WORKING PAPERS

Equilibrium Fast Trading

Sophie Moinas, Bruno Biais and Thierry Foucault

INSTITUT D'ECONOMIE INDUSTRIELLE

Toulouse School of Economics

Equilibrium Fast Trading[1]

Bruno Biais
Toulouse School of Economics
(CNRS-CRM and FBF-IDEI Chair on Investment Banking and Financial Markets)

Thierry Foucault
HEC, Paris

Sophie Moinas
Toulouse School of Economics
(IAE-CRM and FBF-IDEI Chair on Investment Banking and Financial Markets)

February 7, 2013

Abstract

High–speed market connections and information processing improve financial institutions' ability to seize trading opportunities, which raises gains from trade. They also enable fast traders to process information before slow traders, which generates adverse selection. We first analyze trading equilibria for a given level of investment in fast–trading technology and then endogenize this level. Investments can be strategic substitutes or complements. In the latter case, investment waves can arise, where institutions invest in fast–trading technologies just to keep up with the others. When some traders become fast, it increases adverse selection costs for all, i.e., it generates negative externalities. Therefore equilibrium investment can exceed its welfare–maximizing counterpart.

---

[1]Corresponding author: Bruno Biais, Toulouse School of Economics, 21 Allée de Brienne, 31500 Toulouse, France. Tel: 33 5 61 12 85 98. Fax: 33 5 61 12 86 37. email: bruno.biais@univ-tlse1.fr

Equilibrium Fast Trading

February 8, 2013

Abstract

High–speed market connections and information processing improve financial institutions' ability to seize trading opportunities, which raises gains from trade. They also enable fast traders to process information before slow traders, which generates adverse selection. We first analyze trading equilibria for a given level of investment in fast–trading technology and then endogenize this level. Investments can be strategic substitutes or complements. In the latter case, investment waves can arise, where institutions invest in fast–trading technologies just to keep up with the others. When some traders become fast, it increases adverse selection costs for all, i.e., it generates negative externalities. Therefore equilibrium investment can exceed its welfare–maximizing counterpart.

1 Introduction

Investors and traders in financial markets must process very large amounts of information: both fundamental information such as news announcements and information stemming from the markets themselves. For example, Wall Street institutions, trading an NYSE listed stock must closely monitor the dynamics of the order book, order flow and trades in this stock on the Exchange. In addition, they must collect and process information on this stock from other markets, such as, e.g., Nasdaq, Bats or Direct Edge, and from derivatives markets. This is particularly important because markets are now fragmented, and trading occurs on several different market venues (see, e.g., O'Hara and Ye, 2011).

To cope with this huge flow of information and the fragmentation of the markets, financial institutions invest in fast connections to trading venues and high–speed information processing capacities. For example, trading firms can buy the right to place their computers next to the exchange's servers. This practice, known as co-location, reduces to a few milliseconds the latency, i.e., the delay between the time at which a message is sent and that at which it is received. Another example is offered by fiber optic cables, across the Atlantic or between Chicago and New York, enabling their users to get information from, and send orders to, markets around 5 milliseconds before their competitors. Yet another example is the venture between the inter-dealer broker BCG Partners and the high–frequency trading group Tradeworx announced in August 2012 to explore data transmission through the air on microwave radio signals, as the latter travel nearly 50 percent faster than light through optical fiber. Other forms of such investments include the purchase of powerful computers and the development of smart programs collecting and comparing data from several markets and automatically firing orders based on this data.[1]

Investments in fast trading technologies are expensive. For example, the cost of Project Express, which drew a new and faster fiber optic cable across the Atlantic, to connect Wall Street to the City, was $300 million. Against this cost, investments in fast trading technology have both positive and negative consequences for the functioning of markets.

On the one hand, they help traders cope with market fragmentation. Thus, financial

---

[1]This also involves investment in human capital. For example, about a third of the employees of Renaissance Technologies (a hedge fund that is extremely active in fast computerized trading) have Ph.Ds.

institutions can seize trading opportunities before they vanish. This can increase the ability of market participants to reap mutually beneficial gains from trade. On the other hand, thanks to investment in high–speed connections and information processing, fast traders access value relevant information before slow traders. Thus, fast traders' orders are superiorly informed. For example, Hendershott and Riordan (2011) find that the permanent impact of fast orders is larger than that of slow ones, and that fast traders' informational advantage is sufficient to overcome the bid–ask spread. Kirilenko et al. (2011) note that "possibly due to their speed advantage or superior ability to predict price changes, high–frequency traders are able to buy right as the prices are about to increase." While this informational advantage generates profits for fast traders, it also generates losses for slow traders. For example, Baron, Brogaard and Kirilenko (2012) find that aggressive, liquidity–taking, high–frequency traders earn short–term profits at the expense of other market participants. Thus, Brogaard, Hendershott and Riordan (2012) write: *"Our results are consistent with concerns about high–frequency traders imposing adverse selection on other investors".*[2]

These remarks combined with the observation that fast trading has grown significantly raise a number of issues: Does the prevalence of fast traders enhance or deteriorate the functioning of markets? Do market forces lead to an optimal amount of investment in fast trading technologies? Is policy intervention called for?

To examine these issues, and in particular the costs and benefits of fast trading, we consider a simple model, with a continuum of financial institutions, with two potential motivations for trade: i) heterogeneity in private valuations, e.g., due to differences in tax or regulatory status,[3] and ii) differences in private information about common values. While the former creates the scope for gains from trade, the latter generates adverse selection. Initially, financial institutions decide whether to invest in fast trading technology or not, at cost C. Fast institutions then always find a trading counterparty, and, before they do, observe advance information about the common value of the asset. Slow institutions,

---

[2] They also point to the benefits of high–frequency traders with regards to price efficiency. They note, however, that such benefits might be limited, as they find that high–frequency traders' orders predict prices only on very short horizons, of less than 4 seconds.

[3] This is in line with Bessembinder, Hao, and Zheng (2012), where private valuation shocks induce gains from trade and hence transactions between rational agents.

in contrast, do not observe such information, and only find a trading counterparty with probability $\rho < 1$. Our assumption about the probabilities of finding trading counterparties is made to capture the idea that fast trading technologies help investors seize trading opportunities before they vanish.

We first analyze trading equilibrium for a given fraction, α, of fast traders. As α increases, the market impact of trades rises. This effect reduces the expected gains from trade for each type of institution, both fast and slow. It reflects the negative externality fast traders inflict upon the others by increasing adverse selection in the marketplace. In the extreme, as α gets very large, slow traders are evicted from the market, i.e., for them, there is a market breakdown. As a result an increase in the fraction of fast traders can have ambiguous effects on trading volume:[4] On the one hand, it tends to increase volume, as it helps financial institutions seize trading opportunities. On the other hand, it can reduce volume by crowding out slow traders.[5]

Second, we analyze equilibrium investment in fast trading technologies. Financial institutions decide to incur the investment cost if it is not larger than the difference between the expected gains of fast and slow traders. The latter depends on the anticipated fraction on fast traders. Hence, the equilibrium fraction of fast traders is the solution of a fixed point problem. While the expected gains of fast and slow traders both decline with the fraction of fast traders (α), the *difference* between the expected gains of the two types of traders can be decreasing or increasing in this fraction. In the former case, investments in fast trading are strategic substitutes, while in the latter case they can be strategic complements. When investments in fast trading are strategic complements, there can be multiple equilibria. In that case, interior equilibria can coexist with a corner equilibrium at $\alpha = 1$. Because the latter is a stable equilibrium, this raises the possibility of investment waves in fast trading, whereby all financial institutions decide to become fast, to avoid being side–lined by the others. In presence of such strategic complementarities, investments in fast trading technologies have the flavour of an arms' race, reminiscent of the analysis of

---

[4]This is consistent with the empirical findings of Jovanovic and Menkveld (2011).

[5]Our theoretical result that fast trading can induce crowding out of slow traders echoes the concerns raised in the IOSCO consultation report on "Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency" (July 2011, page 10): *"some market participants have also commented that the presence of high–frequency traders discourages them from participating as they feel at an inherent disadvantage to these traders' superior technology."*

Glode, Green and Lowery (2012). When financial institutions expect all the others (or the vast majority of the others) to be fast, they anticipate to be crowded out if they remain slow. To avoid this, they also decide to bear the investment cost necessary to be fast.

Third, we analyze the level of investment maximizing utilitarian welfare, and compare it to its equilibrium counterpart. When financial institutions decide whether to become fast or not, they take into account the private costs and benefits of this decision. They do not, however, account for the above mentioned negative externality their decision inflicts upon the others. Hence, equilibrium investment can be higher than its utilitarian welfare maximizing counterpart, i.e., there can be excessive investment in equilibrium.

Our theoretical analysis has several empirical implications. In the short term (i.e., holding the investment in fast trading technologies fixed), increased volatility, or increased market fragmentation, raise the fraction of trades stemming from fast traders, and hence the informativeness of trades. This increase in fast trading is in line with the finding by O'Hara and Ye (2011) that fragmentation increases execution speed. In the long term, increases in the cost of fast–trading reduce investment in this technology and correspondingly the fraction of trading volume stemming from fast institutions. This reduction lowers the permanent impact of trades and can have ambiguous consequences on trading volume. Also, sustained increases in volatility raise investment in fast trading when α is low, but reduce it when α is high. Finally, increases in market fragmentation adversely affect both fast and slow traders, and thus can have ambiguous consequences on the level of investment in fast trading.

Our theoretical analysis also yields policy–relevant insights. Hirshleifer (1971) studies "foreknowledge," the knowledge of events that will in due time occur and be observable by all. The ability of fast traders to collect and process market information that will be also available to slow traders, but a few seconds or milliseconds later, is a form of foreknowledge. Now, as written by Hirshleifer (1971),

> "the distributive aspect of access to superior information... provides a motivation for the acquisition of private information that is quite apart from any social usefulness of that information... There is an incentive for individuals to expend resources in a socially wasteful way in the generation of such information."

We show that the social costs induced by investment in fast trading technologies can go beyond the waste of the resources thus invested. By inducing adverse selection, and crowding out slow traders, investment in fast trading technologies can actually *decrease* overall gains from trade.

Since our analysis implies that equilibrium investment in fast trading can be excessive, one policy response would be to tax observable investments, such as fiber optic cables or co-location. Setting the level of such taxes, however, would be difficult, since it would require estimating the negative externality imposed by fast trading. An alternative, market based, response would be to offer trading platforms accessible to slow traders only. While such a strategy could be fruitful, it could be hard to implement, due to the difficulty of attracting order flow to new platforms. Yet another market–based response would be to offer the possibility to place, on existing markets, limit orders that would be executable against slow orders only.

The next section discusses the relation between our analysis and previous theoretical literature on information acquisition. The third section presents our model. Section 4 presents the trading equilibrium for a given α. Section 5 studies the equilibrium level of fast trading, while Section 6 studies its welfare maximizing counterpart. Section 7 describes empirical implications of the model and Section 8 concludes.

2   Related theoretical literature

Our analysis is in line with the seminal paper of Grossman and Stiglitz (1980). In both cases, the fraction of informed agents affects the outcome of the trading process and is determined in equilibrium. The two main differences between our model and theirs are the following: First, in our framework, all participants are rational and make optimal decisions, i.e., there are no noise traders. Second, investment in fast trading does not only generate advance information, it also enhances the ability to seize trading opportunities. These differences in modelling approaches yield differences in results. First, as all traders are rational and have well defined preferences, we can perform a welfare analysis of investment in fast trading technologies, comparing equilibrium investment to its socially optimal

counterpart.

Second, while in Grossman and Stiglitz (1980) investments in information acquisition are always strategic substitutes, in our model they can also be strategic complements. This arises because both fast and slow institutions adjust their trading strategies to α. This can lead to situations where the difference between the expected gains of the former and the latter is increasing in α. This raises the possibility of equilibrium multiplicity and investment waves, unlike in Grossman and Stiglitz (1980). In a CARA–Gaussian model, Ganguli and Yang (2009) analyze the case where traders observe imperfect private signals on aggregate supply as well as on common values. Breon–Drish (2010) extends Grossman and Stiglitz (1980) to the non Gaussian case. In both models, complementarity in information acquisition arises when prices become less informative as the number of informed investors increases.[6] This interesting mechanism is completely different from that at play in our model, whereby financial institutions decide to be fast because they anticipate many others to also be fast, and thus fear to be crowded out if they remain slow.

3   Model

Consider a unit mass continuum of risk–neutral, profit maximizing financial institutions, trading in the market for one asset.

Asset valuations: Institutions can buy or sell one share of the asset or abstain from trading. Their valuations for the asset are the sum of a common value component and a private value component. The common value, v, can be equal to $\mu + \epsilon$ or $\mu - \epsilon$ with equal probability. Private values are i.i.d across institutions and can be equal to $\delta > 0$ or $-\delta$ with equal probability. Differences in private values capture in a simple way that other considerations than expected cash–flows affect the willingness of investors to hold assets. For example, regulation can make it costly or attractive for certain investors, such as insurance companies, pension funds, or banks to hold certain asset classes.[7] Differences

---

[6] In contrast, in our model, the informativeness of prices always increases in α.

[7] For instance, by regulatory requirements, some institutional investors can only hold investment grades bonds. Thus, they value these bonds at a premium relative to other investors .

in tax regimes can also induce differences in private values. All institutions optimally decide to trade or not by comparing their valuation for the asset to asset prices, in equilibrium (see below). Hence, there are no noise traders in our setting (trades can be mutually profitable because of differences in private valuations among traders). This feature of the model is important for welfare analysis (all investors' welfare is well defined here). It also implies that all investors optimally adjust their trading strategies when the market structure (e.g., the fraction of fast institutions; see below) changes.

