Price efficiency and High Frequency Trading in call auctions

EUROFIDAI WORKING PAPER

Panagiotis Anagnostidis[a], Patrice Fontaine[b,*], Christos Varsakelis[c]

[a] *Institut Europlace de Finance, Palais Brongniart 28 place de la Bourse, 75002 Paris, France*
European Financial Data Institute, 150 rue de la chimie, 38400 Campus de Saint Martin d'Hères, Grenoble, France
[b] *European Financial Data Institute, 150 rue de la chimie, 38400 Campus de Saint Martin d'Hères, Grenoble, France*
Léonard de Vinci Pôle Universitaire, Research Center, 92916 Paris La Défense, France
[c] *Université catholique de Louvain, Institute of Mechanics, Materials and Civil Engineering, 1348 Louvain–la–Neuve, Belgium*

Abstract

What is the effect of high frequency trading on the information content of call auction prices? To answer this question, we model a transparent electronic call market where High Frequency and Non High Frequency Traders (HFTs and NON HFTs) coexist. Because of their increased processing power and their ability to access multiple data streams within very short time intervals, HFTs have more precise information on fundamentals compared to NON HFTs and, thus, their bidding activity accelerates price efficiency. However, when HFTs behave strategically, considering the impact of their quotes on the market, price efficiency deteriorates. Subsequently, we test the predictions of our model regarding the effect of HFTs on call auction price efficiency, using a data set that includes a unique HFT flag for order and trade messages for the NYSE Euronext Paris opening call. Our results indicate that orders submitted by HFTs are more informative compared to those placed by NON HFTs. Moreover, early in the preopening, HFTs act strategically by submitting orders which they later (prior to opening) adjust or cancel. On the other hand, just before the opening, HFTs increase competition sharply by submitting large, aggressive and more informative orders with the intention of having them filled. Consequently, market prices are less informative early in the preopening, whereas they reflect information at the open. Our empirical results indicate that HFTs lead the price discovery process in the Paris market.

Keywords: High Frequency Trading, Call Auction, Rational Expectations, Price Efficiency

JEL Classification: G1, G14

---

*Corresponding author

Email addresses: panagiotis.anagnostidis@eurofidai.org (Panagiotis Anagnostidis), patrice.fontaine@eurofidai.org (Patrice Fontaine), christos.varsakelis@sabic.com (Christos Varsakelis)

1 Introduction

In the past decade, a burst of algorithmic and high frequency trading activity has been evidenced in financial markets. This is mainly attributed to i) the ever-increasing computational power that has rendered the implementation of order placement strategies at ultra–high speeds (e.g., millisecond accuracy) feasible, and ii) the improvement of electronic network communication systems in terms of speed, accessibility and quality that has allowed investors to collect accurate information about fundamentals and, thus, to enhance their risk assessment process and to increase their profits from trade. Nowadays, a substantial proportion of total volume of transactions in electronic markets (e.g., stock, currency and commodity markets) is triggered by High Frequency Trading programs (hereafter referred to, also, as HFTs). HFTs receive market information, analyze it and, subsequently, submit, adjust or cancel orders into the electronic system, aiming for capital gain opportunities that arise from temporal asset price movements. Moreover, algorithmic trading is often used by investors and fund managers as a means of reducing the cost of trading, via the application of 'slice and dice' techniques, as well as by market makers for various order handling and liquidity supply services (Brogaard, 2010; Jovanovic and Menkveld, 2016; Hendershott et al., 2011; Chlistalla, 2012; O'Hara, 2015; Foucault et al., 2016; Serbera and Paumard, 2016).

The presence of HFTs and their coexistence with Non High Frequency Traders (hereafter referred to, also, as NON HFTs) have raised important questions regarding the efficiency and the stability of financial markets. Even though a consensus on the role of HFTs has not been reached, several pros and cons have been identified thus far. As regards the positive contribution, there exists evidence that HFTs reduce transaction costs and improve information speed and price efficiency, while enhancing liquidity through their ability to provide simultaneous access to interlinked electronic trading platforms (Hasbrouck and Saar, 2013; Conrad et al., 2015). On the negative side, it is argued that the presence of HFTs may increase the possibility of abnormal price variations, through the rapid dissemination of quotes, or, even worse, of market failures (e.g., flash crashes) (Huang and Wang, 2009; Madhavan, 2012; Hasbrouck and Saar, 2013; Chordia et al., 2013; Kirilenko et al., 2017).

Literature on HFTs is quite extensive but almost exclusively concerns the continuous double auction mechanism.[1] A possible explanation for this focus stems from the fact that the main continuous session is rather attractive for HFTs, as they are able to take advantage of their speed to exploit arbitrage opportunities that arise within very short time intervals, often on the expense of slower and less informed traders (Biais et al., 2015). However, since most efforts have been devoted to the continuous double auction, there is

[1]Section 2 presents a detailed review of the literature on HFTs.

a notable dearth of studies for the second most important trading mechanism in organized markets, the single call auction. Consequently, relevant questions concerning the incentive of HFTs to participate in call auctions as well as their subsequent effect on price quality remain open.

The present study constitutes a step towards closing this gap in the literature by investigating first theoretically and then empirically the effect of HFTs on the informativeness of prices generated from the single call auction. The call auction mechanism differs from the continuous double auction in that orders are aggregated during a specific time period while trading is absent, whereas clearing takes place at a single auction price and at a predetermined point in time. To this extent, the single call is often adopted by Exchanges as an efficient mechanism in aggregating disperse information about fundamentals after extended periods of no trade (e.g., preopening) or after periods of higher volatility (e.g., trading halts); see, for example, Economides and Schwartz (1995). Our study investigates the presence of HFTs in such a trading environment and, in particular, it examines the effect of HFTs' order placement strategies on market price quality.

To model the clearing market, we develop a theoretical framework based on the notion of noisy rational expectations equilibria (Grossman and Stiglitz, 1976; Grossman, 1976; Grossman and Stiglitz, 1980; Hellwig, 1980; Admati, 1985; Madhavan and Panchapagesan, 2000). In our analysis, a finite number of uninformed, privately informed and liquidity traders participate in a transparent automated call auction market.[2] Thus, all traders have access to the market price, while market makers are absent.[3] Although this paradigm is classical, our theoretical framework differentiates from the state of the art in certain ways.

First, we introduce two types of privately informed traders: slow (human) and fast HFTs (machines). Each trader in the former group possesses a private signal about fundamentals as in the traditional rational expectations framework (Grossman and Stiglitz, 1980). Members of the second group, HFTs, view the same private signal as human informed traders but with greater precision, due to their ability to collect and process information over multiple data streams and electronic trading networks. This assumption implies that HFTs may have an incentive to participate in the call market not because of their increased order processing speed, but due to their informational advantage. Indeed, because bidding is only theoretical until the clearing time, momentum fast trading is significantly limited in the call auction mechanism. Thus, it is reasonable to assume that HFTs may enter the call market to profit from their private signals, without

[2] Because call auctions are less popular compared to continuous trading, herein we do not opt for a continuum of traders (Aumann, 1964) but rather a finite set of traders. By doing so, we differentiate from the bulk of related microstructure studies where a continuum of traders is preferred; see, for instance, Vives (1995), Cespa and Foucault (2013) and Biais et al. (2015).

[3] In Vives (1995), for example, competitive market makers set the clearing price by observing the order flow.

necessarily being ultra–fast.[4] Our approach is close to that of Cespa and Foucault (2013) and Easley et al. (2016) who consider different types of informed traders depending on their decision to invest in obtaining access to market data (e.g., realized transaction prices). In our model, however, informed traders are endowed exogenously with private signals of different precision, whereas all traders observe the market price that is publicly available. Further, uninformed traders view the market price as a weighted average of the private signals of the two types of informed traders and the amount of noise trading. Hence, in equilibrium, the belief of uninformed traders about the private signals of HFTs and human informed traders plays a significant role in the price formation process.

Second, herein we assume that HFTs can behave strategically, considering the impact of their bids on the market price (e.g., Kyle, 1985). This assumption is in line with one of the well–documented properties of HFTs, the implementation of frequent submit–cancel strategies (e.g., O'Hara, 2015). However, in a market with a finite number of traders, as the one considered herein, liquidity can be limited and, thus, prices can be more sensitive to significant supply–demand swifts. Thus, when agents place their bids strategically, additional noise is induced into prices. To incorporate the effect of strategic trading into our model, we add a price elasticity term in the standard Grossman and Stiglitz (1980) framework, as in Rindi (2008) and Kovalenkov and Vives (2014). Our choice is consistent with the analysis of Kovalenkov and Vives (2014) who show that, when the number of market participants is small, a competitive equilibrium can be utilized as an approximation of the true strategic equilibrium only when agents are risk averse but not close to risk neutral. Otherwise, considering a pure competitive equilibrium (i.e., as in Grossman and Stiglitz (1980)) can be problematic.

Several of our model predictions are qualitatively similar to those of the competitive model of Grossman and Stiglitz (1980). Price efficiency improves with the number of informed traders in the market increased. Also, noise trading and the magnitude of risk aversion are inversely correlated with price efficiency.

Regarding HFTs, when they are present in the market price efficiency improves significantly faster compared to a market of slow traders; this is due to the increased information content of HFTs' quotes. The model predicts, also, that when uninformed traders form their expectations by considering the accuracy of HFTs' signals more promptly, compared to the precision of human informed traders, price efficiency improves. On the other hand, when uninformed traders form their expectations by placing more weight on human informed traders' signals, prices are more noisy. The aforementioned predictions are valid to the

[4] Bellia et al. (2017), for example, find that HFTs earn profits from orders executed at the opening auction in the Paris market. Further, the authors attribute this finding to the fact that HFTs are able to collect information from various data sources, thus forming more precise signals on fundamentals compared to the rest of the trading public.

extent that market participants do not behave strategically. In this respect, we show that price efficiency is inversely correlated with the price elasticity factor for the two types of informed traders (HFTs and NON HFTs). Hence, when traders become strategic, price efficiency deteriorates significantly.

In the empirical part, we contribute to the literature on call auction mechanisms by providing insight on the behavior of HFTs in the preopening period of the Euronext Paris opening call. We employ a unique set of high frequency data that includes the entire trading and order placement history for the CAC 40 stocks in year 2013. A significant advantage of the database used in the present study is that it includes flagged messages that pertain to HFT activity, proxied by the ratio of individual order lifetime over the average order lifetime for order cancellations.[5] Therefore, we are able to distinguish directly between HFTs and NON HFTs.

For the purposes of our analysis we re–construct the order book and calculate preopening indicative prices and opening prices on the basis of: i) the full set of order messages, that is the actual market or, else, the 'Full order book', ii) the HFT set of order messages which we term as the 'HFT market' or the 'HFT order book' and iii) the set of messages related to the bidding activity of NON HFTs which, accordingly, we term as the 'NON HFT market' or the 'NON HFT order book'. By doing so, we opt to disentangle HFT from NON HFT activity and, in turn, to compare the informativeness of quotes placed by the two types of traders.[6] We examine empirically the information content of the re–constructed indicative and auction prices using the Weighted Price Contribution (WPC) statistic, originally introduced by Barclay and Warner (1993) in their study of NYSE, together with the unbiasedness regression technique proposed by Biais et al. (1999) in their analysis of the Paris opening auction. The former measures the amount of price discovery during the trading process, whereas the latter provides an estimation of the information content of prices, which is the main focus of our study.[7]

We find evidence that HFTs contribute significantly to price efficiency a few minutes before the opening by increasing their participation in the order accumulation process. To this extent, HFTs tend to submit aggressive and large orders with the intention of having them executed. These results are consistent with our model prediction that price efficiency is positively correlated with the number of HFTs in the market.

[5]The HFT classification is provided by the AMF (*Autorité des Marchés Financiers*) (see, Megarbane et al. (2017)).

[6]See, for instance, Madhavan and Panchapagesan (2000) and Anagnostidis et al. (2015) who build different types of order book states for the preopening of the NYSE and the Athens Exchange, respectively, by isolating market from limit orders.

[7]The WPC statistic is a widely popular measure of price discovery and it has been applied in several empirical studies in the past (Cao et al., 2000; Ciccotello and Hatheway, 2000; Barclay and Hendershott, 2003; Ellul et al., 2005; Moshirian et al., 2012; Anagnostidis et al., 2015; Bellia et al., 2016, 2017). Similarly, the unbiasedness regression methodology is routinely employed in the call auction literature for the estimation of price informativeness (Madhavan and Panchapagesan, 2000; Ciccotello and Hatheway, 2000; Barclay and Hendershott, 2003; Comerton-Forde and Rydge, 2006; Moshirian et al., 2012; Anagnostidis et al., 2015; Boussetta et al., 2016; Baruch et al., 2017).

Further, we find that HFTs' orders are more informative compared to those submitted by NON HFTs, consistent with the main assumption of our model that HFTs are able to form more precise signals, compared to the rest of the trading public, by processing data from various sources and at extremely high speeds. Our results indicate, also, that HFTs act strategically when placing their orders; they enter the market early in the preopening, whereas they modify and/or cancel their quotes aggressively later, just before the opening. In this respect, we find evidence of a strategic reversal in the bidding behavior of HFTs who tend to raise indicative prices 10 minutes before the clearing time, causing therefore positive returns, whereas they increase their selling activity a few minutes before the opening, generating negative returns. Accordingly, our evidence suggests that prices are rather noisy 10 minutes prior to opening, whereas price efficiency increases sharply just before the auction. Thus, our results corroborate our model prediction that strategic quoting (early in the preopening) deteriorates price efficiency, whereas competitive biding (prior to opening) accelerates it.

The paper is organized as follows: Section 2 provides a review of the literature. Section 3 develops the analytical framework. Section 4 presents an empirical application on the Paris opening call market. Finally, Section 5 concludes the paper.

2 Literature review

2.1 High frequency and algorithmic trading

The increasing interest on the effect of HFTs on the quality of financial markets has attracted considerable attention and has thus resulted in a large and growing body of the microstructure literature that focuses, mainly, on the continuous double auction mechanism. Although HFTs are frequently labelled as a potential source of abnormal price movements, there exists evidence that trading by means of machine learning methods may actually improve market quality. For instance, Brogaard (2010), Hasbrouck and Saar (2013), Hagströmer and Norden (2013), Carrion (2013) and Brogaard et al. (2014) study the low–latency trading environment in the NASDAQ market and indicate, collectively, that HFTs accelerate price discovery, reduce price volatility and enhance the provision of liquidity.

Hendershott et al. (2011) examine empirically the NYSE market and find that the rapid increase of algorithmic trading during the past decade has narrowed the spreads and reduced adverse selection costs for traders. Moreover, the authors find that algorithmic trading has improved market liquidity. In the order–driven setting, Hendershott and Riordan (2013) find that algorithmic trading in the DAX market improves price efficiency. Similarly, Boehmer et al. (2015) show that algorithmic trading has, on average, a positive effect on price efficiency and liquidity in 42 international equity markets around the world (order and quoted driven systems).

At this point it is pertinent to note that observers often refer to algorithmic trading and high frequency trading as two equal entities. Nonetheless, HFTs are designed to act, exclusively, within the domain of millisecond or microsecond accuracy, whereas algorithmic trading machines (also known as ATs) aim, mostly, to carry out various order handling and liquidity providing strategies on behalf of investors and market makers, without necessarily being ultra fast. Essentially, high frequency trading is a subset of algorithmic trading, with emphasis placed both on order handling and execution speed (Chlistalla, 2012).

Although there is ample empirical evidence on the positive effect of HFTs on traditional aspects of market quality, like liquidity and price efficiency, there are studies suggesting that, under circumstances, HFTs may contribute to the occurrence of market failures, such as flash crashes, the sharpening of price variations (e.g., widening of spreads) and the rise of systematic risk. Kirilenko et al. (2017), for example, investigate the NASDAQ flash crash on May, 2010. The authors show that the emergence of the sharp volatility spike at 2:45 is significantly related to the fast reaction of HFTs on the rapid depletion of liquidity. Boehmer et al. (2015) also find that algorithmic trading intensity is positively correlated with price volatility. In a theoretical paper, Biais et al. (2015) show that high speed connections (i.e., HFTs) enable investors to profit from trading by searching for desirable quotes at the spot and, by doing so, they impose additional adverse selection costs on slow traders. Chaboud et al. (2014) study algorithmic data from the EBS electronic trading system for currency pairs (euro–dollar, dollar–yen and euro–yen). The results of this study suggest that machine trades are often correlated, giving rise to systematic market risk. Brogaard (2010) investigates data from the NASDAQ and finds that order flows generated from HFTs are correlated, possibly imposing additional non–diversifiable (il)liquidity risk on investors' portfolio selection strategies. Similarly, Jain et al. (2016) examine the Tokyo Stock Exchange low–latency platform and find evidence that HFTs rise systematic liquidity risk.

Thus far, the majority of related empirical studies have focused on the continuous trading process in the US Exchanges, mainly the NASDAQ and the NYSE, whereas order–driven call auction systems have attracted less attention. From this point of view, the present study aims to shade further light on the impact of high frequency trading on the quality of automated electronic markets focusing on the call auction mechanism.

2.2 *Call auction trading*

Call auction venues differ significantly from continuous systems in that orders are aggregated without trading, leading to the emergence of crossed supply–demand schedules. Subsequently, at a pre–specified point in time, buy and sell orders are matched and executed at a single equilibrium price. Due to their ability to aggregate disperse information about fundamental values into one single price, call auctions are typically adopted by Exchanges as ideal trading mechanisms during periods of increased market stress, such as the opening and the closing.

Because bidding is only theoretical during the order accumulation period, pre–call communication games may emerge between investors. In particular, the pre–trade period in transparent call mechanisms may offer a learning environment where investors submit, adjust or cancel their orders by observing the flow of indicative clearing prices and the dynamics of the prevailing bids and asks. This is the case in the models of Jordan (1982), Vives (1995) and Biais et al. (1999), where competitive agents (i.e., traders, brokers and/or market makers) react to the disclosed information and drive the asset price to its equilibrium value. On the other hand, Medrano and Vives (2001) and Biais et al. (2013) show that the presence of privately informed strategic traders, who attempt to manipulate the market, may add noise to the price discovery process, thus slowing down the speed of information. Also, Madhavan and Panchapagesan (2000) show that the presence of liquidity traders may add noise to the price discovery process, through the submission of aggressive price–inelastic orders (e.g., market orders). Our theoretical analysis contributes to the aforementioned literature by considering the effect of HFTs in the call market.

Information revelation and price discovery in the call auction mechanism have been studied empirically in several studies in the past. Ciccotello and Hatheway (2000), Cao et al. (2000), Barclay and Hendershott (2003), Barclay and Hendershott (2008) and Pagano et al. (2013) investigate the NASDAQ dealers' market opening. Madhavan and Panchapagesan (2000) examine the NYSE opening procedure and the role of specialists in price discovery and price efficiency. In the order–driven trading framework, Biais et al. (1999) and Pagano and Schwartz (2003) and Hillion and Suominen (2004) analyze the opening and the closing auctions of the Euronext Paris stock market, respectively. Comerton-Forde and Rydge (2006) and Moshirian et al. (2012) study the opening and closing auctions in the Australian order–driven stock market. Hauser et al. (2012) examine the behavior of opening prices in the Tel Aviv Stock Exchange after the introduction of a random opening time, while Kalay et al. (2004) investigate, in the same market, the elasticity of clearing prices during the preopening period. Lastly, Anagnostidis et al. (2015) analyze the opening price discovery process in the Greek order–driven market. The aforementioned empirical studies indicate, altogether, the importance of market stability at periods of increased uncertainty, such as the opening or the closing. Additionally, they hint that the call auction is probably, until now, the most efficient mechanism for revealing prices after periods of no trade.

Our study is part of a new stream of the microstructure literature that investigates the effect of HFTs on the quality of call auction markets. Bellia et al. (2016) and Bellia et al. (2017) have recently investigated the role of HFTs in the preopening process of the Tokyo Stock Exchange and the Paris Stock Exchange, respectively. Although price efficiency is not their main focus, these studies provide, collectively, evidence that HFTs participate drastically in the preopening period and that they have a leading role in the price formation process and the liquidity dynamics in the call auction market, in line with our analysis. Boussetta et al. (2016) examine the overall role of the preopening period in the price discovery process across fragmented markets. Using high frequency data from the Paris market and the BATS and Chi–X multilateral trading facilities, they find, among others, that slow traders play a significant role in the price discovery process across the different trading venues. Our empirical analysis complements the aforementioned studies in that we provide insight on the information content of HFTs' quotes, as well as the impact of HFTs' order placement activity on the informativeness of prices generated from the single call auction mechanism.

3 Analytical backround

3.1 Derivation of the equilibrium

We consider an automated transparent one–period call market where a single risky asset with random ex–post liquidation value, $u \sim N(0, \sigma_u^2)$,[8] and a risk–less asset are traded. Furthermore, we normalize the return of the risk free asset to zero and, for the sake of simplicity, we assume that market agents do not hold an initial endowment.

The trading public consists of K fast HFTs, M slow–human informed, L uninformed and Z liquidity/noise traders ($K, M, L, Z \in \mathbb{N}$). HFTs are in possession of a unique set of information collected through various electronic networks. Therefore, in our analysis, HFTs are considered as privately informed traders. Slow (human) informed traders hold the same piece of information about fundamentals as HFTs but they view the asset price with less precision. That is, we assume that HFTs hold more precise information, compared to human informed traders, as they are able to have access to a wide range of data sources as well as increased processing power. Uninformed traders observe the publicly available market price, whereas they

[8]Without loss of generality we assume that $E(u) = 0$ (e.g., Kovalenkov and Vives, 2014).

do not hold any private information. Noise traders arrive in the market for liquidity purposes (e.g., to close a short position).

In the market under consideration, informed and uniformed traders act strategically, i.e. they take into account the impact of their own orders on the clearing price. Incorporating price elasticity in our analysis is deemed important since, herein, we assume a finite amount of traders and it is a priori possible that clearing prices could be strategically influenced by market participants' order placement activities (Medrano and Vives, 2001; Biais et al., 2013).

Informed and uninformed traders are assumed to be risk averse and, therefore, they have negative exponential expected utility functions of the form $U(W) = -\exp^{-AW}$, with A and W being the risk aversion coefficient and the terminal wealth, respectively; for simplicity, we assume that all agents have the same risk–aversion parameter (A). Further, HFTs, slow informed investors, uninformed investors and noise traders have, respectively, the per capita demand functions Q_F (F indicates Fast), Q_S (S indicates Slow), Q_U (U indicates Uninformed) and Q, with $Q \sim N(0, \sigma_Q^2)$. Because traders have access to information about the price–volume pair (i.e., market transparency), we assume that they behave rationally in the sense that, before posting their quotes, they update their expectations by observing the disseminated information; risk averse traders submit limit orders whereas liquidity/noise traders submit market orders.[9] After bidding, the market clears and trading takes place.