Investment in information collection and processing technology: Before trading occurs, institutions simultaneously decide whether to invest in infrastructures (computers, co-location, etc.) and intellectual capital (skilled traders, codes, etc.) that increase the speed and the accuracy with which institutions receive and process information from markets. The cost of this investment in "fast trading technologies" is denoted by C. We hereafter refer to the institutions who invest in this technology as fast, and to the others as slow. The fraction of fast institutions is denoted by α. It is endogenized in Section 5 where we analyze the decision of each institution to become fast or remain slow. Investment in a fast trading technology helps institutions in two ways.

First, they access and process information on v before slow institutions. To capture this in the simplest possible way, we assume that, before trading, fast institutions observe whether $v = \mu + \epsilon$ or $\mu - \epsilon$. This assumption is consistent with the empirical evidence suggesting fast traders are superiorly informed about short-term changes in valuations. For example, Brogaard, Hendershott and Riordan (2012) find that market orders placed by high–frequency traders forecast very short-run changes in prices, and that, for this reason, high–frequency traders' marketable orders are profitable even after accounting for transaction costs (including the bid–ask spread and trading fees).

Second, fast institutions are more likely to find trading opportunities. Regulations such as the MiFID in Europe or RegNMS in the U.S. led to competition but also fragmentation between trading platforms. As a result, quotes for the same security are posted in various trading venues.[8] Thus, investors have to search for the best price among multiple trading

---

[8]For instance, in May 2011, the three most active competitors of the London Stock Exchange, namely Chi-X, BATS Europe and Turquoise, reached a daily market share in FTSE 100 stocks of 27.5%, 7.4% and 5.2%, respectively while that of the London Stock Exchange was 51%. Source: http://www.ft.com/trading-room.

venues and to compare trading opportunities among several markets. Fast access to market data reduces search costs for best execution prices and help investors locating attractive quotes before they have been hit or withdrawn.[9] To capture this, we assume slow institutions are less likely to find a trading opportunity than fast institutions. Namely, slow institutions find a trading counterparty with probability $\rho < 1$, while fast institutions find it with probability 1.

Trading: When an institution gets a trading opportunity, it decides whether to buy one share ($\omega = 1$), sell one share ($\omega = -1$), or abstain from trading ($\omega = 0$). We assume that the trade of the institution is executed at a price equal to the expectation of the common value conditional on its order: $E(v|\omega)$, computed with rational expectations about all equilibrium strategies, as in Glosten and Milgrom (1985) and Easley and O'Hara (1987).

A literal interpretation of our trading process, in which the fast traders would exploit their private information via market orders, is in line with empirical findings. For example, Brogaard, Hendershott and Riordan (2012) find that high–frequency traders trade in the direction of permanent price changes through market orders, while their limit orders are adversely selected. Similarly, Baron, Brogaard and Kirilenko (2012) find that most of the high–frequency traders' profits are generated by aggressive, liquidity–taking, trades.

However, our conclusions apply more generally, whether institutions trade with market or limit orders. Indeed, the driving force behind our model is the information asymmetry between fast and slow traders, which hurt the latter. This effect will be present independently of the type of orders used by fast traders. The assumption that all trades take place at $E(v|\omega)$ is the simplest specification of the outcome of the trading process that is consistent with rationality and participation constraints.[10] This simple specification is assumed for tractability.

Timing: The timing of decisions in the model is as follows:

At $\tau = 0$, each institution decides whether to pay C, and become fast, or not.

At $\tau = 1$, each institution observes its private valuation δ or $-\delta$, and, if it is fast,

---

[9]In line with this hypothesis, Garvey and Wu (2010) find that traders who get quicker access to the NYSE because of their geographical proximity pay smaller average effective spreads.

[10]Participation constraints hold since the trading counterparties of the institutions we consider earn non–negative profits on average. One could introduce a transaction cost incurred by the institutions we analyze, without qualitatively affecting the results.

observes the realization of v: $\mu + \epsilon$ or $\mu - \epsilon$. Then each institution receives a trading opportunity or not and, if it does, optimally chooses $\omega = 1$, $\omega = -1$, or $\omega = 0$. Finally, the institution's trade, ω, is executed at price $E(v|\omega)$.

At $\tau = 2$, v is realized.

We solve the model starting from date $\tau = 1$, for a fixed value of α (Section 4). After characterizing the optimal trade and prices at date $\tau = 1$, we deduce the expected profits of slow and fast institutions for each value of α. We can then endogenize the equilibrium fraction of institutions that decide to become fast (Section 5).

4   Trading with fast and slow investors

This section analyzes equilibrium transaction prices and trading volume at date $\tau = 1$, for a given level of α. As a benchmark, it is useful to first discuss the case in which all institutions are slow ($\alpha = 0$). In this case, their orders do not convey any information and are therefore executed at price μ. Institutions with a high private valuation buy, while those with a low valuation sell. Trading volume (denoted by *Vol*) is equal to the fraction of institutions that find a counterparty, ρ.

When $\alpha > 0$, the analysis is more complex. At date $\tau = 1$, after institutions observe their valuations for the asset and before trading, there are six types of institutions: (i) fast institutions with good news and high private valuations (which we denote by GH), (ii) fast institutions with good news and low private valuations (GL), (iii) fast institutions with bad news and high private valuations (BH), (iv) fast institutions with bad news and low private valuations (BL), (v) slow institutions with high private valuation (H), and (vi) slow institutions with low private valuation (L). For brevity, we focus the analysis on the case where an institution decides to buy the asset or abstain from trading. The corresponding price is denoted by a. The case of sales (at price b) is symmetric, e.g., the markup at which institutions buy ($a - \mu$) is equal to the discount at which they sell ($\mu - b$).

As will be clear below, equilibria can involve pure or mixed strategies for the trading decision taken by an institution given his type. Hence, we denote by β_j^F the (endogenous) probability that, conditional on finding a trading opportunity, a fast institution

$j \in \{GH, GL, BH, BL\}$ buys. Similarly, β_j^S is the probability that, conditional on finding a counterparty, a slow institution $j \in \{H, L\}$ buys. As transaction prices are given by $E(v|\omega)$, they cannot exceed $\mu + \epsilon$ or be smaller than $\mu - \epsilon$. Hence, fast institutions with good news on v and high private valuation always buy, i.e., $\beta_{GH}^F = 1$. Symmetrically, fast institutions with bad news and low private valuation never buy, i.e., $\beta_{BL}^F = 0$. Applying Bayes law, one obtains the following lemma.

Lemma 1 *Buy orders execute at price*

$$a = E(v|\omega = 1) = \mu + \frac{\alpha(1 + \beta_{GL}^F - \beta_{BH}^F)}{2\left((1 - \alpha)\rho(\beta_H^S + \beta_L^S) + \frac{\alpha}{2}(1 + \beta_{GL}^F + \beta_{BH}^F)\right)}\epsilon. \tag{1}$$

First suppose that $\delta > \epsilon$. In this case the valuation of the asset by a fast institution with a low private valuation and good news, that is, $\mu + \epsilon - \delta$, is smaller than a since $a \geq \mu$. Thus, institutions with low private valuations never buy the asset, whether fast or slow, i.e., $\beta_{GL}^F = \beta_L^S = 0$. In contrast, slow institutions with a high private valuation always buy the asset since $a \leq \mu + \epsilon < \mu + \delta$. Thus, $\beta_H^S = 1$ when $\delta > \epsilon$. The behavior of fast institutions with a high private valuation and bad news depends on a. However, if $a = \mu$, buying is optimal for these institutions. The next proposition states that there is always an equilibrium of this type when $\delta > \epsilon$.[11]

Proposition 1 *When $\delta > \epsilon$ there exists a trading equilibrium in which: (i) all institutions buy (resp. sell) if and only if they have a high (resp. low) private valuation, (ii) all trades take place at price μ, and (iii) trading volume equals $\alpha + (1 - \alpha)\rho$.*

When $\delta > \epsilon$, news about v are small relative to private valuation shocks. Hence, there is an equilibrium in which prices and allocations are identical to those that prevail when

---

[11]The proof of the proposition is straightforward, and skipped for brevity. The equilibrium is not necessarily unique when $\delta > \epsilon$, as will be the case when $\delta < \epsilon$. For instance, suppose that a fast institution with bad news and a high private valuation does not buy the asset ($\beta_{BH}^F = 0$). In this case, $a = \mu + \frac{\alpha}{2(1-\alpha)\rho+\alpha}\epsilon$ since $\beta_H^S = 1$ (see equation (1)). This price is greater than the valuation of a fast institution with a high private valuation if $\left(\frac{2(\alpha+(1-\alpha)\rho)}{\alpha+2(1-\alpha)\rho}\right)\epsilon > \delta$, so that not buying is indeed optimal for this fast institution. The possibility of multiple equilibria for prices and allocations at $\tau = 1$ is in line with the analyses of Glosten and Milgrom (1985) and Dow (2005). They underscore the possibility of virtuous circles (traders anticipate the market will be liquid, hence they submit lots of orders, hence the market is liquid) or vicious circles (where illiquidity is a self–fulfilling prophecy). We have checked that the equilibrium in Proposition 1 Pareto dominates other equilibria with low liquidity (that is, with $a > \mu$).

all institutions are slow, that is, when $\alpha = 0$. However, trading volume is higher because some institutions are more likely to find a counterparty. Therefore, if $C = 0$, $\alpha > 0$ Pareto dominates the benchmark case, $\alpha = 0$.

In the rest of the paper, we assume $\delta < \epsilon$, that is, possible changes in the common value, v, are larger than traders' private valuation shocks. Hence, fast institutions' trades are more driven by their information than by their private valuation. Thus, adverse selection problems are more severe than when $\delta > \epsilon$. To simplify the analysis and reduce the number of possible cases that can arise in equilibrium, we hereafter assume

$$\frac{\epsilon}{2} < \delta < \epsilon. \tag{2}$$

That is, the volatility of the common value is higher than the dispersion in private valuations ($\delta < \epsilon$), but the latter is still significant ($\frac{\epsilon}{2} < \delta$).[12] Equation (2) implies

$$\mu < \mu + \epsilon - \delta < \mu + \delta < \mu + \epsilon < \mu + \epsilon + \delta. \tag{3}$$

The first term from the left is the unconditional expectation of v. The second one is the valuation of the security for a fast investor with good news but a low private valuation for the asset. The third term is the valuation of the security for a slow investor with positive private value. The fourth term is the common value of the security given good news on the fundamental. The fifth and last term is the valuation of the security for a fast investor with good news and positive private value. The ranking of institutions' possible valuations for the asset in equation (3) implies there are five possible types of equilibria, corresponding to increasingly high ask prices. In all candidate equilibria, $\beta_{GH}^F = 1$ and $\beta_{BL}^F = \beta_L^S = 0$, as already explained. Furthermore, $\beta_{BH}^F = 0$ since $a > \mu > \mu - \epsilon + \delta$. We denote the five possible types of equilibria by P1, M1, P2, M2 and P3, respectively and we outline their characteristics below:

- **P1:** If $\mu \leq a < \mu + \epsilon - \delta$, fast institutions with good news buy, whatever their private valuation, while slow institutions buy if and only if their private valuation is high.

---

[12]We have checked that when $\frac{\epsilon}{2} \geq \delta$ the qualitative results are similar to those derived under condition (2).

Hence, $\beta^F_{GL} = 1$ and $\beta^S_H = 1$.

- **M1:** If $a = \mu + \epsilon - \delta$, fast institutions with good news and high private valuation buy. So do slow institutions with high private valuation, i.e., $\beta^S_H = 1$. Fast institutions with good news but low private valuation are indifferent between buying and not trading. They play mixed strategies, buying with probability $0 \leq \beta^F_{GL} \leq 1$.

- **P2:** If $\mu + \epsilon - \delta < a < \mu + \delta$, fast institutions buy if they have good news and high private valuation, but they do not trade if their private valuation and their information on v conflict, i.e., $\beta^F_{GL} = 0$. Slow institutions with high private valuation buy, i.e., $\beta^S_H = 1$.

- **M2:** If $a = \mu + \delta$, fast institutions with good news and high private valuation buy, but they do not trade if their private valuation and their information on v conflict, i.e., $\beta^F_{GL} = 0$. Slow institutions with a high private valuation are indifferent between buying or not trading. They play a mixed strategy, buying with probability $0 \leq \beta^S_H \leq 1$.

- **P3:** If $a = \mu + \epsilon$, fast institutions with good news and high private valuation buy. Other types choose not to buy. Hence, $\beta^F_{GL} = 0$, and $\beta^S_H = 0$.

Equilibrium P3 features "crowding out" as slow institutions are sidelined and only fast institutions trade. In this equilibrium, only a small fraction of the gains from trade can be reaped, especially if there are few fast institutions (α small) since slow institutions choose not to trade. Unfortunately, such equilibrium can be pervasive. Suppose it is expected that only fast institutions with good news and a high private valuation buy. Correspondingly, $a = \mu + \epsilon$. As a result, slow institutions choose not to trade since $\delta < \epsilon$ and so do fast institutions whose private value and signal on v conflict. Thus, the expectation that only fast institutions with high private valuation and good news will buy is self-fulfilling for all parameter values. This yields our next proposition.[13]

---

[13] The proposition follows directly from the arguments we just gave. Hence, for brevity, its proof is omitted.