Uninformed traders do not hold private information and, thus, they use their prior beliefs conditional on the observed indicative price, p, to decide about their future investments. In particular, they maximize their expected utility

$$E[U(W)] = E[-e^{-AW}] \tag{1}$$

where $W = Q_U(u - p^*)$ is the terminal wealth and p^* is the clearing price. Note that since u is normally distributed, W is also normal

$$W \sim N(Q_U(E(u|p) - p^*), Q_U^2 Var(u|p)). \tag{2}$$

Using the properties of normal distribution, it is straightforward to show that maximizing the expected utility is equivalent to maximizing the following quantity (certainty equivalent)

$$E(W) - \frac{A}{2} Var(W). \tag{3}$$

[9] We assume that the demand of liquidity trading is exogenous.

Thus, to maximize expected utility, uninformed traders use the following first order condition:

$$E(u|p) - p - \frac{\partial p}{\partial Q_U} Q_U - AQ_U Var(u|p) = 0. \tag{4}$$

Notice that the price p is a function of the demand Q_U (strategic effect) and, therefore price impact (i.e., $\frac{\partial p}{\partial Q_U}$) is introduced in the model. By solving equation (4), the following demand function is derived for uninformed traders

$$Q_U = \frac{E(u|p) - p}{\omega_U + A\, Var(u|p)}, \tag{5}$$

where $\omega_U = \frac{\partial p}{\partial Q_U}$ is the price elasticity term; in a purely competitive market this term is identically zero.

In contrast to uninformed traders, both slow (human) informed investors and HFTs receive a private signal s with $s \sim N(0, \sigma_s^2)$ that corresponds to the sum of the true ex-post liquidation value u and a white noise term i.e.

$$s = u + e \tag{6}$$

where $e \sim N(0, \sigma_e^2)$ is a residual term and $e \perp u$ ($\sigma_e^2 > 0$).[10] What differentiates HFTs from slow informed traders is the precision with which they view the signal, i.e. we assume that the variance of the white noise term e, σ_e^2, is different between the two groups of informed traders. In other words, HFTs and human informed traders do not hold different opinions on the value of the risky asset u (i.e., they all know the true distribution of u), but rather receive different levels of noise.[11] The rationale behind this assumption is simple: HFTs have fast access to multiple data streams and increased processing power. Consequently, HFTs can build a private signal that improves their precision against human traders (informed and uninformed).

The informational advantage of HFTs over slow informed (and uniformed) traders translates into the following series of conditions:

$$E_F(e) = E_S(e) = 0, \tag{7}$$

$$Var_S(e) = \sigma_S^2 > Var_F(e) = \sigma_F^2, \tag{8}$$

$$cov_S(s, u) = cov_F(s, u) = Var(u) = \sigma_u^2, \tag{9}$$

where here and throughout this paper the indices S and F designate Slow and Fast, respectively. Note also

[10] We further assume that s and Q are uncorrelated.

[11] Our approach is similar to that of Hirshleifer and Luo (2001) who consider a competitive market with rational overconfident and liquidity traders. In their framework, rational traders perceive the true distribution of e, whereas overconfident traders believe, falsely, that the variance of e is smaller than the true variance.

that the correlation of the signals of the HFTs and slow informed traders is positive independently from whether the white noise terms have a non-zero covariance or not. Henceforth, we will assume that the white noise terms are independent.

In view of the above, the demand function for informed traders assumes the following form,

$$Q_{F,S} = \frac{E(u|s,p) - p}{\omega_{F,S} + A\,Var(u|s,p)}\,, \tag{10}$$

where $\omega_{F,S} = \dfrac{\partial p}{\partial Q_{F,S}}$ is the price elasticity with respect to HFTs and slow informed traders respectively. Notice that equation (10) can be derived from the following first order condition, as in the case of uninformed traders:

$$E(u|s,p) - p - \frac{\partial p}{\partial Q_{F,S}} Q_S - A Q_{F,S} Var(u|s,p) = 0. \tag{11}$$

We can now study the equilibrium of the market at hand. This is established in the following Proposition:

Proposition 3.1. *There exists a unique linear equilibrium such that the price p is given by the following linear rule*

$$p = \alpha s + \beta Q. \tag{12}$$

The coefficient α is uniquely determined as the solution to a linear equation and is given by the following expression

$$\alpha = \frac{\dfrac{C_1}{L} + \dfrac{\sigma_u^2}{\omega_U \left(\sigma_u^2 + \sigma_U^2 + \left(\dfrac{Z}{C_1}\right)^2 \sigma_Q^2\right) + A\,\sigma_u^2 \left(\sigma_U^2 + \left(\dfrac{Z}{C_1}\right)^2 \sigma_Q^2\right)}}{\dfrac{\sigma_u^2 + \sigma_U^2 + \left(\dfrac{Z}{C_1}\right)^2 \sigma_Q^2}{\omega_U \left(\sigma_u^2 + \sigma_U^2 + \left(\dfrac{Z}{C_1}\right)^2 \sigma_Q^2\right) + A\,\sigma_u^2 \left(\sigma_U^2 + \left(\dfrac{Z}{C_1}\right)^2 \sigma_Q^2\right)} + \dfrac{C_2}{L}}. \tag{13}$$

The coefficient β is given by the following expression

$$\beta = \alpha \times \frac{Z}{C_1}. \tag{14}$$

Here,

$$\begin{aligned} C_1 =& \frac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2} + \frac{M}{\dfrac{\omega_S}{\gamma_S} + A\sigma_S^2} \\ C_2 =& \frac{K}{\omega_F + A\gamma_F\sigma_F^2} + \frac{M}{\omega_S + A\gamma_S\sigma_S^2} \end{aligned}$$

and σ_U^2 reflects the precision with which uniformed traders choose to build their expectations; see begining of the proof.

Proof. Following standard practices, we assume that uninformed traders view the asset price as a function of the private signal and the demand of noise traders (Rindi, 2008; De Jong and Rindi, 2009). The normality of the variables considered herein asserts that this function must be linear, so that,[12]

$$\hat{p} = \alpha s + \beta Q, \tag{15}$$

[12]Linearity in s and Q is often considered an assumption. Nonetheless, within a rational expectation framework where all random variables are normal this is actually a corollary.

for some real numbers α and β.[13] A subtle point in this respect concerns the variance that uniformed traders will utilize when formulating their expectations. Uniformed traders are aware (in our framework) that HFTs and slow informed traders view the same private signal, albeit with different quality. They are, however, unable to extract σ_S^2 and σ_F^2. Moreover, linearity dictates that they can only consider a signal s with precision being a linear combination of σ_S^2 and σ_F^2, i.e. the precision of slow informed traders and HFTs, correspondingly. Herein, we assume that uniformed traders form their expectations based on a weighted average precision

$$\sigma_U^2 = Var_U(e) = \lambda\sigma_F^2 + (1-\lambda)\sigma_S^2, \tag{16}$$

with $\lambda \in [0,1]$. In other words, uniformed traders conjecture that there exists a signal with (weighted) average levels of noise.

Based on the above, the conditional expectation and variance of u can be calculated. In view of the theorem of projection for normal variables, the following conditional moments are obtained:

$$E(u|p = \alpha s + \beta Q) = \frac{\alpha\sigma_u^2}{\alpha^2\sigma_u^2 + \alpha^2\sigma_U^2 + \beta^2\sigma_Q^2} p \tag{17}$$

and

$$Var(u|p = \alpha s + \beta Q) = \sigma_u^2 \left(\frac{\alpha^2\sigma_U^2 + \beta^2\sigma_Q^2}{\alpha^2\sigma_u^2 + \alpha^2\sigma_U^2 + \beta^2\sigma_Q^2} \right). \tag{18}$$

The conditional expectations (and variances) for the informed traders are more involved but the joint normality of u and e substantially simplifies the algebra. The final relations read:

$$E_{F,S}(u|s,p) = s\frac{\sigma_u^2}{\sigma_u^2 + \sigma_{F,S}^2} = \gamma_{F,S}s = E_{F,S}(u|s) \tag{19}$$

and

$$Var_{F,S}(u|s,p) = \sigma_{F,S}^2\frac{\sigma_u^2}{\sigma_u^2 + \sigma_{F,S}^2} = \gamma_{F,S}\sigma_{F,S}^2 = Var_{F,S}(u|s) \tag{20}$$

where $\gamma_{F,S} = \sigma_u^2/(\sigma_u^2 + \sigma_{F,S}^2)$ and, as before, F and S correspond to fast HFT and slow informed investors.

The equalities $E_{F,S}(u|s,p) = E_{F,S}(u|s)$ and $Var_{F,S}(u|s,p) = Var_{F,S}(u|s)$ that stem from equations (19) and (20) do not contradict (10), where it was a priori assumed that HFTs and slow informed investors form their expectations based on both variables (s and p). These equalities are a consequence of the linear price rule and the fact that informed traders observe the same signal, albeit with different precision.

[13] More generally, one should include an additive constant in (15) as done, for instance, in García and Urošević (2013). However, due to the fact that $E(u) = 0$, this constant turns out to be zero and for this reason it has been omitted.

The market condition to calculate the clearing price is that the excess demand is zero:

$$KQ_F + MQ_S + LQ_U + ZQ = 0 \tag{21}$$

Using the conditional expectations derived earlier, equation (21) becomes, after some simplification,

$$K\frac{\gamma_F s - p}{\omega_F + \gamma_F A\,\sigma_F^2} + M\frac{\gamma_S s - p}{\omega_S + \gamma_S A\,\sigma_S^2} + L\frac{(\alpha-\alpha^2)\sigma_u^2 - \alpha^2\sigma_U^2 - \beta^2\sigma_Q^2}{\omega_U(\alpha^2\sigma_u^2 + \alpha^2\sigma_U^2 + \beta^2\sigma_Q^2) + A\,\sigma_u^2(\alpha^2\sigma_U^2 + \beta^2\sigma_Q^2)}p + ZQ = 0. \tag{22}$$

Next, we rearrange equation (22) to acquire

$$\begin{aligned}\left(\frac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2} + \frac{M}{\dfrac{\omega_S}{\gamma_S} + A\,\sigma_S^2}\right)s + ZQ = &-L\frac{(\alpha-\alpha^2)\sigma_u^2 - \alpha^2\sigma_U^2 - \beta^2\sigma_Q^2}{\omega_U(\alpha^2\sigma_u^2 + \alpha^2\sigma_U^2 + \beta^2\sigma_Q^2) + A\,\sigma_u^2(\alpha^2\sigma_U^2 + \beta^2\sigma_Q^2)}p \\ &+ \left(\frac{K}{\omega_F + \gamma_F A\sigma_F^2} + \frac{M}{\omega_S + \gamma_S A\,\sigma_S^2}\right)p,\end{aligned} \tag{23}$$

which is equivalent to

$$\left(\frac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2} + \frac{M}{\dfrac{\omega_S}{\gamma_S} + A\,\sigma_S^2}\right)s + ZQ = Rp = R(\alpha s + \beta Q) = R\alpha s + R\beta Q, \tag{24}$$

for

$$R = -L\frac{(\alpha-\alpha^2)\sigma_u^2 - \alpha^2\sigma_U^2 - \beta^2\sigma_Q^2}{\omega_U(\alpha^2\sigma_u^2 + \alpha^2\sigma_U^2 + \beta^2\sigma_Q^2) + A\,\sigma_u^2(\alpha^2\sigma_U^2 + \beta^2\sigma_Q^2)} + \left(\frac{K}{\omega_F + \gamma_F A\sigma_F^2} + \frac{M}{\omega_S + \gamma_S A\,\sigma_S^2}\right). \tag{25}$$

At the final step of our analysis, we impose rational expectations by matching the coefficients on the left and the right hand sides of equation (24).[14] By doing so, we immediately obtain the following relations,

$$R = Z/\beta, \quad R\alpha = \frac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2} + \frac{M}{\dfrac{\omega_S}{\gamma_S} + A\,\sigma_S^2}, \quad \frac{\alpha}{\beta} = \frac{K}{Z\left(\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2\right)} + \frac{M}{Z\left(\dfrac{\omega_S}{\gamma_S} + A\,\sigma_S^2\right)}. \tag{26}$$

Since in view of conditions (26) α is a linear function of β, for the determination of these coefficients it suffices to solve either of the equalities emerging from the coefficient matching. We opt to solve the

[14]Realizations coincide with expectations in equilibrium.

algebraic equation for the coefficient of s,

$$\left(\frac{K}{\dfrac{\omega_F}{\gamma_F}+A\sigma_F^2}+\frac{M}{\dfrac{\omega_S}{\gamma_S}+A\,\sigma_S^2}\right)=R\alpha, \tag{27}$$

which, using equation (25), can be written as

$$\begin{aligned}\left(\frac{K}{\dfrac{\omega_F}{\gamma_F}+A\sigma_F^2}+\frac{M}{\dfrac{\omega_S}{\gamma_S}+A\,\sigma_S^2}\right)=&-L\frac{(\dfrac{1}{\alpha}-1)\sigma_u^2-\sigma_U^2-\left(\dfrac{\beta}{\alpha}\right)^2\sigma_Q^2}{\omega_U(\sigma_u^2+\sigma_U^2+\left(\dfrac{\beta}{\alpha}\right)^2\sigma_Q^2)+A\,\sigma_u^2(\sigma_U^2+\left(\dfrac{\beta}{\alpha}\right)^2\sigma_Q^2)}\alpha\\ &+\left(\frac{K}{\omega_F+\gamma_F A\sigma_F^2}+\frac{M}{\omega_S+\gamma_S A\,\sigma_S^2}\right)\alpha.\end{aligned} \tag{28}$$

In equation (28), we have divided each term of the first part of the right–hand side by α^2 ($\alpha\neq 0$). Next, we introduce the third relation of (26) into (28) and upon setting

$$C_1=\frac{K}{\dfrac{\omega_F}{\gamma_F}+A\sigma_F^2}+\frac{M}{\dfrac{\omega_S}{\gamma_S}+A\sigma_S^2} \tag{29}$$

$$C_2=\frac{K}{\omega_F+A\gamma_F\sigma_F^2}+\frac{M}{\omega_S+A\gamma_S\sigma_S^2} \tag{30}$$

we obtain,

$$\frac{C_1}{L}=\frac{(\alpha-1)\sigma_u^2+\alpha\sigma_U^2+\alpha\left(\dfrac{Z}{C_1}\right)^2\sigma_Q^2}{\omega_U\left(\sigma_u^2+\sigma_U^2+\left(\dfrac{Z}{C_1}\right)^2\sigma_Q^2\right)+A\,\sigma_u^2\left(\sigma_U^2+\left(\dfrac{Z}{C_1}\right)^2\sigma_Q^2\right)}+\frac{C_2}{L}\alpha.$$

This is a first order equation with respect to α with the following (unique) solution:

$$\alpha=\frac{\dfrac{C_1}{L}+\dfrac{\sigma_u^2}{\omega_U\left(\sigma_u^2+\sigma_U^2+\left(\dfrac{Z}{C_1}\right)^2\sigma_Q^2\right)+A\,\sigma_u^2\left(\sigma_U^2+\left(\dfrac{Z}{C_1}\right)^2\sigma_Q^2\right)}}{\dfrac{\sigma_u^2+\sigma_U^2+\left(\dfrac{Z}{C_1}\right)^2\sigma_Q^2}{\omega_U\left(\sigma_u^2+\sigma_U^2+\left(\dfrac{Z}{C_1}\right)^2\sigma_Q^2\right)+A\,\sigma_u^2\left(\sigma_U^2+\left(\dfrac{Z}{C_1}\right)^2\sigma_Q^2\right)}+\dfrac{C_2}{L}}. \tag{31}$$

Using the first two relations of (26), β is given by the following expression

$$\beta = \alpha \times \frac{Z}{C_1}, \tag{32}$$

and the price p is uniquely determined. □

3.2 Characterization of price efficiency

In this Section we provide the definition of price efficiency as well as its characterization with respect to the model parameters described in Section 3.1.[15]

Proposition 3.2. *Price efficiency is given by*

$$Var(u|p = \alpha s + \beta Q)^{-1} = \frac{1}{\sigma_u^2} + \frac{1}{\sigma_U^2 + (\beta/\alpha)^2\sigma_Q^2}, \tag{33}$$

with

$$\frac{\beta}{\alpha} = \frac{Z}{\dfrac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2} + \dfrac{M}{\dfrac{\omega_S}{\gamma_S} + A\sigma_S^2}}. \tag{34}$$

Proof. Price efficiency can be derived by re–writing equation (18) to as follows,

$$\begin{aligned} &Var(u|p = \alpha s + \beta Q)^{-1} = \\ \left[\sigma_u^2\left(\frac{\alpha^2\sigma_U^2 + \beta^2\sigma_Q^2}{\alpha^2\sigma_u^2 + \alpha^2\sigma_U^2 + \beta^2\sigma_Q^2}\right)\right]^{-1} &= \frac{1}{\sigma_u^2} + \frac{1}{\sigma_U^2 + (\beta/\alpha)^2\sigma_Q^2}, \end{aligned} \tag{35}$$

with

$$\frac{\beta}{\alpha} = \frac{Z}{\dfrac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2} + \dfrac{M}{\dfrac{\omega_S}{\gamma_S} + A\sigma_S^2}}. \tag{36}$$

Equation (35) represents the informativeness of p about the liquidation value, u, of the asset. □

Although the relations involving α and β are quite cumbersome, equation (35) for price efficiency, which is the focus of our study, is quite simple and can be analyzed as follows:

[15] A list of the model parameters is presented in Appendix A at the end of the manuscript.

1) Price efficiency with respect to K and M:

As the number of HFTs, K, increases in the market, price efficiency improves $\big(\partial Var(u|p)^{-1}/\partial K > 0\big)$.[16] Accordingly, as the number of human informed traders, M, increases, clearing prices reflect more information $\big(\partial Var(u|p)^{-1}/\partial M > 0\big)$. To keep the analysis tractable and for the sake of space preservation, we present analytical formulas for the partial derivatives of $Var(u|p)^{-1}$ with respect to the model parameters in the Internet Appendix (attached at the end of the manuscript).

We recall that HFTs view the same private signal as privately informed NON HFTs but with greater precision. However, in view of our assumption that both HFTs and NON HFTs can act strategically, we cannot conclude that price efficiency improves faster with K increased as opposed to M increased. Rather, there is an interesting interplay that is quantified by the following proposition.

Proposition 3.3. *For price efficiency given by* (35)*, the following hold true,*

i) If the equilibrium is competitive ($\omega_F = \omega_S = 0$) then

$$\partial Var(u|p)^{-1}/\partial K > \partial Var(u|p)^{-1}/\partial M, \tag{37}$$

i.e. the rate at which information is incorporated into prices with respect to K will always be greater than the corresponding rate for slow informed traders.

ii) If the equilibrium is strategic then, if

$$\omega_F > \frac{A(\sigma_S^2 - \sigma_F^2)\sigma_u^2}{\sigma_F^2 + \sigma_u^2} + \omega_S \frac{\sigma_S^2 + \sigma_u^2}{\sigma_F^2 + \sigma_u^2}, \tag{38}$$

we have that

$$\partial Var(u|p)^{-1}/\partial K < \partial Var(u|p)^{-1}/\partial M \tag{39}$$

i.e. the rate at which information is incorporated into prices with respect to K is less than the rate at which information is induced into prices with respect to M.

[16] Without loss of generality, to calculate the partial derivative of price efficiency with respect to K (or M, Z), we extend the domain of K (M, Z) from $\mathbb{N}$ to $\mathbb{R}$.

Proof. We treat (i) and (ii) separately. As regards (i), using the analytical formulas from the Internet Appendix, it is simple to show that

$$\frac{\dfrac{\partial Var(u|p)^{-1}}{\partial M}}{\dfrac{\partial Var(u|p)^{-1}}{\partial K}} = \frac{A\,\sigma_F^2 + \dfrac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}}{A\,\sigma_S^2 + \dfrac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}} < 1, \tag{40}$$

since $\sigma_F^2 < \sigma_S^2$ and $\omega_F = \omega_S = 0$.

Case (ii) is a little more involved. Again, using the analytical formulas for the partial derivatives, we start from the following inequality

$$\frac{\dfrac{\partial Var(u|p)^{-1}}{\partial M}}{\dfrac{\partial Var(u|p)^{-1}}{\partial K}} = \frac{A\,\sigma_F^2 + \dfrac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}}{A\,\sigma_S^2 + \dfrac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}} > 1, \tag{41}$$

which can be written as

$$A\,\sigma_F^2 + \frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2} > A\,\sigma_S^2 + \frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}. \tag{42}$$

After some simple algebraic manipulations we obtain

$$\omega_F > \frac{A(\sigma_S^2-\sigma_F^2)\sigma_u^2}{\sigma_F^2+\sigma_u^2} + \omega_S\frac{\sigma_S^2+\sigma_u^2}{\sigma_F^2+\sigma_u^2}, \tag{43}$$

with

$$\frac{\sigma_S^2+\sigma_u^2}{\sigma_F^2+\sigma_u^2} > 1 \tag{44}$$

since

$$\sigma_S^2 > \sigma_F^2. \tag{45}$$

□

Therefore, in a strategic equilibrium, under condition (43) the rate at which information is induced into prices as the number of HFTs (K) increases is lower compared to the corresponding rate for NON HFTs.

2) Price efficiency with respect to $\omega_{F,S}$:

As the price impact of informed traders increases, price efficiency diminishes $\left(\partial Var(u|p)^{-1}/\partial\omega_{F,S} < 0\right)$. In other words, when traders consider the effect of their demands on the price formation process, they are willing to trade less aggressively and, therefore, information is incorporated into prices with a smaller rate. By contrast, when traders increase competition ($\omega_{F,S}$ decreases), information is incorporated into prices faster (Rindi, 2008); see, Figure 1. In the extreme case where ω is equal to zero for all types of traders, we obtain the pure competitive equilibrium.