Proposition 2 *When $\delta < \epsilon$, there always exists a crowding out equilibrium (P3).*

The equilibrium with crowding out is not the only possibility, however. To spell out the conditions under which other equilibria than P3 exist, denote:

$$\alpha_1 = \frac{\rho(\epsilon - \delta)}{\rho(\epsilon - \delta) + \delta}, \quad \alpha_2 = \frac{\rho(\epsilon - \delta)}{\rho(\epsilon - \delta) + \frac{\delta}{2}}, \alpha_3 = \frac{2\rho\delta}{2\rho\delta + \epsilon - \delta}.$$

Using these notations, and observing that $\alpha_1 < \alpha_2 < \alpha_3$, we state our next proposition.

Proposition 3 *When $\delta < \epsilon$,*

1. *If $0 < \alpha \leq \alpha_3$, there exists an M2 equilibrium, in which $\beta_H^S = \frac{\alpha}{2(1-\alpha)\rho\delta}(\epsilon - \delta)$.*

2. *If $\alpha < \alpha_1$, there exists a P1 equilibrium, in which $a = \mu + \frac{\alpha}{\alpha + (1-\alpha)\rho}\epsilon$.*

3. *If $\alpha_1 \leq \alpha \leq \alpha_2$, there exists an M1 equilibrium, in which $\beta_{GL}^F = \frac{2(1-\alpha)\rho}{\alpha\delta}(\epsilon - \delta) - 1$.*

4. *If $\alpha_2 < \alpha < \alpha_3$, there exists a P2 equilibrium, in which $a = \mu + \frac{\alpha}{\alpha + 2(1-\alpha)\rho}\epsilon$.*

Figure 1 shows the equilibria that can be obtained for each value of α. It highlights that there are three equilibria when $0 < \alpha < \alpha_3$. However, as claimed in the next lemma, the equilibria with low trading volume (that is, P3 and M2) are Pareto dominated by the others (P1, M1, or P2 depending on the value of α).

Lemma 2 *For each value of α, there is a unique Pareto dominant trading equilibrium: P1 when $0 \leq \alpha < \alpha_1$, M1 when $\alpha_1 \leq \alpha \leq \alpha_2$, P2 when $\alpha_2 < \alpha < \alpha_3$, M2 when $\alpha = \alpha_3$, and P3 when $\alpha > \alpha_3$.*

Hereafter, for each value of α, we focus on the Pareto dominant equilibrium. Figure 2 shows the evolution of the price impact of buy orders, $a - \mu$, as a function of α. In our framework, this impact is a measure of the informativeness of trades since trades take place at $E(v|\omega)$. It weakly increases in α because: (i) the fraction of investors with information

14

increases and (ii) slow institutions trade less when α is large. Thus, the model implies that the informational impact of trades should increase with the fraction of fast institutions.[14] Our next corollary states the effect of α on trading volume per institution.[15]

Corollary 1 *Equilibrium trading volume is* $Vol(\alpha) = \alpha + (1 - \alpha)\rho$ *for* $0 \leq \alpha < \alpha_1$, $Vol(\alpha) = (1 - \alpha)\rho\epsilon/\delta$ *for* $\alpha_1 \leq \alpha \leq \alpha_2$, $Vol(\alpha) = \alpha/2 + (1 - \alpha)\rho$ *for* $\alpha_2 < \alpha \leq \alpha_3$, *and* $Vol(\alpha) = \alpha/2$ *for* $\alpha > \alpha_3$. *Hence, volume is non monotonic in* α.

An increase in α has two effects. On the one hand, more institutions find a trading opportunity. On the other hand, the cost of exploiting this opportunity increases as the fraction of fast institutions increases because the price impact of trades increases with α, due to greater adverse selection. As a result, some institutions may choose not to trade even if they find a trading opportunity. More specifically, fast institutions trade less or do not trade when their signal and private valuations conflict when $\alpha > \alpha_1$, and slow institutions stop trading when $\alpha > \alpha_3$. The first effect tends to increase volume while the second tends to decrease volume.

Thus, the net effect of an increase in the fraction of fast institutions on the volume of trades is ambiguous and, for this reason, trading volume is not monotonic in α. As Figure 3 shows, trading volume increases in α when α is very low or sufficiently high. Indeed, in these cases, the positive effect of fast trading on the likelihood of finding a trading opportunity dominates the price impact effect. In contrast, for intermediate values of α, trading volume can decrease in α.[16] In addition, there is always a discrete drop in

---

[14] At first glance, the empirical findings in Hendershott, Jones and Menkveld (2011) do not support this implication of our model. They find that the informational impact of trades has declined on the NYSE after a change in market structure that made algorithmic trading easier on the NYSE. However, it is not clear whether Hendershott, Jones and Menkveld (2011)'s proxy for algorithmic trading captures the effect of fast trading rather than other forms of algorithmic trading. The market structure change considered in Hendershott, Jones and Menkveld (2011) may have helped slow traders to better find trading opportunities (an increase in ρ). Our model predicts that for a fixed value of α, the informational impact of trades declines in ρ (see Section 7).

[15] In our set-up, the trading volume per institution is the likelihood that an institution buys multiplied by two (since each institution is equally likely to buy or sell, ex-ante).

[16] Specifically, when $\rho > 1/2$, a small increase in the fraction of fast institutions increases the trading volume when $\alpha < \alpha_1$ or $\alpha > \alpha_3$ and it decreases trading volume when $\alpha_1 \leq \alpha \leq \alpha_3$. When $\rho \leq 1/2$, a small increase in the fraction of fast institutions increases the trading volume when $\alpha < \alpha_1$ or $\alpha > \alpha_2$ and it decreases trading volume when $\alpha_1 \leq \alpha \leq \alpha_2$.

trading when α increases beyond α_3, due to the fact that at this point slow institutions stop trading.[17] The possibility of a negative effect of fast trading on the volume of trade is in line with Jovanovic and Menkveld (2011), who find that for Dutch stocks the entry of a fast trader on Chi-X led to a drop in volume.[18]

5   Equilibrium investment in fast trading technologies

In the previous section, we have derived equilibrium prices and allocations for a fixed value of α. We now study the equilibrium determination of α. To do so, we first compare the gains of fast and slow institutions. Then we study the interdependence of investment decisions among financial institutions. We finally define and characterize equilibrium investment in fast trading technologies.

Comparing the gains of fast and slow institutions: Denote the expected profits of slow and fast institutions at date $\tau = 0$, gross of the investment cost C, by $\psi(\alpha)$ and $\phi(\alpha)$, respectively. Using Proposition 3, we obtain the following corollary.

Corollary 2 *At date $\tau = 0$, the gross expected profits of slow and fast institutions are $\psi(\alpha) = (\delta - \frac{\alpha}{\alpha+(1-\alpha)\rho}\epsilon)\rho$ and $\phi(\alpha) = \frac{(1-\alpha)\rho}{\alpha+(1-\alpha)\rho}\epsilon$ for $0 \leq \alpha < \alpha_1$, $\psi(\alpha) = (2\delta - \epsilon)\rho$ and $\phi(\alpha) = \delta$ for $\alpha_1 \leq \alpha \leq \alpha_2$, $\psi(\alpha) = (\delta - \frac{\alpha/2}{\alpha/2+(1-\alpha)\rho}\epsilon)\rho$ and $\phi(\alpha) = \frac{1}{2}(\delta + \frac{(1-\alpha)\rho}{\alpha/2+(1-\alpha)\rho}\epsilon)$ for $\alpha_2 < \alpha \leq \alpha_3$, and $\psi(\alpha) = 0$ and $\phi(\alpha) = \delta/2$ for $\alpha > \alpha_3$. Hence, the expected profits of fast and slow institutions decrease in α.*

Figure 4 illustrates that the expected gains of slow *and* fast institutions decline with α. As α increases, institutions buy at a higher markup or sell at more discounted prices and institutions realize less gains from trade. For instance, fast institutions with low private valuations but good news trade less or stop trading when $\alpha > \alpha_1$ because their impact on prices is too high. Similarly, slow institutions pull out from the market when $\alpha > \alpha_3$.

---

[17] A small increase in α at $\alpha = \alpha_3$ implies that trading volume drops from $\frac{\alpha_3}{2} + (1-\alpha_3)\rho$ to $\frac{\alpha_3}{2}$.

[18] Anecdotal evidence also suggests that, as high–speed trading expands, trading volume can increase or decrease. For example, an article entitled *"Electronic trading slowdown alert"* published in the Financial Times on September 24, 2010 (page 14) describes a sharp drop in trading volume in 2010 from a high of about \$7,000 billions in April 2010 to a low of \$4,000 billions in August 2010. The article explicitly points to changes in market structures as a cause for this reversal in trading volume.

For these reasons, the entry of a new fast institution reduces the gains of all the other institutions, i.e., it exerts a *negative externality*. Fast institutions, however, always get greater expected gains than slow ones because (i) they are more likely to find a trading opportunity, and (ii) they profit from their advance information.

Strategic substitutability or complementarity: For a given level of α, the expected gain of an institution incurring the investment cost C is: $\phi(\alpha) - C$, while if it does not invest, its expected gain is: $\psi(\alpha)$. Thus, an institution is better off investing if and only if: $\phi(\alpha) - \psi(\alpha) \geq C$. As ϕ and ψ vary with α, the profitability of investment for one institution depends on the investment decisions of the other institutions. Thus, investment choices are interdependent. As mentioned above, an increase in α hurts both fast and slow institutions. However, what matters for the decision to invest is the *difference* between the profits of fast institutions and slow institutions. If $\phi - \psi$ decreases in α, then fast institutions loose more than slow ones when α goes up. In that case, institutions' investment decisions in fast trading technologies are strategic substitutes: the greater the fraction of institutions that invest, the less profitable it is to invest. In contrast, if $\phi - \psi$ increases in α, slow institutions are hurt more than fast ones by an increase in α and institutions' investment decisions are then strategic complements: the greater the fraction of fast institutions, the more profitable it is to invest in fast trading technologies. In that case, the institutions' decisions to invest are mutually reinforcing. The next proposition describes the behavior of $\phi - \psi$, and states the condition under which investment decisions are (locally) strategic substitutes or complements.

Proposition 4 *The function $\phi - \psi$ is continuous in α except at α_3, where it jumps downward. When $0 \leq \alpha < \alpha_1$,*

$$\frac{\partial(\phi(\alpha) - \psi(\alpha))}{\partial \alpha} < 0,$$

i.e., institutions' investment decisions in fast technologies are local strategic substitutes. If $\alpha_1 \leq \alpha \leq \alpha_2$ or $\alpha > \alpha_3$, $\phi(\alpha) - \psi(\alpha)$ is constant with α. Finally, when $\alpha_2 < \alpha \leq \alpha_3$,

$$\frac{\partial(\phi(\alpha) - \psi(\alpha))}{\partial \alpha} > 0 \ \text{if} \ \rho > \frac{1}{2}, \ \text{and} \ \frac{\partial(\phi(\alpha) - \psi(\alpha))}{\partial \alpha} < 0 \ \text{if} \ \rho \leq \frac{1}{2}.$$

17

Thus, institutions' investment decisions in fast technologies are local strategic complements if $\rho > \frac{1}{2}$ and local strategic substitutes otherwise.

Expected gains from trade are equal to the probability of a transaction multiplied by the gains reaped in that case. An increase in the fraction of fast institutions increases the markup at which institutions buy $(a - \mu)$ and the discount at which they sell $(\mu - b)$. This effect explains in part why institutions' expected profits decline with α, whether they are fast or slow. However, the adverse change in markups per trade is costlier for institutions that trade more. That is, the negative effect of an increase in α on expected profits is relatively higher for institutions that are more likely to trade. It turns out that whether fast or slow institutions trade more depends on which trading equilibrium prevails, which explains why the effect of α on $\phi(\alpha) - \psi(\alpha)$ is ambiguous.

When $0 \leq \alpha < \alpha_1$, *P1* prevails. In that case slow institutions buy if and only if they have located a trading opportunity and their private valuation is high, which occurs with probability $\frac{\rho}{2}$. In contrast, fast traders buy if and only if they have observed good news about the common value, which occurs with probability $\frac{1}{2}$. Thus, slow investors trade less often than fast ones, and are therefore less affected by an increase in α. In this case, investment decisions are local strategic substitutes: an increase in the fraction of fast traders affect (negatively) more fast traders than slow traders because the former are more likely to trade.

In contrast, when $\alpha_2 < \alpha < \alpha_3$, *P2* prevails. While slow institutions still buy with probability $\frac{\rho}{2}$, fast institutions buy only if they have good news and a high private valuation, which occurs with probability $1/4$. Hence, when $\rho > \frac{1}{2}$ slow investors trade more often than fast ones, and are therefore relatively more affected by the increase in α. In this case investment decisions are local strategic complements: an increase in α increases the benefit of becoming a fast trader because those who remain slow are hurt relatively more than those who are fast.

Given that $\phi - \psi$ is discontinuous at α_3, in addition to local complementarity, we must consider global complementarity, comparing the benefit of becoming fast, $\phi - \psi$, for non–adjacent values of α_3. For $\alpha > \alpha_3$, slow traders are crowded out, hence ψ is equal to 0, and correspondingly the benefit of becoming fast is large. Thus it can be that $\phi - \psi$ is

smaller for intermediate values of α than for $\alpha > \alpha_3$. We will return to this form of global complementarity below and study its impact on the equilibrium decision to invest in fast trading technologies.