To demonstrate the result established in Proposition 3.3, Figure 2 plots price efficiency with respect to ω_F for a NON HFT market ($K = 0, M = 30$) and for the following HFT markets ($M = 0, K = 30, 50, 100$). From equation (43) we calculate that the breakpoint for the markets ($K = 0, M = 30$) and ($M = 0, K = 30$) is $\omega_F = 1.5364$ for the model parameters fixed in the example (see Figure 2). As ω_F increases in the HFT market ($M = 0$) price efficiency drops, however it is greater compared to the NON HFT market ($K = 0$) until the breakpoint is reached. After the breakpoint is reached, HFTs become significantly more strategic compared to NON HFTs and, thus, the information contained in their quotes is induced into prices at a smaller rate compared to the case of NON HFTs. In such a scenario, the amount of HFTs should be significantly larger compared to that of NON HFTs in order for the improved precision of HFTs' signal to have a greater effect on price efficiency. This can be observed in Figure 2 where by increasing the number of HFTs, K, in the market price efficiency drops at a smaller rate with ω_F increased.

It is worth noting that when the equilibrium is purely competitive ($w_F = 0$ and $w_S = 0$) and traders are risk neutral ($A = 0$) then price efficiency is constant and equal to $1/\sigma_u^2 + 1/\sigma_U^2$, since $\beta/\alpha = 0$. By contrast, when $A = 0$ and traders are strategic ($w_F \neq 0$ and/or $w_S \neq 0$) price efficiency differs significantly from $1/\sigma_u^2 + 1/\sigma_U^2$. This feature can be observed in Figure 3 which plots price efficiency against K and M for different levels of ω_F and for $A = 0$. On the other hand, while $A = 0$ and $w_F \neq 0$ and/or $w_S \neq 0$, if we let $K \to +\infty$ and $M \to +\infty$ (i.e., as in a market with a continuum of traders) then price efficiency converges to $1/\sigma_u^2 + 1/\sigma_U^2$; that is, the strategic case collapses to the competitive case. Thus, when market participants are few (finite) and risk neutral (or close to risk neutral), the competitive equilibrium can deviate significantly from the true strategic equilibrium, whereas only when the market is large enough the true strategic equilibrium is well approximated by the competitive one (Kovalenkov and Vives, 2014).

[Figures 1, 2 and 3 around here]

3) Price efficiency with respect to $\sigma^2_{F,S}$:

Prices become more efficient as the accuracy of HFTs' signal increases; that is, as σ_F^2 decreases price informativeness improves $\left(\partial Var(u|p)^{-1}/\partial\sigma_F^2 < 0\right)$. Symmetrically, the same result holds for slow informed traders $\left(\partial Var(u|p)^{-1}/\partial\sigma_S^2 < 0\right)$. Figure 4 illustrates the effect of σ_F^2 decreased on price efficiency. Notice, also, that in this particular example the market is competitive ($\omega_F = \omega_S = 0$) and thus price efficiency improves faster as the number of HFTs increases, compared to NON HFTs.

[Figure 4 around here]

4) Price efficiency with respect to σ_U^2:

As σ_U^2 increases (decreases), price efficiency diminishes (increases) $\big(\partial Var(u|p)^{-1}/\partial\sigma_U^2 < 0\big)$. Thus, price efficiency depends on the weight (λ) that uninformed traders place on the signals of privately informed HFTs and NON HFTs, when forming their expectations about the price; see, Figure 5.

[Figure 5 around here]

5) Price efficiency with respect to A:

The risk aversion coefficient A is inversely correlated with market efficiency $\big(\partial Var(u|p)^{-1}/\partial A < 0\big)$. When traders are more/less risk averse (that is, the risk aversion coefficient A increases/decreases) market efficiency is lower/higher. This result is natural in the sense that bids become less/more aggressive and, therefore, information is impounded into prices with lower/higher speed. This is illustrated in Figure 6 which shows price efficiency, $Var(u|p)^{-1}$, as a function of the number of informed traders in the market (M and K) and for different levels of risk aversion. Notice, also, that the effect of privately informed HFTs and NON HFTs on price efficiency is asymmetric due to the difference in the precision of the signals and the fact that $\omega_F = \omega_S = 0$. In this respect, information is incorporated into prices faster when HFTs are present in the market; consider, for example, the marginal case of a HFT market ($M = 0$ and $K \neq 0$) compared to a market with only human informed traders ($M \neq 0$ and $K = 0$).

[Figure 6 around here]

6) Price efficiency with respect to Z:

Noise trading (i.e., Z) is inversely correlated with price efficiency $\big(\partial Var(u|p)^{-1}/\partial Z < 0\big)$ (Grossman and Stiglitz, 1976, 1980). Figure 7 plots $Var(u|p)^{-1}$ as a function of the number of informed traders in the market (M and K) and for different levels of noise. Price efficiency diminishes significantly when noise trading increases. Observe, again, that the effect of noise trading on price efficiency is asymmetric considering the presence of HFTs (K) and NON HFTs (M) in the market, as the simulation corresponds to a competitive equilibrium ($\omega_F = \omega_S = 0$) and $\sigma_F^2 < \sigma_S^2$. Therefore, information is induced into prices faster through HFTs' quotes.

Note that, either in the competitive or in the strategic framework, the level of price efficiency is always greater when both types of traders are present in the market ($K \neq 0, M \neq 0$), compared to price efficiency in a pure HFT market ($M = 0, K \neq 0$) or a pure NON HFT market ($K = 0, M \neq 0$), under the same market conditions. Formally, the following inequalities hold for all values of K and M:

$$Var(u|p = \alpha s + \beta Q)^{-1} =$$
$$\frac{1}{\sigma_u^2} + \frac{1}{\sigma_U^2 + \left(\dfrac{Z}{\dfrac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2} + \dfrac{M}{\dfrac{\omega_S}{\gamma_S} + A\sigma_S^2}} \right)^2 \sigma_Q^2} > \frac{1}{\sigma_u^2} + \frac{1}{\sigma_U^2 + \left(\dfrac{Z}{\dfrac{M}{\dfrac{\omega_S}{\gamma_S} + A\sigma_S^2}} \right)^2 \sigma_Q^2} \tag{46}$$

and

$$Var(u|p = \alpha s + \beta Q)^{-1} =$$
$$\frac{1}{\sigma_u^2} + \frac{1}{\sigma_U^2 + \left(\dfrac{Z}{\dfrac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2} + \dfrac{M}{\dfrac{\omega_S}{\gamma_S} + A\sigma_S^2}} \right)^2 \sigma_Q^2} > \frac{1}{\sigma_u^2} + \frac{1}{\sigma_U^2 + \left(\dfrac{Z}{\dfrac{K}{\dfrac{\omega_F}{\gamma_F} + A\sigma_F^2}} \right)^2 \sigma_Q^2}, \tag{47}$$

given that all the other parameters of the model are fixed. This feature is of particular importance when the level of noise (Z) is considerably elevated and market participants are few. Consider, for example, Figure 8 which plots $Var(u|p)^{-1}$ as a function of the number of informed traders in the market (M and K) and for different levels of elevated noise. In the case where $Z = 100$, price efficiency reaches its pick when $M = K = 30$, whereas for ($M = 0, K = 30$) or ($M = 30, K = 0$) price efficiency is significantly less.

[Figures 7 and 8 around here]

3.3 Discussion

In the present Section we have presented a theoretical background that describes the bidding activity in a transparent order–driven single call market. The trading public consists of fast informed HFTs, slow (human) informed, uninformed and liquidity traders. HFTs are able to view the same private signal as human informed traders but with greater precision due to their ability to collect information from various data sources. The model predicts that HFTs accelerate price discovery and price efficiency due to the increased information content of their quotes, to the extent, however, that they do not behave strategically.

In the subsequent analysis, we are interested to examine empirically the predictions of our model on price efficiency, using an intraday data set from the order driven Paris stock market opening call. The Paris preopening period is transparent and, therefore, market participants are able to observe the evolution of the indicative price–volume pair and, in turn, to submit or adjust their orders. Thus, the order accumulation process can be viewed as a sequence of theoretical market clearings where rational agents drive the price to its equilibrium value as time evolves. Consequently, the preopening period constitutes a natural laboratory to examine our model findings.

Our empirical study is closely related to that of Biais et al. (1999), who find that preopening indicative prices in the Paris Bourse become more informative as the clearing time approaches; the authors refer to this feature as a 'learning pattern'. In this respect, they argue that investors learn from each other by observing the preopening order flow, thus submitting more informative orders as the opening time approaches. On the other hand, because bidding is only theoretical in the preopening, strategic trading may affect negatively the price discovery process by inducing noise into prices. Herein, we are interested to examine the presence of learning patterns and, more important, to investigate the magnitude of the effect of HF quotes on the informativeness of preopening and opening prices in the Paris market. According to our model findings, call auction prices should predict better the true value of the asset when competitive HFTs are active in the market, whereas they should reflect noise when HFTs become strategic.

4 An empirical application

4.1 The Paris Euronext market

Trading at the Euronext Paris platform is conducted in two main ways: a) the order–driven market model and b) the LP quote driven market model. The former is a continuous trading model where liquidity is provided through individual brokers' orders with the enhancement of Supplementary Liquidity Providers (SLPs), whereas the latter concerns securities which are traded continuously via the quotation of designated Liquidity Providers (LPs).

The order–driven system includes either continuous or periodic auction trading. The first mechanism concerns the more liquid securities, like those comprising the CAC 40 Index, whereas the second is for the less liquid securities. The continuous double auction mechanism, examined herein, is operated under the following daily time (CET) schedule:[17]

1. 07:15–09:00 Preopening phase – Order accumulation period
2. 09:00 Opening auction
3. 09:00–17:30 Main trading session: Continuous session

[17] The trading day schedule can be found at: `https://www.euronext.com/en/trading-calendars-hours`

4. 17:30–17:35 Preclosing phase – Order accumulation period
5. 17:35 Closing auction
6. 17:35–17:40 Trading at the last phase (at the close)
7. 17:40–07:15 After hours trading

The opening call auction procedure lasts 1 hour and 45 minutes. During this time period investors are allowed to submit, modify, or cancel orders, while observing the disclosed information on the evolution of the indicative clearing price–volume pair and the prevailing bid–ask quotes. Since trading is absent, all orders are stored into the central limit order book with price–time execution priority. Three main types of orders are allowed during the preopening period: a) market on opening orders, b) pure market orders and c) limit orders.[18] Also, market on opening orders and pure market orders have priority against limit orders at the time of the auction. Figure 9 illustrates the formulation of crossing supply and demand lines, due to the absence of trade, for a hypothetical set of limit prices and quantities. Notice that because stock prices are discrete, it is possible that more than one equilibrium values are present.

[Figure 9 around here]

After the end of the accumulation period, the electronic system considers the supply–demand schedule formed by the queuing orders, seeking for the price that maximizes the trading volume; that is, the equilibrium value. If the maximum volume principle suggests more than one equilibrium prices, then the opening price is set according to the minimum volume surplus principle. Lastly, if more than one prices satisfy the minimum surplus principle, then the system fixes as the opening price the one that is closer to the reference price; the latter is usually the price of the last trade before the preopening period. After the opening price is set, buy and sell orders are matched and executed in a single trade and at a single opening price. Unexecuted market or limit orders are sent forth to the main session with the original price and time priority; market orders are stored as limit orders at the opening price.

It is important to mention that the opening time in the Paris market was fixed at 09:00 until the 15/08/2015, whereas after 15/08/2015 the uncrossing takes place randomly between 09:00:00 and 09:00:30.

4.2 Data sample

The data sample used in the present study is retrieved from the AMF–BEDOFIH Paris Euronext high frequency database and includes 36 stocks from the CAC 40 Index composition on 3 January 2013 and for

[18]Because they do not include price preference, market and on open sell (buy) orders are aggregated at the best ask (bid).

the entire year 2013 (254 trading days).[19] For four stocks transactions take place on the Paris platform, whereas their negotiation is conducted on other platforms (i.e., the securities are dual listed) (see Table 2). We have excluded these stocks from our sample as the database does not contain information on the related quoting activity.

The sample encompasses two main files: i) trades and ii) orders. The first contains information about the trading history in the Euronext Paris market. More specifically, the data set includes information about the time (accurate to the microsecond), the price and the quantity of negotiations. The second includes information about the order placement history; time of submission, price, size, side, duration, type, validity and time of release from the system (either because of execution or because of cancellation).

It is worth mentioning that in our sample period (2013) the opening time is fixed at 09:00 (see, also, Section 4.1). This feature is rather important for investors as it allows them to adjust or cancel their orders literally just before the opening, without fearing of having them executed at an undesirable clearing price; see, for instance, Medrano and Vives (2001) for a detailed analysis of the effect of a random opening time on investors' order placement strategies.[20]

An important aspect of the data set, crucial to our work, is a unique HFT flag that accompanies order and trade messages. In particular, and in line with the AMF documentation (see, Megarbane et al. (2017)), each message is categorized according to the following list:

a) HFT: High Frequency trader
b) NON HFT: Non–High Frequency trader
c) MIXED: Mixed trader (investment bank account applying HFT)

The definition of the HFT flag, as provided by AMF, is based on the following two conditions:

1) A market participant (with a specific ID) is classified as a HFT if: i) the average lifetime of her canceled orders is less than the average lifetime of all canceled orders in the book, and ii) she has canceled at least 100,000 orders during the year.

2) A market participant (with a specific ID) is classified as a HFT if: i) he has canceled at least 500,000 orders with a lifetime of less than 0.1 second and ii) the top percentile of the lifetime of his canceled orders is less than 500 microseconds.

[19]The Paris market data included in the BEDOFIH database are provided by the AMF. Further Information on the BEDOFIH database can be found at: `https://www.eurofidai.org/en`. Note, also, that the Euronext Paris AMF–BEDOFIH database has been used recently by Boussetta et al. (2016) and Bellia et al. (2017) in their empirical analyses.

[20]For additional details on the Euronext Paris stock market opening and main sessions, see, also, Biais et al. (1999) and Biais et al. (1995), respectively.

Note that, according to the AMF classification, the market participant is classified as a HFT if she meets one of these conditions. If, additionally, the participant operates on behalf of a large investment bank, she is classified as a MIXED trader. It is important to mention that once a trader is classified as HFT, NON HFT or MIXED, this flag is immutable.

4.3 Preliminary analysis

4.3.1 Order flow during the preopening

This Section provides a preliminary analysis of the preopening order flow and opening trading activity in the Paris market to motivate further the empirical analysis of the present study.

[Table 1 around here]

Table 1 reports summary statistics about submissions, modifications and cancellations during the preopening period in 15 minute frequency and for each trader type; HFT, MIXED and NON HFT. The first thing to notice is that the activity of HFTs and MIXED traders is rapidly increased after 08:30. For HFTs, the across stocks and days average number of submissions in the intervals 08:30–08:45 and 08:45–09:00 is 220.8 and 112.6, respectively, whereas earlier it is practically zero. For MIXED traders, order submissions are significantly elevated between 08:45 and 09:00; the corresponding (stock–day) average number of submissions is 129.9. On the other hand, human traders (NON HFTs) enter the market earlier in the morning as significant bidding starts at 07:15. Similar to HFTs and MIXED traders, however, NON HFTs increase their bidding activity between 08:30 and 08:45.

Regarding modifications and cancellations, HFTs and MIXED traders sharply increase their activity before the opening, in the interval between 08:45 and 09:00; the average number of modifications and cancellations is, respectively, 36.6 and 30.6 for HFTs and 58.9 and 33.3 for MIXED traders. Human traders also increase their modification activity minutes before the opening (17.5 modifications on average), whereas their cancellations are less compared to those submitted by HFTs and MIXED traders (6 on average).[21]

Interestingly, a significant increase in the average order size is observed for all types of traders just before the opening time. For example, HFTs increase the order size from 192.1 to 667.9 shares from 08:30 to 09:00. Accordingly, MIXED and human traders increase their bidding intensity from 1,824.3 and 779.8 to 5,938.6 and 1,373.8 shares, respectively, from 08:30 to 09:00.

[21] Our summary results on the HFTs' order flow are consistent with those reported by Bellia et al. (2017) for the Paris preopening period and for the CAC 40 sample in year 2013. The authors report that HFTs enter the call market after 08:30, whereas they increase their order placement activity just before the clearing time.

A few comments are in order regarding the summary statistics on the order flow. First, all types of traders act, to some extent, strategically by submitting orders earlier in the preopening and then modifying and/or canceling their bids prior to opening. This feature implies a possible deterioration in the price discovery process, consistent with the prediction of our model (i.e., ω increased, price efficiency decreased). Second, the increase in the average number of order submissions as well as the average order size a few minutes before the clearing time, indicates that traders submit true orders with the intention of having them executed. Thus, such orders should convey significant information which, in turn, is incorporated into prices. For example, in Easley and O'Hara (1987) investors increase the size of their order to profit from their private information. Accordingly, our model predicts that price efficiency should increase drastically before the opening (competitive bidding).

It is important to note that the percentage of quotes placed by designated Supplementary Liquidity Providers during the preopening is less than 2%, according to our computations, which is practically negligible. Hence, we postulate that the rest of traders (HFTs and/or NON HFTs) entering the market before the clearing time are willing to trade for individual and not for market making purposes.

Given that our overall purpose is to investigate the effect of HFTs on the informational content of auction prices, for the remaining of our empirical analysis we shall often consider HFT and MIXED flagged quotes as a single group, the HFT group. Further, our focus will be placed on the last 30 minutes of the preopening (08:30–09:00) when HFTs enter the market rapidly.[22]

4.3.2 The last 30 minutes

This Section focuses on the last 30 minutes of the preopening. In particular, we provide some visual evidence on the order flow and market depth characteristics attributed to the three types of traders in our data set. Figure 10 illustrates the evolution of the across days and stocks average number of submissions, modifications and cancellations for each type of trader (HFT, MIXED and NON HFT) in 1 minute time intervals. HFTs enter the market rapidly at 08:30 with a stock–day average number of submissions equal almost to 140. Submissions for human traders are also elevated at 08:30, whereas MIXED traders exhibit a very low rate of submissions at that point in time.

[Figure 10 around here]

After 08:55, that is 5 minutes until the opening, all types of traders increase their order submission activity.

[22] We have also conducted our empirical analysis over the entire preopening period and results indicate that before 08:30 there is considerable noise in the market.

The same holds for order modifications and cancellations, especially for HFTs and MIXED traders. The implementation of frequent submit–cancel strategies by HFTs has been well–documented in the microstructure literature, although exclusively in the continuous double auction mechanism (e.g., Chlistalla, 2012; O'Hara, 2015). Accordingly, we postulate for the call auction mechanism that HFTs increase their cancellations and modifications just before the clearing time, possibly taking advantage of their speed, while re–submitting larger and more aggressive orders with the intention of having them executed.

[Figure 11 around here]

Figure 11 depicts the evolution of the market depth that is attributed to HFTs (here, HFT and MIXED shares are altogether) and NON HFTs during the last 30 minutes of the order accumulation process. An interesting insight is revealed by the top graph, regarding the percentage of HFT shares standing on the book. Both on the buy and the sell side, NON HFT shares are the vast majority during the most of the preopening time period, whereas the inverse is observed after 08:57. Further, and according to the middle graph, this change on the percentage of standing shares is attributed to the fact that HFTs increase sharply their participation in shares on the order book. Indeed, according to Table 1, the average number of submissions as well as the average order size are significantly increased at that time of the preopening. Hence, we infer that HFTs become more aggressive before the opening, submitting orders with the intention of having them filled. The latter argument is further corroborated by the bottom graph in Figure 11 which illustrates the percentage of shares on the executable side of the book, attributed to HFTs.[23] Prior to opening, the percentage of aggressive HFT shares on the buy (sell) side is almost 42% (70%). Hence, a significant amount of HFT orders is expected to be executed at the opening trade.

4.3.3 Opening trade

This Section discusses summary statistics at the opening. Specifically, Table 2 reports means for the opening volume, number of trades, value of transactions, percentage of value of transactions relative to total daily value of transactions, as well as HFT, NON HFT and MIXED trading activity, for our CAC 40 stock sample. The opening trade accounts for, approximately, 1.4% of total daily traded volume.[24] This percentage is rather low compared to the 10% reported by Biais et al. (1999) for the opening auction of the

[23]Given that the clearing price is p, executable buy (sell) orders are those with a limit price higher or equal (lower or equal) than p. Also, market and at–the–open orders which have no price preference are considered as infinitely aggressive and thus they have higher execution priority.

[24]Similar to our findings, Boussetta et al. (2016) report that the average percentage of opening volume with respect to the total daily traded volume ranges from 1.3% for large stocks (CAC 40 or CAC Next 20) to 1.9% for smaller stocks in the Paris market and for the years 2012 and 2013.

CAC 40 stocks in 1993. It is consistent, however, with our theoretical analysis where we have chosen to model the clearing market with a finite number of agents.

Concerning the HFT activity, the percentage of HFT and MIXED flagged trades at the opening is significantly increased; almost 51.5%, on average, of total HFT buying activity at the opening involves HFT or MIXED flags, while on the sell side the corresponding percentage is approximately equal to 45%. These findings are in line with the preopening market depth statistics presented in Section 4.3.2. That is, HFTs increase their aggressiveness prior to opening, submitting orders which are expected to be executed.

[Table 2 around here]

4.4 The effect of HFTs' quotes on call auction prices

Having demonstrated that HFTs participate actively in the preopening period, especially during the last 30 minutes before the clearing time, in this Section we address two emerging questions: How large is HFTs' contribution to total price discovery, and, more important, do HFTs' quotes improve price informativeness?

4.4.1 Price discovery during the preopening

To study the contribution of HFTs to the price formation process during the preopening period, we employ the Weighted Price Contribution (WPC) statistic introduced by Barclay and Warner (1993), which is defined as follows:

$$WPC_k = \sum_{i=1}^{N_{stocks}} \left(\frac{|\Delta P_i|}{\sum_{i=1}^{N_{stocks}} |\Delta P_i|} \right) \times \left(\frac{\Delta P_{k,i}}{\Delta P_i} \right), \tag{48}$$

where ΔP_i is the total logarithmic price change for the period of interest (that is, the last 30 minutes of the preopening between 08:30 and 09:00) for stock i ($i = 1, ..., N_{stocks}$) and $\Delta P_{k,i}$ is the logarithmic price change for interval k within the period of interest. The first parenthesis in equation (48) is the weighting factor for each stock to control for potential heteroscedasticity in preopening returns, whereas the second parenthesis is the contribution of interval k to total price adjustment (Barclay and Hendershott, 2003). Note that the WPC statistic, as defined above, is the weighted average price contribution across stocks. Subsequently, this quantity is averaged across days to obtain the overall WPC estimate. Similar results can be acquired by averaging first across days, within equation (48), and then across stocks. In this case, however, it is likely that the common market component in stock returns will add bias to the average estimate of the WPC statistic. Therefore, following Barclay and Hendershott (2003), we select to calculate the WPC statistic day by day and then to average across days.[25]

[25] Bellia et al. (2017) apply a variation of the WPC statistic, the WPDC (Weighted Price Discovery Contribution), for the French stock market. In particular, the authors calculate the WPDC order by order to separate the contribution of HFTs

To separate the contribution of different types of traders on the price discovery process, we consider the following three cases:

- HFT market: order book calculated exclusively on the basis of HFT and MIXED orders that pertain directly to HFT activity.
- NON HFT market: order book calculated on the basis of slow traders' quotes, flagged as NON HFTs.
- Actual market: the full order book.