Equilibrium definition: Building on this analysis, we now study the equilibrium determination of the fraction of fast institutions. First, consider corner equilibria. If

$$\phi(1) - \psi(1) > C, \tag{4}$$

then institutions prefer to invest when they expect all the others to do so. Hence, $\alpha^* = 1$ is an equilibrium if Condition (4) holds. Symmetrically, if

$$\phi(0) - \psi(0) < C, \tag{5}$$

then institutions prefer not to invest when they expect the others also won't. Hence, $\alpha^* = 0$ is an equilibrium if Condition (5) holds. Next, turn to interior equilibria. The fraction α^* of fast institution is an interior equilibrium if, when institutions expect that a fraction α^* of institutions will invest, they are indifferent between investing and not investing, that is,

$$\phi(\alpha^*) - \psi(\alpha^*) = C. \tag{6}$$

Thus, an equilibrium fraction of fast institutions exists if Condition (4) or (5) or (6) holds. Since $\phi(\alpha) - \psi(\alpha)$, is discontinuous in α at α_3, it may be that none of these conditions hold. This technical problem however is an artifact of the assumption that we work with a continuum of institutions, making decisions simultaneously. To guarantee existence of an equilibrium, suppose that, instead of a continuum, there are N institutions. Their arbitrary ranking: $i \in \{1, 2, ...N\}$ determines the order in which they sequentially decide whether to invest or not. Denote by $k \in \{0, 1, ...N\}$ the number of institutions deciding to invest. The equilibrium number of fast investors is 0 if (5) holds, N if (4) holds, and otherwise k^* s.t. $\phi(\frac{k}{N}) - \psi(\frac{k}{N}) \geq C, \forall k \leq k^*$ and $\phi(\frac{k}{N}) - \psi(\frac{k}{N}) < C, \forall k > k^*$. Taking the limit, as N goes to infinity, either equilibrium is pinned down by (4), (5), or (6), or α_3 is the equilibrium.

19

Equilibrium characterization: The analysis differs according to whether ρ is smaller than $1/2$ or larger than $1/2$. In the former case, institutions' investment decisions are strategic substitutes and therefore the equilibrium fraction of fast institutions is unique. In the latter case, investment choices can be strategic complements so that there can be multiple equilibrium investment levels.

Proposition 5 *If $\rho \leq \frac{1}{2}$, then the equilibrium fraction of fast institutions is generically unique and as follows:*

1. *If $C \geq \epsilon - \delta\rho$, the equilibrium fraction of fast institutions is 0.*

2. *If $(1 - 2\rho)\delta + \epsilon\rho < C < \epsilon - \delta\rho$, there exists an interior equilibrium in $(0, \alpha_1)$.*

3. *If $\frac{\epsilon}{2} < C < (1 - 2\rho)\delta + \epsilon\rho$, there exists an interior equilibrium in $[\alpha_2, \alpha_3)$.*

4. *If $\frac{\delta}{2} < C \leq \frac{\epsilon}{2}$, the equilibrium fraction of fast institutions is α_3.*

5. *If $C < \frac{\delta}{2}$, the equilibrium fraction of fast institutions is 1.*

Figure 5 illustrates the determination of the equilibrium fraction of fast traders when $\rho \leq \frac{1}{2}$. This fraction is obtained at the intersection of i) the horizontal line that gives the value of C and ii) the downward sloping curve $\phi(\alpha) - \psi(\alpha)$. There are ranges of values for α for which $\phi(\alpha) - \psi(\alpha)$ is constant. When C is equal to this constant (namely, if $C = (1 - 2\rho)\delta + \epsilon\rho$ or $C = \frac{\delta}{2}$), then there are multiple possible values of α. This situation however is not generic since it occurs for only two specific values of C. Now, we consider the case in which $\rho > \frac{1}{2}$.[19]

Proposition 6 *If $\rho > \frac{1}{2}$ then the equilibrium fraction of fast institutions is as follows:*

1. *If $C > \max[\frac{\epsilon}{2}, \epsilon - \delta\rho]$, equilibrium is unique and equal to 0.*

2. *If $\frac{\epsilon}{2} < C < \epsilon - \delta\rho$, equilibrium is unique and in $(0, \alpha_1]$.*

---

[19]In stating Proposition 6, we ignore cut-off values for C (e.g., $C = \frac{\delta}{2}$ or $(1 - 2\rho)\delta + \epsilon\rho$) that delineate the various cases considered in the proposition. This simplifies the presentation without affecting the conclusions since these cases are not generic.

3. If $\max[\frac{\delta}{2}, \delta(1-2\rho)+\epsilon\rho] < C < \frac{\epsilon}{2}$, there are two equilibria, one in $[0, \alpha_1)$, the other in (α_2, α_3).

4. If $\delta(1-2\rho)+\epsilon\rho < C < \frac{\delta}{2}$, there are three equilibria, one in $[0, \alpha_1)$, the second in (α_2, α_3) and the third equal to 1.

5. If $\frac{\delta}{2} < C < \delta(1-2\rho)+\epsilon\rho$, equilibrium is unique and equal to α_3.

6. If $C < \min[\frac{\delta}{2}, \delta(1-2\rho)+\epsilon\rho]$, equilibrium is unique and equal to 1.

In contrast with the case $\rho \leq \frac{1}{2}$, there can be multiple equilibria, as illustrated in Figure 6. Multiplicity can arise because $\phi - \psi$ can increase with α, i.e., because of the strategic complementarity in institutions' investment decisions. In this context, when one institution expects many others to be fast, it perceives that it is costly to remain slow and it is likely to choose to be fast. Hence, investment in fast trading can be contagious, which implies the possibility of investments waves.

This is particularly striking when $\epsilon - \delta\rho \leq C < \frac{\delta}{2}$. Indeed, this condition means that $\phi(0) - \psi(0) \leq C < \phi(1) - \psi(1)$. Thus, no institution finds it optimal to invest if each expects the others not to invest. Yet all institutions find it optimal to invest if they expect the others to invest. The reason is that slow institutions anticipate being crowded out if their competitors invest in the fast trading technology. Hence, the loss in profits of remaining slow is high and they are better off investing as well. This self-reinforcing mechanism for institutions' investment decisions can be interpreted as an arms' race, which is reminiscent of the situation analyzed in Glode, Green and Lowery (2012).

Equilibrium stability: To examine if investment waves are likely to occur, we now analyze if they correspond to stable equilibria. To do so, consider the following tâtonnement process (in line with the analysis of adaptive dynamics in Manzano and Vives, 2012). Start from any value of α. If, at this point, $\phi(\alpha) - \psi(\alpha) < C$, reduce α, while if $\phi(\alpha) - \psi(\alpha) > C$, raise α. Iterate this process until (6) holds, or α reaches 0, 1 or α_3. Inspecting Figure 6, one can see that the corner equilibria ($\alpha = 0$ or 1) are stable, as well as the interior equilibrium when it is in $(0, \alpha_1)$.[20] Equilibria in (α_2, α_3), however, are not stable. Hence, when they

---

[20]The equilibrium with $\alpha = 1$ arises from complementarity in information acquisition decisions by institutions when there are multiple equilibria. The fact that this equilibrium is stable differs from the

coexist with an equilibrium where all institutions become fast, the latter is likely to prevail. To see this, consider Panel 2 in Figure 6 and suppose that the equilibrium fraction of fast institution is in (α_2, α_3). As this equilibrium is not stable, a small increase in the fraction of fast institutions at this point triggers a domino effect that leads all institutions to be fast.

Equilibrium and investment costs: Using Propositions 5 and 6, Figure 6 shows the relationship between the costs of investing in the fast trading technology and the equilibrium level of investment in this technology. Panel 1 of Figure 6 depicts the case where investments are strategic substitutes, as $\rho \leq 1/2$. In this case, consistent with a priori intuition, equilibrium investment is decreasing in investment cost. Panel 2 depicts the link between equilibrium and investment when $\rho > 1/2$. In that case, the possibility of strategic complementarity in investment decisions generates patterns that could seem, a priori, less intuitive. The equilibrium level of investment can be increasing in C. This possibility, however, arises only for unstable equilibria. Nevertheless, as long as C remains smaller than $\delta/2$, there exists one equilibrium where α stays at 1 while C increase. In those equilibria, in spite of the increase in C, financial institutions continue to invest in fast trading technology, because they anticipate the others to do and fear to be side lined.

6   Social Optimum

The analysis in the previous section suggests that the expectation of large investment in fast trading is self–fulfilling, which can generate investment waves. In this section we study whether such investment waves maximize utilitarian welfare. The latter is equal to:

$$W(\alpha) = \alpha \left(\phi(\alpha) - C \right) + (1 - \alpha)\psi(\alpha)(\alpha). \tag{7}$$

Relying on Corollary 2 and equation (7), the following proposition obtains:

Proposition 7 *If*

$$C < (\frac{1}{2} - \rho)\delta, \tag{8}$$

_____

result in Manzano and Vives (2012), who find that only equilibria in which traders' actions are strategic substitutes are stable.

then utilitarian welfare is maximized when $\alpha = 1$. Otherwise utilitarian welfare is maximized for $\alpha = 0$.[21]

As stated in Proposition 7 and illustrated in Figure 7, utilitarian welfare is maximized for $\alpha = 0$ or 1. That such corner solutions emerge is due to the linearity of the technologies we assume, in particular with respect to investment costs. Which of the two corners is optimal depends on Condition (8). The proposition emphasizes that the level of investment maximizing utilitarian welfare depends on the investment cost C, the probability that slow institutions find trading opportunities ρ, and the gains from trade in case of match δ. The social benefit of fast trading is that it raises the probability that institutions find an opportunity from ρ to 1. When ρ is larger than $\frac{1}{2}$, this increase is not very valuable and utilitarian welfare maximization dictates $\alpha = 0$, for all values of C, including $C = 0$. That is, the negative externalities imposed by fast trading are so large that the welfare gains of institutions that become fast is always more than offset by the resulting welfare losses of slow institutions, so that utilitarian welfare declines with α, even before accounting for the cost of investing in fast trading technologies.

In contrast, when ρ is smaller than $\frac{1}{2}$, investment in fast trading is always socially optimal when $C = 0$. Intuitively, fast institutions' ability to find new trading opportunities generates a welfare gain that more than offsets welfare losses due to larger price impacts. Of course, when C increases, aggregate welfare decreases and if C is high enough relative to the improvement in gains from trade when the fraction of fast traders move from zero to one (that is, $(\frac{1}{2} - \rho)\delta$), then no investment in fast trading technologies becomes socially optimal.

Figure 7 shows the equilibrium level of investment in the fast trading technology (red curve) and the socially optimal level of investment (blue dashed curve). Clearly, there exists a wide range of values for C for which the equilibrium level of investment in the fast trading technology is excessive relative to the social optimum. We formalize this observation in the corollary.

---

[21] When $C = (\frac{1}{2} - \rho)\delta$, utilitarian welfare is identical when $\alpha = 0$ and $\alpha = 1$. Thus, either no investment or maximal investment in the fast trading technology is socially optimal.

Corollary 3 *There exists a stable equilibrium with overinvestment if and only if $(\frac{1}{2} - \rho)\delta \leq C < \max[\frac{\delta}{2}, \epsilon - \delta\rho]$. Otherwise the equilibrium level of investment in the fast trading technology coincides with the socially optimal level of investment.*

When C is very low (smaller than $(\frac{1}{2} - \rho)\delta$), the socially optimal level of investment in the fast trading technology is maximal and equal to one (provided that $\rho \leq \frac{1}{2}$). This is also the level of investment that is achieved when institutions' decisions are decentralized. Similarly, when C is very high (above $\max[\frac{\delta}{2}, \epsilon - \delta\rho]$) the equilibrium and socially optimal levels of investment also coincide and are equal to zero (for all values of ρ). For intermediate values of C, however, there is a discrepancy between the equilibrium level of investment in the fast trading technology and the social optimum. Investing in fast trading is not socially optimal, either because C is high, or because the benefit of an increase in trading opportunities is too low. Yet, C is not high enough to deter all institutions from investing. Hence, overinvestment arises because, when making their investment decisions, institutions take only their own profits into account, and fail to internalize the negative externality they impose on the others. In contrast, there is no underinvestment in our framework, because investment in fast trading technologies does not generate any positive externality.

7   Empirical implications

Three parameters in our model affect fast trading and its consequences: i) ϵ, which measures the volatility of the asset common value, and could be proxied by weekly returns' volatility,[22] ii) ρ, which decreases as markets get more fragmented, and iii) C, the cost of investing in fast trading technologies, which varies with technological progress, taxes, regulatory costs, and exchanges' co-location or data fees.

In the short run, trading technology, and correspondingly the fraction α of fast institutions, is fixed. Yet, volatility or market fragmentation can vary. Our first two empirical implications spell out the consequences, in our model, of changes in volatility and market fragmentation in the short term.[23]

---

[22] Weekly volatility is more appropriate than higher frequency volatility, because it is less likely to be affected by microstructure effects.

[23] We say that trading volume increases with a parameter if there is a range of α for which trading

Implication 1:   *In the short term, increased volatility (ϵ) raises the fraction of trades stemming from fast traders.*

Implication 1 is in line with the positive correlation between high–frequency trading and volatility evidenced by Chaboud et al.(2009) and Brogaard (2011). In our model, however, the presence of fast institutions does not affect the fundamental volatility of the asset. The causality runs the other way: high volatility increases the informational advantage of fast institutions, who correspondingly trade more, and increases adverse selection for slow institutions, who trade less.

Implication 2:   *In the short–term, increased market fragmentation (lower ρ) raises the fraction of trades stemming from fast traders and the permanent impact of trades.*[24]

As ρ goes down, slow institutions trade less. Fast institutions respond by also adjusting their activity downward, but less than slow institutions. Hence trades are more likely to stem from fast institutions. Correspondingly, trades are more likely to be motivated by information, and their permanent impact on prices increases.