For each of the three cases, we calculate indicative prices per 1 minute, from 08:30 to 09:00 (opening) and, subsequently, the logarithmic intraday returns $\Delta P_{k,i}$ are used in equation (48).[26] We calculate, however, the 08:30 to 09:00 return, ΔP_i, using prices exclusively from the actual market, in all three cases. By doing so, we opt to proxy the contribution of each type of trader (and for each preopening interval) to total price discovery between 08:30 and 09:00.[27]

[Figure 12 around here]

The upper graph in Figure 12 illustrates the evolution of the cumulative Weighted Price Contribution statistic within the preopening period for the aforementioned three cases. Evidently, HFTs lead the price formation process almost over the entire period. At the end of the preopening, HFT prices can explain approximately 73% of total price discovery, whereas NON HFT price adjustments can explain only 48% of total price discovery.

Interestingly, prices in the HFT market seem to reverse at about 08:56. From the bottom graph in Figure 12 we observe that prior to opening HFTs generate negative returns. This feature can be explained by the fact that HFTs tend to place sell quotes aggressively; almost 70% of aggressive sell quotes just before the clearing time are attributed to HFTs (see Figure 11). Nonetheless, earlier at 08:50 the HFT order book generates excessively positive returns, implying that HFTs tend to place buy quotes more aggressively. Notice, also, the influence of HFTs on the actual returns (the pick in the bottom graph for the FULL

and NON HFTs to total price discovery. Similarly, Ciccotello and Hatheway (2000) employ the WPC statistic per quote to separate the contribution of each NASDAQ dealer to total price discovery. Our approach, however, differs in that we construct tentative prices on the basis of the different sets of orders that pertain to HFT and NON HFT activity and, subsequently, we use them to estimate the standard WPC statistic.

[26] When the market is empty (i.e., when supply does not cross demand) we fix as the indicative price the mid–point value (Hasbrouck, 1991). This feature is rather often for HFTs since before 08:30 their participation in the market is practically negligible. When there are no orders on the buy and/or the sell side we omit the calculation.

[27] In a similar way, Madhavan and Panchapagesan (2000) and Anagnostidis et al. (2015) separate market from limit orders in the NYSE and the Athens Exchange market opening, respectively, to compare the information content of the limit order book against the actual system price which includes price–inelastic market and at–the–open orders.

BOOK case), which tend to be positive at 08:50 on average. The aforementioned evidence indicates a strategic change in the quoting behavior of HFTs which, according to our model, should deteriorate the information content of indicative prices.

4.4.2 Price efficiency

Thus far, our empirical results suggest that HFTs play a leading role in the price discovery process during the preopening, especially toward the clearing time. Further, we find evidence that HFTs act strategically by submitting orders earlier in the preopening which they later adjust or cancel, affecting, therefore, the price formation process. In this Section we investigate the effect of HFTs on price informativeness which is the central question of our study. In particular, we test the main prediction of our model that accurately informed HFTs improve price efficiency when they bid competitively, whereas they induce noise into prices when they behave strategically.

For the purposes of our analysis we employ the unbiasedness regression technique of Biais et al. (1999). To do so, we use the current day closing price as a representation of the fair value of the stock, u. This proxy is based on the assumption that at the end of the trading day prices reflect all market information.[28] Now, consider the information set I_0 at the start of the preopening period (i.e., the start of the trading day). Then, the previous day closing price represents the expected value of the stock at time zero ($t = 0$) conditional on I_0; denote this expectation by $E(u|I_0)$. If traders consider the disclosed information on the indicative price and the order book dynamics, then as time progresses price P_t should become more informative; that is, $P_t = E(u|I_t)$ with I_t being the information set at time $t > 0$. On the other hand, if the price P_t incorporates noise, then it should reflect the information set I_0 plus a noise term η_t. Thus, $P_t = E(u|I_0) + \eta_t$ with $\eta_t \perp u$.

To conduct the econometric test, we initially compute theoretical prices over 1 min frequency in the preopening period for the three cases: HFT market, NON HFT market and the full order book (the actual market). Subsequently, we utilize the following overnight logarithmic returns: 1) previous day close to current day preopening time t, R_{ct} and 2) previous day close to current day close, R_{cc}. The first type proxies the difference between the preopening price P_t and the equilibrium value at the start of the day $E(u|I_0)$, $[P_t - E(u|I_0)]$, whereas the second type proxies the change of the equilibrium price of the stock, $[u - E(u|I_0)]$, after the end of the trading day. The idea is to examine the correlation between the two types of returns over the sample trading days, at each preopening time stamp, to infer on price efficiency.[29]

[28] We have also used the current day opening price as a proxy of the fair value of the asset and results, available upon request, are qualitatively similar to those reported herein. Note, further, that other prices obtained during the trading day can be employed as fair value proxies; mid–day prices for example (Madhavan and Panchapagesan, 2000).

[29] Boussetta et al. (2016) also apply the unbiasedness regression technique to infer price efficiency in the Paris market. Our approach, however, differs in that we construct separate prices on the basis of HFT and NON HFT quotes.

We estimate the following linear regression at each preopening time t and over the sample trading days:

$$R_{cc} = a + bR_{ct} + \epsilon. \tag{49}$$

If the slope coefficient b is equal to unity, then stock preopening tentative prices are efficient as they reflect all market information at time t (I_t). By contrast, if b is different from unity, then indicative prices reflect information plus some noise.[30] In other words, the estimated b coefficient can be viewed as a signal to noise ratio (Barclay and Hendershott, 2003). Notice that the variance of the residual term in equation (49) is comparable to the residual variance obtained in equation (35), after conditioning on the linear rule $p = \alpha s + \beta Q$.

[Figure 13 around here]

Figure 13 plots the average, across stocks and days, estimated b coefficients during the last 30 minutes of the preopening for the three cases: i) Full book, ii) HFT book, iii) NON HFT book (upper graph). For the actual market (the Full book), it is evident that prices gradually incorporate information until they become fully efficient at the opening. According to the reported confidence intervals at the opening, the b coefficient is statistically equal to unity at 09:00.[31] Thus, there is evidence of a 'learning' pattern in the market, where traders adjust their quotes according to the available public and private information. This learning pattern is similar to that reported by Biais et al. (1999) and Boussetta et al. (2016) for the Paris market, as well as by Ciccotello and Hatheway (2000) and Barclay and Hendershott (2003) for the NASDAQ, Comerton-Forde and Rydge (2006) and Moshirian et al. (2012) for the Australian Exchange and Anagnostidis et al. (2015) for the Athens Exchange.[32] On the other hand, early prices are rather noisy, a feature that reflects the fact that the significant bidding activity after 08:30 (see Section 4.3) does not convey information about fundamentals, but rather reflects the strategic behavior of market participants.

[30] Note that stock price efficiency at the opening can also be examined using variance ratio statistics over daily data (e.g., opening and closing prices); see, for example, Stoll and Whaley (1990). The unbiasedness regression technique, however, is advantageous, as it enables us to examine the dynamics of stock prices within the trading day. Further, by running a separate regression for each consecutive time interval within the preopening, we avoid nonstationarity issues that arise due to price adjustments as the price discovery evolves.

[31] Barclay and Hendershott (2003) point out that if stock returns are serially uncorrelated and measured without any errors, then the slope coefficient b in equation (49) should be statistically equal to unity. Otherwise, because of the presence of various microstructure effects that induce correlation into returns (e.g., temporary pricing errors and the non–synchronous trading effect), b should become noisy, deviating from unity. As in Barclay and Hendershott (2003), in the present study we report confidence intervals using the time–series standard errors from the coefficient estimates. We have also replicated the regression analysis using Newey–West standard errors corrected for serial correlation and heteroscedasticity. Results, available upon request for the interested reader, are qualitatively similar with those reported herein.

[32] Boussetta et al. (2016) utilize the SBF 120 Index constituents for years 2012 and 2013, whereas we utilize the CAC 40 Index (for year 2013) which includes the most liquid securities in the French market.

Regarding the HFT market, early orders induce a significant amount of noise, a feature that suggests their strategic behavior. Notice, for example, the sudden drop in price efficiency at 08:50, when the price reversal occurs in the HFT market (see Section 4.4.1). Nonetheless, prior to opening HFTs increase their order aggressiveness (i.e., competition for execution) and, as predicted by our model, they drive the price toward its fair value. The information content of HFT indicative prices increases rapidly within the last two minutes of the preopening, as b approaches unity. At the open, however, it is not statistically equal to unity, suggesting that HFT prices still reflect some noise.

As for NON HFTs, the b coefficient pattern increases monotonically with time, hinting that NON HFTs add information to the price formation process. It remains, however, rather low during the entire 30 minute period, as it ranges approximately from 0.2 to 0.45. Thus, the signal to noise ratio is rather low in the NON HFT market, compared to the HFT market. Notably, around 08:30 NON HFT indicative prices are more efficient compared to HFT prices, although order submission activity for HFTs is significantly greater compared to that of NON HFTs (see, Figure 10). This feature implies further the strategic effect of HFTs; information is induced into prices with a smaller rate via HFTs' orders, compared to those placed by NON HFTs'. This argument is consistent with the model analysis presented in Section 3.2.

The visual evidence on the evolution of the estimated coefficient b for the three types of prices, depicted in the upper graph in Figure 13, provides some first indication on the increased information content of HFTs' quotes and their positive contribution to price efficiency. Nonetheless, a formal econometric test is needed to compare the b coefficients and, therefore, to establish this result. To do so, we cannot use directly the standard errors of the unbiasedness regressions, as in all cases the dependent variable is the close–to–close return and, thus, the b coefficients are correlated. To test formally the difference between the information content of the three types of prices, we first conduct a normality test for the difference between the across stock average estimated b coefficients, for the three different cases: i) $mean(\hat{b}^{HFT})$=$mean(\hat{b}^{NONHFT})$, ii) $mean(\hat{b}^{HFT})$=$mean(\hat{b}^{FULLBOOK})$ and iii) $mean(\hat{b}^{NONHFT})$=$mean(\hat{b}^{FULLBOOK})$ (Barclay and Hendershott, 2003). Figure 13 (middle graph) plots the evolution of the absolute value of the t-statistic during the last 30 minutes of the preopening. Notice that, on average, the full book contains more information compared to the HFT and the NON HFT order books. For most 1 minute time intervals the average $\hat{b}^{FULLBOOK}$ coefficient is statistically different (and greater according to the upper graph of Figure 13) from the rest of the average estimated coefficients at the 5% probability level. We attribute this feature to the fact that the full book aggregates information from both types of informed traders.[33] Nonetheless,

[33]Consider, for instance, Figure 8 from the model analysis where under the presence of considerable noise (e.g., $Z = 100$),

observe the difference between the estimated mean coefficients prior to open and at the open for the HFT and the NON HFT market. This result suggests that HFT orders convey more information compared to NON HFT quotes, consistent with the main assumption of our model that HFTs are able to generate private signals with higher information content compared to the rest of the human trading public. Hence, we infer that HFTs improve significantly price efficiency during the last minutes of the preopening by submitting aggressive and meaningful, in terms of information, orders.