Now, consider the long–run effect of changes in C or ϵ, i.e., their effect on equilibrium investment in fast trading technology (α) and its consequences for the market.[25] Empirically, α could be proxied by the number of trading desks that are co-located, or equipped for algorithmic trading (as in Chaboud et al., 2009).

Implication 3:   *Increases (resp. decreases) in C reduce (resp. increase) investment in fast–trading technology and correspondingly the fraction of trading volume stemming from fast institutions. This lowers (resp. raises) the permanent impact of trades and can have ambiguous consequences on trading volume.*

---

volume *strictly* increases in this parameter and *no* value of α for which it decreases. Implications 1 and 2 directly follow from our analysis above, in particular the definition of α_1, α_2 and α_3, as well as the expressions for volume in Corollary 1.

[24]In line with the literature (e.g., Hasbrouck (1991)), the permanent impact of a trade is its informational content (i.e., $|E(v|\omega) - \mu|$ in our set up).

[25]When there are multiple equilibria, we focus on the stable ones. Implications 3, 4 and 5 directly follow from our analysis above, in particular the expression of $\phi - \psi$ in the proof of Proposition 4.

Increases in investment in fast–trading arise, e.g., when institutions develop softwares and human capital to engage in high–frequency trading. Decreases in the investment in fast–trading by financial institutions arise, e.g., when they stop purchasing co-location services, or when they stop subscribing to high–speed fiber optic connections. Implication 3 is consistent with anecdotal evidence in an October 14, 2012, New–York Times article entitled "High-Speed Trading no Longer Hurtling Forward". One interviewee explains cut in investments in high–speed trading by "*the rising cost of technology [...]. One example is the data sold by the exchanges. Until recently, most firms bought a data package from Nasdaq that cost about $1,000 a month and gave traders a feed that includes all orders submitted to the exchange. In August, Nasdaq introduced a new, more comprehensive data package that costs about $25,000 a month. Many firms say that they have to sign up for new technologies if their competitors do.*"

Implication 4: *Sustained increases in volatility raise investment in fast trading when the latter is low, but reduce it when the latter is high.*

When the fraction of fast institutions is relatively low, sustained increases in volatility raise the profitability of fast informed trades, as well as the adverse selection cost borne by slow traders. This raises investment in fast trading technology. Conversely, drops in volatility should reduce such investment. This is consistent with anecdotal evidence: while investments in fast–trading technologies soared during the highly volatile crisis period, they seem to have gone down a bit in the calmer recent environment. For instance, the above mentioned article of the New-York Times writes: "*Profits from high-speed trading in American stocks are [...] down 35 percent from last year and 74 percent lower than the peak of about $4.9 billion in 2009, according to estimates from the brokerage firm Rosenblatt Securities. [...] firms large and small have been cutting staff, and in some cases have shut down.[...] For high-speed traders, rising prices are actually a part of the problem: climbing stock markets tend to be calmer stock markets, providing fewer trading opportunities for high-speed firms.*"

When α is large, however, increased volatility makes adverse selection costs so large for slow institutions that they prefer to abstain from trading. This induces a drop in liquidity

26

such that fast traders can't profitably exploit their informational advantage any more. In turn, this induces a decline in the number of institutions willing to bear the costs necessary to be fast.

Implication 5: *When $\rho \geq 1/2$ and $\alpha \geq \alpha_1$ investments in fast trading technology are strategic complements. Hence, an exogenous shock in the investment of some institutions should propagate and positively affect the investment decisions of other institutions.*

To test this prediction, one could rely on econometric approaches developed by IO analyses of network externalities, e.g., Rysman (2004) and Gowrisankaran and Stavins (2004). Investment decisions could be proxied by purchases of co–location services or high–throughput connections to exchanges. To account for the condition $\rho \geq 1/2$ and $\alpha \geq \alpha_1$, one could introduce interaction terms between the variable of interest and the level of fragmentation of the market and of investment in fast–trading technology. Identifying exogenous shocks in total investment in fast–trading technlogy will be more challenging and require finding good instruments.

Finally, increased market fragmentation, captured in our model by a decrease in ρ, adversely impacts slow traders. First, it reduces their ability to seize trading opportunities. Second, it increases the market impact of trades, and can generate market breakdowns for slow traders. One could expect such adverse effects on slow traders would induce a greater fraction of institutions to become fast. This does not hold for all parameter values, however. The reason is that increased market fragmentation also hurts fast traders. By lowering volume stemming from slow traders, market fragmentation reduces liquidity, which, in turn, reduces the ability of fast traders to conduct profitable trades. For some parameter values, this effect is so strong that an increase in market fragmentation actually reduces equilibrium investment in fast trading technologies.

8 Conclusion

Investment in fast trading technology helps financial institutions cope with market fragmentation. To the extent that this enhances their ability to reap mutual gains from trade,

it improves social welfare. On the other hand, fast institutions observe value relevant information before slow ones, which creates adverse selection. Thus, investment in fast trading generates negative externalities, which are detrimental to social welfare. Because financial institutions do not internalize these negative externalities in making their decisions, equilibrium investment in fast trading technologies can be larger than its utilitarian welfare maximizing counterpart. We identify cases where such excessive investment is particularly prevalent, as financial institutions incur the cost of investment in fast technologies just because they expect the others to do so and fear to be evicted from the market if they remain slow.

To avoid such inefficient arms' races, exchanges should offer investors the ability to post quotes requesting execution against slow traders only. Such quotes would be (partially) protected from the adverse selection costs due to the presence of fast traders. This would reduce the private incentives to socially excessive investment in fast trading technologies. This proposal, however, begs the question of the incentives of market organizers. Can we expect the competition between platforms and exchanges to deliver socially optimal outcomes? Could deviations from such outcomes arise? It could be, e.g., that exchanges are better able to extract rents from fast traders (via co-location and connection fees) than from slow ones. It could also be that liquidity coordination effects make it difficult for new, slow-friendly, platforms to enter the market. We leave these issues for further research.

Proofs

Proof of Lemma 1. In the market structure we consider, we have $a = E(v|\omega = 1)$ thus

$$a = \mu + (2\Pr(v = \mu + \epsilon|\omega = 1) - 1)\epsilon. \tag{9}$$

Using Bayes' rule,

$$\Pr(v = \mu + \epsilon|\omega = 1) = \frac{\alpha(\frac{1}{2} + \frac{\beta_{GL}^F}{2}) + (1-\alpha)(\frac{\beta_H^S}{2} + \frac{\beta_L^S}{2})}{[\alpha(\frac{1}{2} + \frac{\beta_{GL}^F}{2}) + (1-\alpha)\frac{\beta_H^S+\beta_L^S}{2}] + [\alpha\frac{\beta_{BH}^F}{2} + (1-\alpha)\frac{\beta_H^S+\beta_L^S}{2}]},$$

which, after simple manipulation, yields (1). ∎

Preliminary remarks for the proofs of Propositions 1, 2 and 3. For the proofs of these propositions, it useful first to write the expected profit, $\overline{\Pi}_j^F$, of a fast institution of type $j \in \{GH, GL, BH, BL\}$ when it buys one share of the security. We obtain $\overline{\Pi}_{GH}^F = (\mu + \epsilon + \delta - a)$, $\overline{\Pi}_{BH}^F = (\mu - \epsilon + \delta - a)$, $\overline{\Pi}_{GL}^F = (\mu + \epsilon - \delta - a)$, and $\overline{\Pi}_{BL}^F = (\mu - \epsilon - \delta - a)$. Similarly, the expected profits, $\overline{\Pi}_j^S$, of a slow institution of type $j \in \{H, L\}$ when it buys one share of the security is $\overline{\Pi}_L^S = (\mu - \delta - a)$ or $\overline{\Pi}_H^S = (\mu + \delta - a)$.

From Lemma 1, we know that $\mu \leq a \leq \mu + \epsilon$. Hence it is never optimal to buy the asset for (i) a slow institution with a low private valuation and (ii) a fast institution with a bad signal and a low private valuation. Thus, $\beta_L^S = 0$ and $\beta_{BL}^F = 0$. By symmetry, the expected profit of (i) a fast institution with a high private valuation and good news or (ii) a slow institution with a high private valuation must be negative if it sells the asset.

Moreover, the expected profit of a fast institution with good news and high private valuation is always strictly positive if it buys the asset (since $a \leq \mu + \epsilon < \mu + \epsilon + \delta$). Consequently, $\beta_{GH}^F = 1$. Finally, if $\epsilon > \delta$, we have $\overline{\Pi}_{BH}^F < 0$. Thus, a fast institution with bad news and a high private valuation does not buy ($\beta_{BH}^F = 0$), and by symmetry neither does a fast institution with good news and a low private valuation sell when $\epsilon > \delta$.

Proof of Proposition 3. As noted above, $\beta_L^S = \beta_{BL}^F = 0$, $\beta_{GH}^F = 1$, and when $\epsilon > \delta$, $\beta_{BH}^F = 0$. Hence, below we just discuss β_H^S and β_{GL}^F. From preliminary remarks, a slow trader with high private valuation and a fast trader with good news and low private valuation never find it optimal to sell. We therefore analyze below whether buying is

profitable for these two types of traders.

Part 1: If $a = \mu + \delta$, then $\overline{\Pi}^F_{GL} < 0$. Hence, $\beta^F_{GL} = 0$. Slow institutions with a high private private valuation are just indifferent between buying or not since $\overline{\Pi}^S_L = 0$. Hence, playing a mixed strategy ($0 \leq \beta^S_H \leq 1$) is optimal for them. Thus, if institutions expect $a = \mu + \delta$ then it is optimal for them to behave as in M2.

Now suppose that institutions behave as described in M2 with $\beta^S_L = \beta^F_{BH} = 0$, $\beta^F_{GL} = 0$, and $\beta^S_H = \frac{\alpha}{2(1-\alpha)\rho\delta}(\epsilon - \delta)$. Then equation (1) yields $a = \mu + \delta$. Thus, M2 equilibrium exists as long as $\frac{\alpha}{2(1-\alpha)\rho\delta}(\epsilon - \delta) \leq 1$, that is, $\alpha \leq \alpha_3$.

Part 2: We first prove that P1 exists iff $\alpha < \alpha_1$. Suppose $a = \mu + \frac{\alpha}{\alpha+(1-\alpha)\rho}\epsilon$. Then we have $a < \mu + \epsilon - \delta$ iff $\alpha < \alpha_1$. Hence, the expected profit of a fast institution with good news and low private valuation is strictly positive if it buys. Therefore, $\beta^F_{GL} = 1$. Furthermore, the expected profit of a slow institution with a high private valuation is positive if it buys the asset because $a < \mu + \delta$. Hence, $\beta^S_H = 1$. Thus, institutions behave as in P1.

Now suppose that institutions behave as in P1, with $\beta^S_L = \beta^F_{BH} = 0$, $\beta^S_H = 1$, and $\beta^F_{GL} = 1$. Then equation (1) yields $a = \mu + \frac{\alpha}{\alpha+(1-\alpha)\rho}\epsilon$. Thus, P1 exists iff $\alpha < \alpha_1$.

Part 3: We now prove that M1 exists iff $\alpha_1 \leq \alpha \leq \alpha_2$. Suppose $a = \mu + \epsilon - \delta$. As $a < \mu + \delta$, $\beta^S_H = 1$. A fast institution with good news and a low private valuation is just indifferent between buying and doing nothing since $\overline{\Pi}^F_{GL} = 0$. Hence purchasing the security with probability $\beta^F_{GL} = \frac{2(1-\alpha)\rho}{\alpha\delta}(\epsilon - \delta) - 1$ is optimal for an institution with type GL. This probability is greater than zero and less than one iff $\alpha_1 \leq \alpha \leq \alpha_2$. Thus, if institutions expect to trade at $a = \mu + \epsilon - \delta$, then it is optimal for them to behave as in M1 iff $\alpha_1 \leq \alpha \leq \alpha_2$.

Now suppose that institutions behave as in M1, with $\beta^S_L = \beta^F_{BH} = 0$, $\beta^S_H = 1$, and $\beta^F_{GL} = \frac{2(1-\alpha)\rho}{\alpha\delta}(\epsilon - \delta) - 1$. Then equation (1) yields $a = \mu + \epsilon - \delta$. Thus, M1 exists iff $\alpha_1 \leq \alpha \leq \alpha_2$.

Part 4: Finally, we prove that P2 exists iff $\alpha_2 < \alpha < \alpha_3$. Suppose $a = \frac{\alpha}{\alpha+2(1-\alpha)\rho}\epsilon$. As $\alpha_3 < \alpha$, we have $a < \mu + \delta$, hence $\beta^S_H = 1$. Moreover, as $\alpha > \alpha_2$, we have $a > \mu + \epsilon - \delta$. Hence, a fast institution with good news and a low private valuation makes a negative expected profit if it buys the asset ($\overline{\Pi}^F_{GL} < 0$). Thus, it does not buy the asset: $\beta^F_{GL} = 0$. This shows that if institutions expect to trade at $a = \frac{\alpha}{\alpha+2(1-\alpha)\rho}\epsilon$ then it is optimal for them

30

to behave as in P2.

Now suppose that institutions behave as in P2, with $\beta_L^S = \beta_{BH}^F = 0$, , $\beta_H^S = 1$, and $\beta_{GL}^F = 0$. Equation (1) yields $a = \mu + \frac{\alpha}{\alpha + 2(1-\alpha)\rho}\epsilon$. Thus, we have shown that P2 exists iff $\alpha_2 < \alpha < \alpha_3$. ∎

Proof of Lemma 2. When $\alpha > \alpha_3$, the unique equilibrium is $P3$. Now consider the case $\alpha \leq \alpha_3$.

First consider the expected profit of fast institutions conditional on buying the asset, $H(\alpha)$. A necessary condition for fast institutions to buy is that they receive good news. In this case, they buy the asset with probability β_{Gk}^F for $k \in \{H, L\}$. As $\beta_{GH}^F = 1$, we deduce:

$$H(\alpha) = \frac{1}{2}(\mu + \epsilon + \delta - a) + \frac{1}{2}(\mu + \epsilon - \delta - a) \times \beta_{GL}^F. \tag{10}$$

The total expected profit for fast institutions is $H(\alpha)$ since sales are symmetric to purchases. Using the expression for the equilibrium value of a and β_{GL}^F in the various types of equilibria, we obtain: $H(\alpha) = \frac{(1-\alpha)\rho}{\alpha + (1-\alpha)\rho}\epsilon$ in P1, $H(\alpha) = \delta$ in M1, $H(\alpha) = \frac{1}{2}(\delta + \frac{(1-\alpha)\rho}{\alpha/2 + (1-\alpha)\rho}\epsilon)$ in P2, $H(\alpha) = \epsilon/2$ in M2, and $H(\alpha) = \delta/2$ in P3. As $\epsilon > \delta$, fast institutions are better off in M2 equilibrium than in P3 for all values of α. Moreover, we have $\frac{(1-\alpha)\rho}{\alpha + (1-\alpha)\rho}\epsilon > \epsilon/2$, for $0 \leq \alpha < \alpha_1$, $\delta > \epsilon/2$, and $\frac{1}{2}(\delta + \frac{(1-\alpha)\rho}{\alpha/2 + (1-\alpha)\rho}\epsilon) > \epsilon/2$, for $\alpha_2 < \alpha < \alpha_3$. Hence when P1, M1 and P2 exist, they generate larger profits for fast than M2 and P3.

Second, consider the expected profit of slow institutions conditional on buying the asset, $S(\alpha)$. A necessary condition for slow institutions to buy the asset is that they have a high private valuations. Hence:

$$S(\alpha) = (\mu + \delta - a) \times \beta_H^S. \tag{11}$$

The expected gains from trade for slow institutions is just $\rho S(\alpha)$ since a slow institution finds a trading opportunity with probability ρ, and sales are symmetric to purchases. Using the expression for the equilibrium value of a and β_H^S in the various types of equilibria, we obtain $\rho S(\alpha) = (\delta - \frac{\alpha}{\alpha + (1-\alpha)\rho}\epsilon)\rho$ in P1, $(2\delta - \epsilon)\rho$ in M1, $(\delta - \frac{\alpha/2}{\alpha/2 + (1-\alpha)\rho}\epsilon)\rho$ in P2, and 0 in M2 and P3. Clearly, slow institutions' expected gains in P1, M1 or P2 are larger than in P3 and M2.

We conclude that when $\alpha < \alpha_3$, P1, M1 and P2 (when they exist) Pareto dominate P3 and M2. Hence, when $\alpha < \alpha_3$, there is a unique Pareto dominant equilibrium (P1, M1, or P2, depending on the exact value of α) since equilibria P1, M1 and P2 cannot be obtained simultaneously. When $\alpha = \alpha_3$, M2 Pareto dominates P3 as fast institutions get a higher profit in M2. Last, for $\alpha > \alpha_3$, there is a unique equilibrium, P3. Therefore, there is a unique Pareto dominant equilibrium for each value of α. ∎

Proof of Corollary 1. In a given equilibrium, the likelihood of trade by a given institution is simply twice the likelihood that it buys the asset (since, by symmetry, buy and sell orders are equally likely in equilibrium). Thus $\text{Vol}(\alpha) = 2\left(\frac{\alpha}{2}(1+\beta_{GL}^F) + \frac{1}{2}(1-\alpha)\rho\beta_H^S\right)$. The expression for $\text{Vol}(\alpha)$ in the corollary then follows by replacing β_{GL}^F and β_H^S by their expression in the Pareto dominant equilibrium (see Lemma 2) for each value of α.∎

Proof of Corollary 2. In the proof of Lemma 2, we give the expression of fast institutions' expected profit in the Pareto dominant equilibrium for each value of α (see equation (10)). Using this expression, we immediately deduce that $\phi(\alpha) = \frac{(1-\alpha)\rho}{\alpha+(1-\alpha)\rho}\epsilon$ for $0 \le \alpha < \alpha_1$, δ for $\alpha_1 \le \alpha \le \alpha_2$, $\frac{1}{2}(\delta + \frac{(1-\alpha)\rho}{\alpha/2+(1-\alpha)\rho}\epsilon)$for $\alpha_2 < \alpha \le \alpha_3$, and $\delta/2$ for $\alpha > \alpha_3$. Observe that $\lim_{\eta\to 0}\phi(\alpha_1 - \eta) = \delta = \phi(\alpha_1)$, $\lim_{\eta\to 0}\phi(\alpha_2 + \eta) = \delta = \phi(\alpha_2)$, and $\lim_{\eta\to 0}\phi(\alpha_3 - \eta) = \frac{\epsilon}{2} = \phi(\alpha_3)$. Thus, $\phi(.)$ is continuous and monotonically decreasing over $[0, \alpha_3]$. Moreover, $\phi(\cdot)$ is discontinuous and experiences a downward jump at $\alpha = \alpha_3$ since $\phi(\alpha_3) = \frac{\epsilon}{2} > \phi(\alpha) = \delta/2$ for $\alpha > \alpha_3$. Thus, $\phi(.)$ decreases in α over $[0,1]$.

Similarly, in the proof of Lemma 2, we give the expression of slow institutions' expected profit in the Pareto dominant equilibrium for each value of α, that is based on their expected profit conditional on buying (see equation (11)). Using this expression, we deduce that $\psi(\alpha) = (\delta - \frac{\alpha}{\alpha+(1-\alpha)\rho}\epsilon)\rho$ for $0 \le \alpha < \alpha_1$, $(2\delta - \epsilon)\rho$ for $\alpha_1 \le \alpha \le \alpha_2$, $(\delta - \frac{\alpha/2}{\alpha/2+(1-\alpha)\rho}\epsilon)\rho$ for $\alpha_2 < \alpha \le \alpha_3$, and 0 for $\alpha > \alpha_3$. Obviously, $\psi(\alpha)$ decreases in α over the following intervals $\alpha \in [0, \alpha_1)$ and $\alpha \in (\alpha_2, \alpha_3]$. Otherwise it is constant. Moreover $\lim_{\eta\to 0}\psi(\alpha_1 - \eta) = (2\delta - \epsilon)\rho = \psi(\alpha_1)$, $\psi(\alpha_2) = \lim_{\eta\to 0}\psi(\alpha_2 + \eta) = (2\delta - \epsilon)\rho$, and $\psi(\alpha_3) = 0 = \lim_{\eta\to 0}\psi(\alpha_3 + \eta)$. Hence $\psi(\cdot)$ is decreasing and continuous over $[0,1]$.∎

Proof of Proposition 4.   Corollary 2 yields

$$
(\phi - \psi)(\alpha) = \begin{cases}
\frac{\rho\epsilon}{\alpha+(1-\alpha)\rho} - \delta\rho & \text{for } 0 \leq \alpha < \alpha_1, \\
\epsilon\rho + \delta(1 - 2\rho) & \text{for } \alpha_1 \leq \alpha \leq \alpha_2, \\
\frac{1}{2}\left(\frac{\rho\epsilon}{\alpha/2+(1-\alpha)\rho} + \delta(1 - 2\rho)\right) & \text{for } \alpha_2 < \alpha \leq \alpha_3, \\
\delta/2 & \text{for } \alpha > \alpha_3.
\end{cases}
\tag{12}
$$

For $\alpha \geq \alpha_3$, $(\phi - \psi)(\alpha) = \phi(\alpha)$ since $\psi(\alpha) = 0$. The first part of the proposition follows from the fact that $\phi(\cdot)$ is discontinuous and jumps downward at $\alpha = \alpha_3$ (see the proof of Corollary 2). The rest of the proposition is obtained by computing and signing the derivative of the function $(\phi - \psi)(\cdot)$ with respect to α.∎

Preliminary remarks for the proofs of Propositions 5 and 6. For all values of ρ, we have (see equation (12)):

$$
(\phi - \psi)(0) = \epsilon - \delta\rho,
$$
$$
(\phi - \psi)(\alpha) = \epsilon\rho + \delta(1 - 2\rho), \text{ for } \alpha \in [\alpha_1, \alpha_2],
$$
$$
(\phi - \psi)(\alpha_3) = \tfrac{\epsilon}{2},
$$
$$
\text{and } (\phi - \psi)(\alpha) = \delta/2, \text{ for } \alpha \in (\alpha_3, 1].
$$

Figures 5 and 6 plot $(\phi - \psi)(.)$, and in particular its values when $\alpha = 0$, $\alpha = \alpha_2$, $\alpha = \alpha_3$, and $\alpha = 1$. The horizontal line gives the value of C. The propositions follow from varying C on the figure.

Proof of Proposition 5.   When $\rho \leq \frac{1}{2}$, we know from Proposition 4 that the function $(\phi - \psi)(\cdot)$ monotonically decreases with α when $\rho \leq \frac{1}{2}$. This yields $(\phi - \psi)(0) > (\phi - \psi)(\alpha_2) > (\phi - \psi)(\alpha_3) > (\phi - \psi)(1)$. The proposition follows directly from varying C on Figure 5.∎

Proof of Proposition 6.   When $\rho > \frac{1}{2}$, in contrast to the case in which $\rho \leq \frac{1}{2}$, $(\phi - \psi)(\alpha)$ does not decrease with α everywhere. In fact, it increases in α over the range $(\alpha_2, \alpha_3]$ (see Proposition 4). For this reason, there are different possible rankings of $(\phi - \psi)(0)$, $(\phi-\psi)(\alpha_2)$, $(\phi-\psi)(\alpha_3)$, and $(\phi-\psi)(1)$, as shown in Figure 6 (which considers two different possible rankings for $(\phi - \psi)(0)$, $(\phi - \psi)(\alpha_2)$, $(\phi - \psi)(\alpha_3)$, and $(\phi - \psi)(1)$). In any case, $(\phi - \psi)(\alpha_3) > (\phi - \psi)(\alpha_2)$ and $(\phi - \psi)(\alpha_3) > (\phi - \psi)(1)$.

33

First, consider the case in which

$$C > \max[(\phi - \psi)(0), (\phi - \psi)(\alpha_3)] = \max[\epsilon - \delta\rho, \frac{\epsilon}{2}].$$

In this case, it is immediate that C is greater than the largest possible values for $(\phi - \psi)(\alpha)$. Hence, the unique equilibrium level of investment in the fast trading technology is $\alpha^* = 0$. Alternatively, if:

$$C < \min[(\phi - \psi)(\alpha_2), (\phi - \psi)(1)] = \min[\epsilon\rho + \delta(1 - 2\rho), \frac{\delta}{2}],$$

then C is smaller than the smallest possible values for $(\phi - \psi)(\alpha)$. Hence, the unique equilibrium level of investment in the fast trading technology is $\alpha^* = 1$. These observations yield cases 1 and 6 of the proposition.

Now consider, the intermediate cases. In case 2 of the proposition (under the condition $\epsilon - \delta\rho > \frac{\epsilon}{2}$), we have $\frac{\epsilon}{2} < C < \epsilon - \delta\rho$, that is, $(\phi - \psi)(\alpha_3) < C < (\phi - \psi)(0)$. As $\epsilon > \delta$ and $\rho > \frac{1}{2}$, we deduce:

$$\max[(\phi - \psi)(\alpha_2), (\phi - \psi)(1)] < (\phi - \psi)(\alpha_3) < C < (\phi - \psi)(0).$$

Hence, there is a unique solution to the equation $C = (\phi - \psi)(\alpha^*)$ and this solution is strictly larger than zero because $C < (\phi - \psi)(0)$ and less than α_1 because $C > (\phi - \psi)(\alpha_2) = (\phi - \psi)(\alpha_1)$ (see Figure 6).

In case 3 of the proposition, we have $\max[\delta(1 - 2\rho) + \epsilon\rho, \frac{\delta}{2}] < C < \frac{\epsilon}{2}$. Hence,

$$\max[(\phi - \psi)(\alpha_2), (\phi - \psi)(1)] < C < (\phi - \psi)(\alpha_3).$$

In this case, the equation $C = (\phi - \psi)(\alpha^*)$ has two solutions. First, as $(\phi - \psi)(\alpha)$ decreases over $[0, \alpha_1)$, there is a solution in $[0, \alpha_1)$ since $C > (\phi - \psi)(\alpha_2) = (\phi - \psi)(\alpha_1)$. Either $C < (\phi - \psi)(0)$ and the solution is in $(0, \alpha_1)$, or $(\phi - \psi)(0) < C < (\phi - \psi)(\alpha_3)$ (under the condition $\epsilon - \delta\rho < \frac{\epsilon}{2}$) and the solution is $\alpha = 0$. Second, as $(\phi - \psi)(\alpha)$ increases over (α_2, α_3), there is a solution in (α_2, α_3) since $C < (\phi - \psi)(\alpha_3)$ and $C > (\phi - \psi)(\alpha_2)$. There cannot be other solutions (see Figure 6).

34

In case 4 of the proposition (under the condition $\delta(1 - 2\rho) + \epsilon\rho > \frac{\delta}{2}$), we have $\delta(1 - 2\rho) + \epsilon\rho < C < \frac{\delta}{2}$. Hence, $(\phi - \psi)(\alpha_2) < C < (\phi - \psi)(1)$, and since $\delta < \epsilon$ and $\rho > \frac{1}{2}$, we have:

$$(\phi - \psi)(\alpha_2) < C < (\phi - \psi)(1) < (\phi - \psi)(\alpha_3).$$

Then using the same arguments as in the previous case, we deduce that the equation $C = (\phi - \psi)(\alpha^*)$ has a solution in $[0, \alpha_1)$ and a solution in (α_2, α_3). In addition, we have $C < (\phi - \psi)(1)$ so that $\alpha^* = 1$ is also an equilibrium.

In case 5 of the proposition (under the condition $\delta(1 - 2\rho) + \epsilon\rho < \frac{\delta}{2}$), we have $\frac{\delta}{2} < C < \delta(1 - 2\rho) + \epsilon\rho$. Hence, as $(\phi - \psi)(\alpha_2) < (\phi - \psi)(0)$ and $(\phi - \psi)(\alpha_2) < (\phi - \psi)(\alpha_3)$, we have

$$(\phi - \psi)(1) < C < \min[(\phi - \psi)(\alpha_2), (\phi - \psi)(0), (\phi - \psi)(\alpha_3)].$$

In this case, there is no corner equilibrium since $(\phi - \psi)(1) < C < (\phi - \psi)(0)$. Moreover, over $[0, \alpha_3]$, the function $(\phi - \psi)(\alpha)$ reaches its minimum at $\alpha = \alpha_2$ (since it is weakly decreasing until α_2 and increasing over $[\alpha_2, \alpha_3]$). Hence, $C < (\phi - \psi)(\alpha)$ for $\alpha \in [0, \alpha_3]$. Thus, the equation $C = (\phi - \psi)(\alpha)$ has no solution since $C < (\phi - \psi)(\alpha)$ for $\alpha \in [0, \alpha_3]$ and $C > (\phi - \psi)(\alpha)$ for $\alpha \in (\alpha_3, 1]$ (as $(\phi - \psi)(\alpha)$ is constant over $(\alpha_3, 1]$). Consequently, none of Conditions (4), (5) and (6) are satisfied. In this case, as explained in the text, $\alpha^* = \alpha_3$ is the unique equilibrium. The six cases considered in the proposition cover all possible configurations for the parameters when $\frac{\epsilon}{2} < \delta < \delta$. Hence we have characterized all possible equilibria when $\rho > \frac{1}{2}$. $\blacksquare$

Proof of Proposition 7. Taking the derivative of W, we obtain:[26]

$$W'(\alpha) = ((\phi - \psi)(\alpha) - C) + \alpha\phi'(\alpha) + (1 - \alpha)\psi'(\alpha). \tag{13}$$

Corollary 2 implies that $W'(\alpha)$ is equal to $-\delta\rho - C$ for $0 \le \alpha < \alpha_1$, $\epsilon\rho + \delta(1 - 2\rho) - C$ for $\alpha_1 \le \alpha \le \alpha_2$, $\delta(\frac{1}{2} - \rho) - C$ for $\alpha_2 < \alpha \le \alpha_3$, and $\delta/2 - C$ for $\alpha > \alpha_3$. Clearly, $W(.)$ is piecewise linear in α. Thus, candidates for the social optimum are $0, \alpha_1, \alpha_2, \alpha_3$ and 1. Clearly, $W'(\alpha) < 0$ for $\alpha \in [0, \alpha_1)$. Thus, there is no social optimum for which $\alpha = \alpha_1$.

---

[26] At $\alpha = \alpha_3$, we take the left hand side derivative as the right hand side derivative is not defined since $W(\cdot)$ is discontinuous on the right hand side at $\alpha = \alpha_3$.

We now consider the remaining possibilities.

First, consider the case in which $\rho > \frac{1}{2}$. In that case, $W(.)$ decreasing on $(\alpha_2, \alpha_3]$ since $\rho > \frac{1}{2}$ and $C \geq 0$. Remaining candidates for the social optimum are therefore $0, \alpha_2,$ or 1 (the latter two when $W(.)$ is increasing on $[\alpha_1, \alpha_2]$ or on $(\alpha_3, 1]$). Calculations yield:

$$W(0) = \delta\rho, \tag{14}$$

$$W(\alpha_2) = \delta\rho + \left((\frac{1}{2} - \rho)\delta - C\right)\frac{\rho(\epsilon - \delta)}{\rho(\epsilon - \delta) + \delta}, \tag{15}$$

$$W(\alpha_3) = \delta\rho + \left((\frac{1}{2} - \rho)\delta - C\right)\frac{2\delta\rho}{2\delta\rho + \epsilon - \delta}, \tag{16}$$

$$W(1) = \frac{\delta}{2} - C. \tag{17}$$

Hence, (a) $W(\alpha_2) < W(0)$, and (b) $W(1) < W(0)$ since $\rho > \frac{1}{2}$ and $C \geq 0$. Therefore, $\alpha = 0$ is the social optimum when $\rho > \frac{1}{2}$.

Next turn to the case where $\rho \leq \frac{1}{2}$. We have the following ordering:

$$(\frac{1}{2} - \rho)\delta < \frac{\delta}{2} < \frac{\delta}{2} + \delta(\frac{1}{2} - \rho) + (\epsilon - \delta)\rho.$$

- If $C \leq (\frac{1}{2} - \rho)\delta$ then $W(.)$ is increasing on $[\alpha_1, \alpha_2]$, increasing or constant on $(\alpha_2, \alpha_3]$, and increasing after a discontinuous downward jump on $(\alpha_3, 1]$. Possible candidates for the optimum are thus $0, \alpha_3,$ and 1. Using equations (16) and (17), we obtain that:

$$W(1) > W(\alpha_3) \iff C < (\frac{1}{2} - \rho)\delta$$

  Besides, $W(1) > W(0)$ if $C < (\frac{1}{2} - \rho)\delta$ and $W(1) = W(0)$ if $C = (\frac{1}{2} - \rho)\delta$. We deduce that the social optimum is $\alpha = 1$ when $C < (\frac{1}{2} - \rho)\delta$. When $C = (\frac{1}{2} - \rho)\delta$, both $\alpha = 0$ and $\alpha = 1$.

- If $(\frac{1}{2} - \rho)\delta < C < \frac{\delta}{2}$, then $W(.)$ is increasing on $[\alpha_1, \alpha_2]$, decreasing on $(\alpha_2, \alpha_3]$, and increasing after a discontinuous downward jump on $(\alpha_3, 1]$. Hence, possible candidates for the social optimum are $0, \alpha_2,$ and 1. Using equations (14), (15) and (16), calculations show that $W(0) > W(1)$ and that $W(0) > W(\alpha_2)$ under $C > (\frac{1}{2} - \rho)\delta$. We deduce that the social optimum is $\alpha = 0$ when $(\frac{1}{2} - \rho)\delta < C < \frac{\delta}{2}$.

36

- If $\frac{\delta}{2} < C < \frac{\delta}{2} + \delta(\frac{1}{2} - \rho) + (\epsilon - \delta)\rho$, then $W(.)$ is increasing on $[\alpha_1, \alpha_2]$, decreasing on $(\alpha_2, \alpha_3]$, decreasing after a discontinuous downward jump on $(\alpha_3, 1]$. Possible candidates for the optimum are thus 0, and α_2. Using equations (14) and (16), we deduce that $W(0) > W(\alpha_2)$. Hence, the social optimum is $\alpha = 0$ when $C > (\frac{1}{2} - \rho)\delta$.

- If $C > \frac{\delta}{2} + \delta(\frac{1}{2} - \rho) + (\epsilon - \delta)\rho$, then $W(.)$ is decreasing on $[0, 1]$. Hence, in this case, the social optimum is $\alpha = 0$.

To sum up, when $\rho \leq \frac{1}{2}$, the social optimum is $\alpha = 0$ when $C > (\frac{1}{2} - \rho)\delta$ and $\alpha = 1$ otherwise. ∎

Proof of Corollary 3. When $\rho > \frac{1}{2}$, the utilitarian welfare optimum is $\alpha = 0$. Consequently, there is over-investment as long as the equilibrium level of investment is strictly positive. As shown in Proposition 6, this is the case when $C < \max(\epsilon - \delta\rho, \frac{\delta}{2})$. When $\rho \leq \frac{1}{2}$, we have the following ordering: $(\frac{1}{2} - \rho)\delta < \frac{\delta}{2} < \epsilon - \delta\rho$. If $C \geq \epsilon - \delta\rho$, the equilibrium level of investment in the fast trading technology is $\alpha = 0$, which is also the socially optimal level of investment in this technology (see Propositions 5 and 7). In contrast, if $C \leq (\frac{1}{2} - \rho)\delta$ then both the equilibrium level of investment and the utilitarian welfare optimum are $\alpha = 1$ since in this case $C < \frac{\delta}{2}$ (see Propositions 5 and 7). However, if $(\frac{1}{2} - \rho)\delta < C < \epsilon - \delta\rho$, the equilibrium level of investment is strictly positive (Proposition 5) but the social optimum level of investment is 0 (Proposition 7). Hence, in this case, there is excessive investment in the fast trading technology when $\rho \leq \frac{1}{2}$. To sum up, we deduce that there is overinvestment in the fast trading technology if and only if: $(\frac{1}{2} - \rho)\delta < C < \max(\epsilon - \delta\rho, \frac{\delta}{2})$. ∎

Bibliography

Baron, Matthew, Jonathan Brogaard, and Andrei Kirilenko (2012), "The trading profits of high–frequency traders," Working paper, University of Washington.

Bessembinder, Hendrik , Jia Hao and Kuncheng Zheng (2012), "Market Making Obligations and Firm Value," Working paper, David Eccles School of Business.

Breon–Drish, Bradyn (2010), "Asymmetric information in financial markets: Anything goes," Working paper, Stanford University.

Brogaard, Jonathan, 2011, "High-frequency trading and volatility," working paper, University of Washington.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (2012), "High frequency trading and price discovery," Working paper, University of California Berkeley.

Chaboud, Alain, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega, 2009, Rise of the machines: Algorithmic trading in the foreign exchange market, working paper, Federal Reserve Board.

Dow, James (2005), "Is liquidity self-fulfilling," *Journal of Business*, 77, 895-908.

Easley, David and Maureen O'Hara (1987), "Price, trade size and information is securities markets," *Journal of Financial Economics*, 19, 69-90.

Ganguli, Jayant, and Liyan Yang (2009), "Complementarities, multiplicity, and supply information," *Journal of the European Economic Association*, 7, 90-115.

Garvey, Ryan, and Fei Wu (2010), "Speed, Distance, And Electronic Trading: New Evidence On Why Location Matters," *Journal of Financial Markets*, 13, 367-396.

Glode, Vincent, Richard Green, and Richard Lowery (2012), "Financial Expertise as an Arms Race," *Journal of Finance*, 67, 1723-1759.

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

Gowrisankaran, Gautam, and Joanna Stavins (2004), "Network externalities and technology adoption: lessons from electronic payments," *Rand Journal of Economics*, 35, 260–276.

Grossman, Sandford, and Joseph Stiglitz (1980), "On the Impossibility of Informationaly Efficient Markets," *American Economic Review*, 70, 393-408.

Hasbrouck, Joel, 1991, Measuring the Information Content of Stock Trades, *Journal of Finance* 46, 179-207.

Hendershott, Terrence, Charles Jones, and Albert Menkveld (2011), "Does Algorithmic trading improve liquidity," *Journal of Finance*, 66, 1-33.

Hendershott, Terrence, and Ryan Riordan, (2011), Algorithmic Trading and Information, Working paper, University of Berkeley.

Hirshleifer, Jack (1971), "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61, 561-574.

IOSCO (2011), "Consultation Report on Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency," July 2011.

Jovanovic, Boyan, and Albert Menkveld (2011), "Middlemen in limit order markets," Working paper, New York University.

Kirilenko, Andrei, Pete Kyle, Mehrdad Samadi, and Tugkan Tuzun (2011), "The Flash Crash: The impact of high-frequency trading on an electronic market," Working paper, University of Maryland.

Manzano, Carolina, and Xavier Vives (2012), "Public and private learning from prices, strategic substitutability and complementarity, and equilibrium multiplicity," *Journal of Mathematical Economics*, 47, 346-369.

O'Hara, Maureen, and Mao Ye (2011), "Is market fragmentation harming market quality," *Journal of Financial Economics*, 100, 459-474.

Rysman, Marc (2004), "Competitition between Networks: A Study of the Market for Yellow Pages," *Review of Economic Studies*, 71, 483-512.

Figure 1: Equilibria for different values of α

Figure 1 shows the equilibria that can be obtained for each value of α. It highlights the multiplicity of equilibria when $\alpha < \alpha_3$.
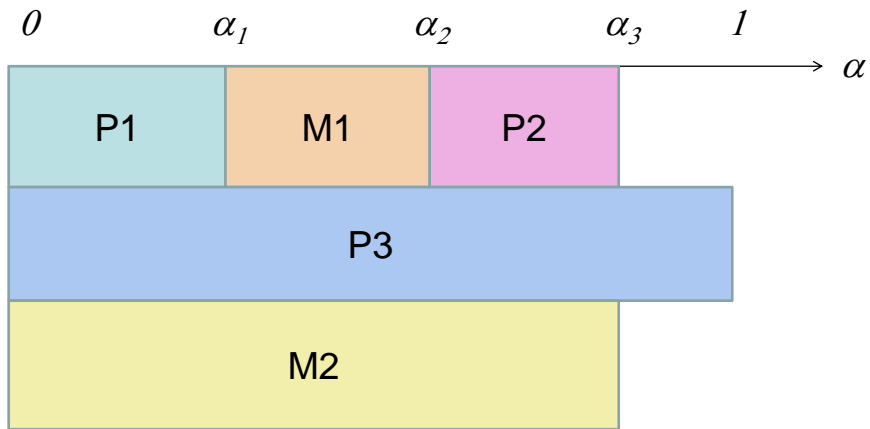

Figure 2: Price impact of a purchase in Pareto dominant equilibrium

Figure 2 shows the evolution of the price impact of buy orders, α-μ, as a function of α.
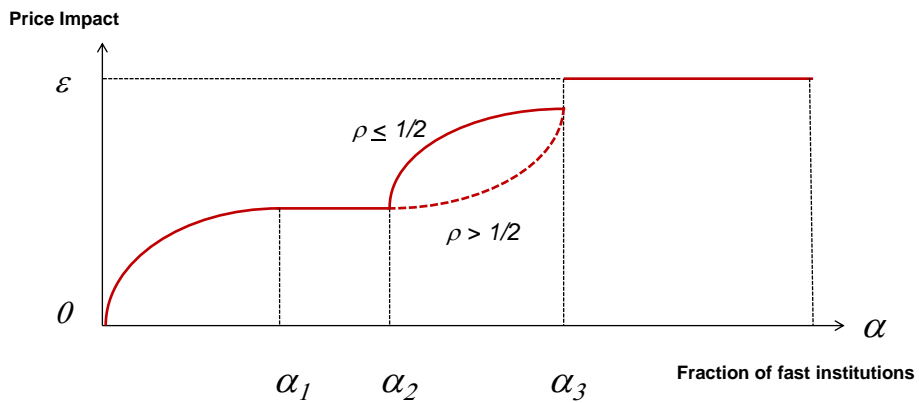
Figure 3: Volume in Pareto dominant equilibrium

Figure 3 shows the evolution of the total trading volume, Vol(α), as a function of α.

Note: The dotted lines depict the case where $\rho > 1/2$.
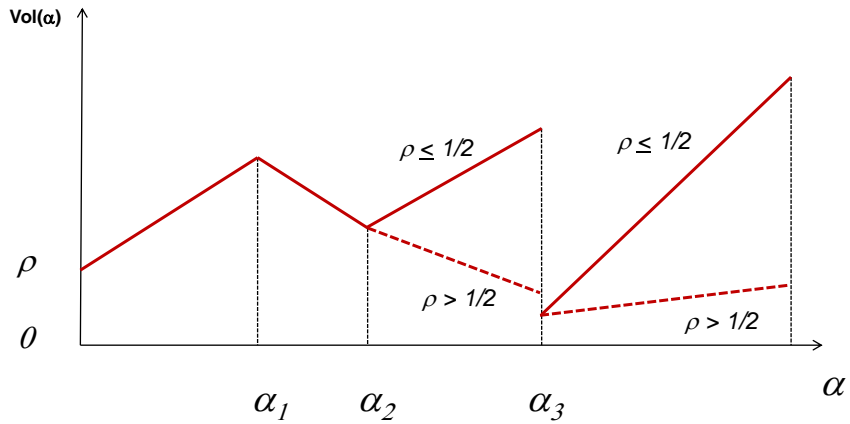
Figure 4: Expected gains of slow and fast institutions

Figure 4 shows the evolution of the expected gains for the fast traders and slow traders, that is, $\phi(\alpha)$ and $\psi(\alpha)$, respectively, as a function of α.

Note: The dotted lines depict the case where $\rho > 1/2$.
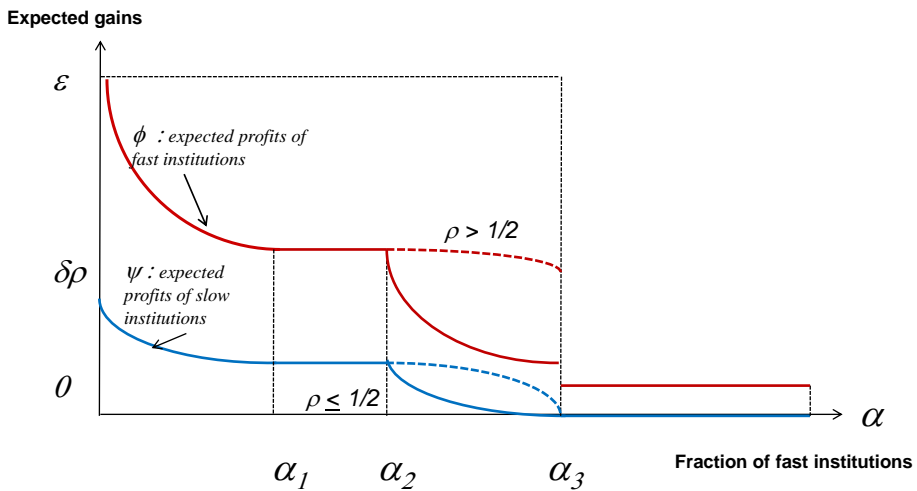
Figure 5: Equilibrium Investment when $\rho \leq \frac{1}{2}$

Figure 5 illustrates the determination of the equilibrium fraction of fast traders when $\rho \leq \frac{1}{2}$. We plot $(\phi-\psi)(.)$ as a function of α. The horizontal line gives the value of C.
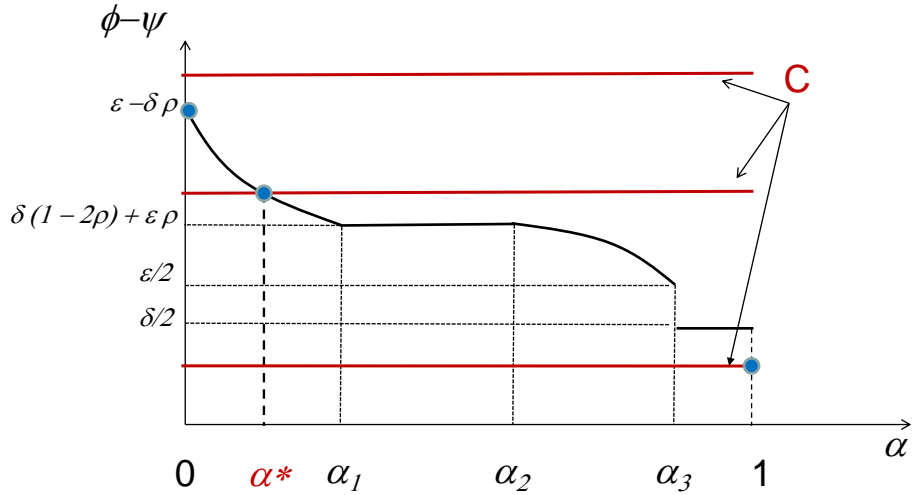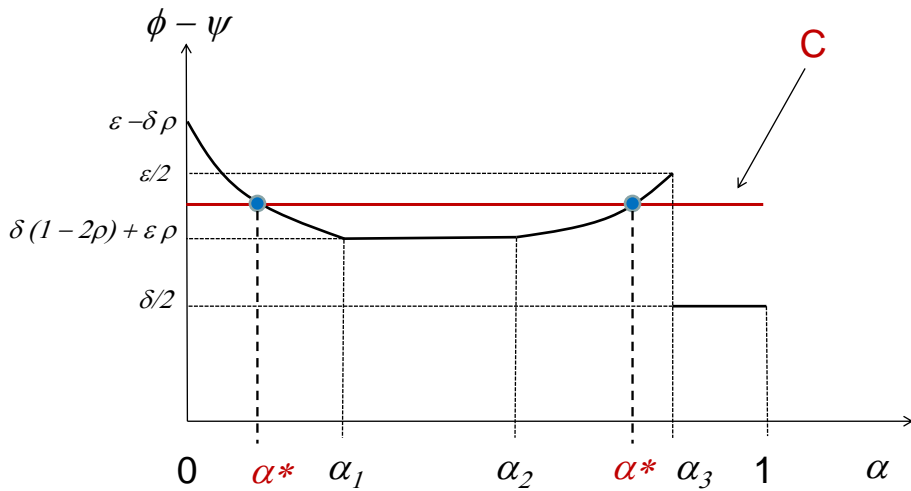

Figure 6, Panel 1

Equilibrium Investment when $\rho > \frac{1}{2}$ and $max(\delta/2,\delta(1-2\rho)+\varepsilon\rho))<min(\varepsilon-\delta\rho,\varepsilon/2)$

Figure 6 illustrates the determination of the equilibrium fraction of fast traders when $\rho > \frac{1}{2}$. We plot $(\phi-\psi)(.)$ as a function of α. The horizontal line gives the value of C.

Note: The figure depicts Case 3 in the Proposition when $\varepsilon/2 < \varepsilon - \delta\rho$, the opposite case is similar.

Figure 6, Panel 2.
Equilibrium when $\rho > \frac{1}{2}$ and $\delta(1-2\rho)+\varepsilon\rho<\delta/2$

Figure 6 illustrates the determination of the equilibrium fraction of fast traders when $\rho > \frac{1}{2}$. We plot $(\phi-\psi)(.)$ as a function of α. The horizontal line gives the value of C.

Note: The figure depicts Case 4 in the Proposition 6 when $\delta/2 < \varepsilon - \delta\rho$; the opposite case is similar.
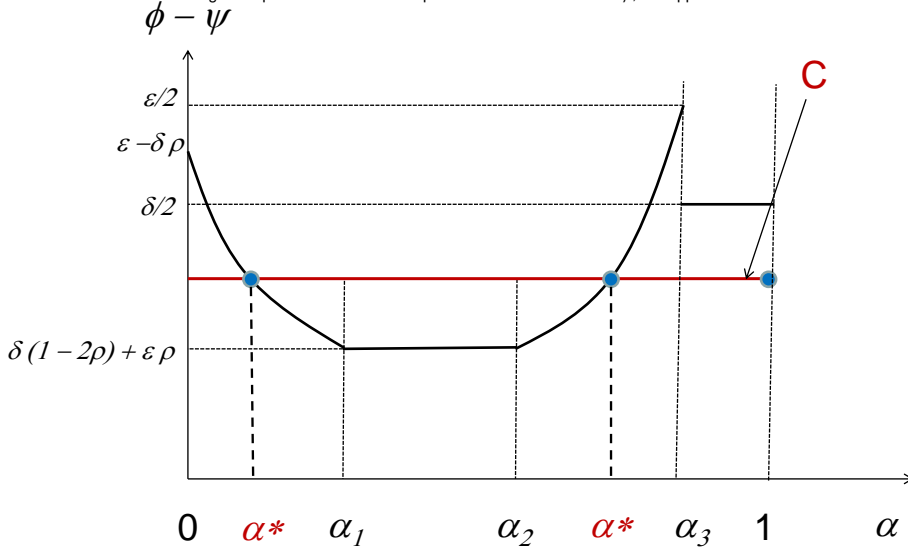

Figure 7, Panel 1: Equilibrium and utilitarian optimum when $\rho \leq \frac{1}{2}$

Figure 7 shows the equilibrium level of investment in the fast trading technology (red curve) and the socially optimal level of investment (blue dashed curve) as a function of the investment cost, C. Panel 1 focuses on the case $\rho \leq \frac{1}{2}$.

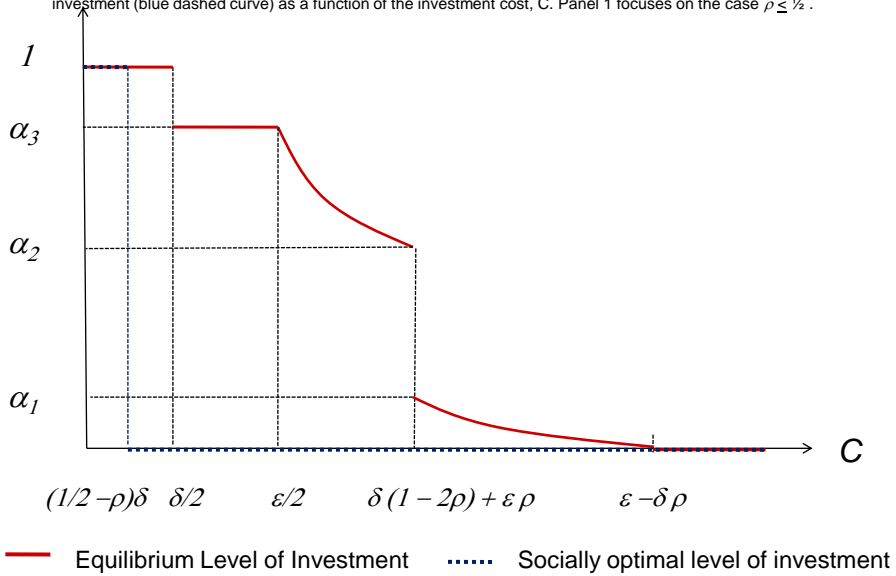—— Equilibrium Level of Investment ...... Socially optimal level of investment

Figure 7, Panel 2: Equilibrium and utilitarian optimum when $\rho > \frac{1}{2}$

Figure 7 shows the equilibrium level of investment in the fast trading technology (red curve) and the socially optimal level of investment (blue dashed curve) as a function of the investment cost, C. Panel 2 focuses on the case $\rho > \frac{1}{2}$.
Note: This figure considers the case in which $\varepsilon - \delta \rho > \varepsilon/2$ and $\delta/2 < \delta(1 - 2\rho) + \varepsilon \rho$; other cases are similar.

| Stable equilibrium level of investment | Unstable equilibrium level of investment | Socially optimal level of investment |