To corroborate our comparative results on price efficiency between HFTs and NON HFTs, we conduct a second test based on the methodology proposed in Madhavan and Panchapagesan (2000). In particular, upon estimating the b coefficient, we estimate a second stage regression using the residuals from the initial regression as the dependent variable. Formally, we estimate the following regressions:

$$\hat{\epsilon}^{NONHFT} = c + dR_{ct}^{HFT} + \epsilon' \tag{50}$$

and

$$\hat{\epsilon}^{HFT} = c + dR_{ct}^{NONHFT} + \epsilon', \tag{51}$$

where $\hat{\epsilon}^{NONHFT}$ ($\hat{\epsilon}^{HFT}$) are the residuals obtained after estimating equation (49) using the NON HFT (HFT) indicative returns, R_{ct}^{NONHFT} (R_{ct}^{HFT}), as independent variable. If the d coefficient is statistically different from zero in equation (50), then HFT prices have additional explanatory power over the NON HFT prices. Conversely, if d is statistically significant in equation (51), then NON HFT prices have significant predictive power beyond HFT prices. Results are plotted in the bottom graph in Figure 13. The percentage of statistically significant d coefficient estimates (from the 36 stock regressions) is consistently higher in the case where we use the R_{ct}^{HFT} returns as an explanatory variable on the residuals from the unbiasedness regressions with the NON HFT returns (first stage). This difference is more apparent at the opening, where in almost 100% of regressions the HFT indicative prices have significant explanatory power on the NON HFT prices. The inverse, however, is not true, since only in 64% of total regressions NON HFT prices have explanatory power on HFT prices. Therefore, we infer that HFT quotes are, up to a certain degree, more informative compared to the NON HFT quotes.

price efficiency is greater when both HFTs and slow informed traders increase their participation in the bidding process (e.g., $K = M = 30$).

5 Conclusions

The present study examines the behavior of High–Frequency Traders (HFTs) in the single call market mechanism. We develop a noisy rational expectations model where HFTs coexist in the market with NON HFTs. In particular, the market includes two types of informed traders: fast HFTs and slow (human) traders. Further, because of their ability to access multiple data streams, HFTs are able to collect information and to form private signals that are more accurate compared to the signals that human informed traders obtain. Our model predicts that when HFTs are present in the market, the information content of prices increases significantly, however, to the extent that HFTs bid competitively. When HFTs become strategic, price efficiency deteriorates.

To test the predictions of our model, we utilize a unique set of high frequency data from the Paris market opening call. Our empirical results suggest that early in the preopening HFTs are absent. As the opening time approaches, however, HFTs increase their bidding activity. Our evidence indicates, also, that HFTs act strategically. In particular, HFTs submit their orders half an hour before the opening time, whereas they modify/cancel a large fraction of these orders just before the opening, while submitting new (or revised) and aggressive orders. Accordingly, and consistent with our model prediction, earlier in the preopening as well as minutes before the opening prices reflect noise (effect of strategic trading). Just before the opening, however, price efficiency increases sharply as HFTs submit aggressive and large orders with the intention of having them executed.

Our results add to a growing part of the microstructure literature that examines the characteristics of HFTs in electronic call markets and the effect that HFTs have on the quality of clearing prices. The call auction mechanism is a useful tool for revealing prices after extensive periods of no trade (e.g., opening of markets) or after periods of market instability (e.g., trading halts) and, therefore, the understanding of HFTs' activities in such a trading environment as well as their effect on the price formation process is of critical importance. Yet, there are several questions to be answered before we reach a consensus on the role of HFTs in call auction trading. Future research could focus on the investigation of potential links between HFTs' order placements strategies in call auction markets and the subsequent continuous trading (e.g., after the opening or after a trading halt).

Acknowledgements

This project has received financial support from the French State under the program 'Investments for the Future' (reference ANR–11–EQPX–0006 – "BEDOFIH"). The authors are grateful to Sonia Jimenez–

Garcés, Terrence Hendershott, Carole Metais and Duc Khuong Nguyen for useful comments on previous versions of the paper, as well as the seminar participants at the 2017 AFFI Conference and at the 2017 Finance and Growth Workshop, ANR FIRE. The authors would also like to thank Clément Michoudet for his research assistance.

References

Admati, A.R., 1985. A noisy rational expectations equilibrium for multi-asset securities markets. Econometrica: Journal of the Econometric Society 53, 629–657.

Anagnostidis, P., Kanas, A., Papachristou, G., 2015. Information revelation in the Greek exchange opening call: Daily and intraday evidence. Journal of International Financial Markets, Institutions and Money 38, 167–184.

Aumann, R.J., 1964. Markets with a continuum of traders. Econometrica: Journal of the Econometric Society 32, 39–50.

Barclay, M.J., Hendershott, T., 2003. Price discovery and trading after hours. Review of Financial Studies 16, 1041–1073.

Barclay, M.J., Hendershott, T., 2008. A comparison of trading and non-trading mechanisms for price discovery. Journal of Empirical Finance 15, 839–849.

Barclay, M.J., Warner, J.B., 1993. Stealth trading and volatility: Which trades move prices? Journal of Financial Economics 34, 281–305.

Baruch, S., Panayides, M., Venkataraman, K., 2017. Informed trading and price discovery before corporate events. Journal of Financial Economics 125, 561–588.

Bellia, M., Pelizzon, L., Subrahmanyam, M.G., Uno, J., Yuferova, D., 2016. Low-Latency Trading and Price Discovery: Evidence from the Tokyo Stock Exchange in the Pre-Opening and Opening Periods. SAFE Working Paper No. 144, 7 June. Available at SSRN 2841242.

Bellia, M., Pelizzon, L., Subrahmanyam, M.G., Uno, J., Yuferova, D., 2017. Coming Early to the Party. SAFE Working Paper No. 182, 20 September. Available at SSRN 3038699.

Biais, B., Bisière, C., Pouget, S., 2013. Equilibrium discovery and preopening mechanisms in an experimental market. Management Science 60, 753–769.

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

Biais, B., Hillion, P., Spatt, C., 1995. An empirical analysis of the limit order book and the order flow in the paris bourse. The Journal of Finance 50, 1655–1689.

Biais, B., Hillion, P., Spatt, C., 1999. Price discovery and learning during the preopening period in the paris bourse. Journal of Political Economy 107, 1218–1248.

Boehmer, E., Fong, K.Y., Wu, J.J., 2015. International evidence on algorithmic trading. Working Paper, 18 September. Available at SSRN 2022034.

Boussetta, S., Lescourret, L., Moinas, S., 2016. The role of pre-opening mechanisms in fragmented markets. Working Paper, 5 May. Available at SSRN 2939502.

Brogaard, J., 2010. High frequency trading and its impact on market quality. Working Paper, 22 November. Northwestern University Kellogg School of Management.

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

Cao, C., Ghysels, E., Hatheway, F., 2000. Price discovery without trading: Evidence from the Nasdaq preopening. The Journal of Finance 55, 1339–1365.

Carrion, A., 2013. Very fast money: High-frequency trading on the NASDAQ. Journal of Financial Markets 16, 680–711.

Cespa, G., Foucault, T., 2013. Sale of price information by exchanges: does it promote price discovery? Management Science 60, 148–165.

Chaboud, A.P., Chiquoine, B., Hjalmarsson, E., Vega, C., 2014. Rise of the machines: Algorithmic trading in the foreign exchange market. The Journal of Finance 69, 2045–2084.

Chlistalla, M., 2012. High-frequency trading: better than its reputation? Research briefing, Deutsche Bank Research.

Chordia, T., Goyal, A., Lehmann, B.N., Saar, G., 2013. High-frequency trading. Journal of Financial Markets 16, 637–645.

Ciccotello, C.S., Hatheway, F.M., 2000. Indicating ahead: Best execution and the Nasdaq preopening. Journal of Financial Intermediation 9, 184–212.

Comerton-Forde, C., Rydge, J., 2006. The influence of call auction algorithm rules on market efficiency. Journal of Financial Markets 9, 199–222.

Conrad, J., Wahal, S., Xiang, J., 2015. High-frequency quoting, trading, and the efficiency of prices. Journal of Financial Economics 116, 271–291.

De Jong, F., Rindi, B., 2009. The microstructure of financial markets. Cambridge University Press.

Easley, D., O'Hara, M., 1987. Price, trade size, and information in securities markets. Journal of Financial Economics 19, 69–90.

Easley, D., O'Hara, M., Yang, L., 2016. Differential access to price information in financial markets. Journal of Financial and Quantitative Analysis 51, 1071–1110.

Economides, N., Schwartz, R.A., 1995. Electronic call market trading. The Journal of Portfolio Management 21, 10–18.

Ellul, A., Shin, H.S., Tonks, I., et al., 2005. Opening and closing the market: Evidence from the london stock exchange. Journal of Financial and Quantitative Analysis 40, 779–801.

Foucault, T., Hombert, J., Roşu, I., 2016. News trading and speed. The Journal of Finance 71, 335–382.

García, D., Urošević, B., 2013. Noise and aggregation of information in large markets. Journal of Financial Markets 16, 526–549.

Grossman, S., 1976. On the efficiency of competitive stock markets where trades have diverse information. The Journal of Finance 31, 573–585.

Grossman, S.J., Stiglitz, J.E., 1976. Information and competitive price systems. The American Economic Review 66, 246–253.

Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. The American Economic Review 70, 393–408.

Hagströmer, B., Norden, L., 2013. The diversity of high-frequency traders. Journal of Financial Markets 16, 741–770.

Hasbrouck, J., 1991. Measuring the information content of stock trades. The Journal of Finance 46, 179–207.

Hasbrouck, J., Saar, G., 2013. Low-latency trading. Journal of Financial Markets 16, 646–679.

Hauser, S., Kamara, A., Shurki, I., 2012. The effects of randomizing the opening time on the performance of a stock market under stress. Journal of Financial Markets 15, 392–415.

Hellwig, M.F., 1980. On the aggregation of information in competitive markets. Journal of Economic Theory 22, 477–498.

Hendershott, T., Jones, C.M., Menkveld, A.J., 2011. Does algorithmic trading improve liquidity? The Journal of Finance 66, 1–33.

Hendershott, T., Riordan, R., 2013. Algorithmic trading and the market for liquidity. Journal of Financial and Quantitative Analysis 48, 1001–1024.

Hillion, P., Suominen, M., 2004. The manipulation of closing prices. Journal of Financial Markets 7, 351–375.

Hirshleifer, D., Luo, G.Y., 2001. On the survival of overconfident traders in a competitive securities market. Journal of Financial Markets 4, 73–84.

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

Jain, P.K., Jain, P., McInish, T.H., 2016. Does high-frequency trading increase systemic risk? Journal of Financial Markets 31, 1–24.

Jordan, J.S., 1982. A dynamic model of expectations equilibrium. Journal of Economic Theory 28, 235–254.

Jovanovic, B., Menkveld, A.J., 2016. Middlemen in limit order markets. Working Paper, 21 June. Available at SSRN 1624329.

Kalay, A., Sade, O., Wohl, A., 2004. Measuring stock illiquidity: An investigation of the demand and supply schedules at the tase. Journal of Financial Economics 74, 461–486.

Kirilenko, A., Kyle, A.S., Samadi, M., Tuzun, T., 2017. The flash crash: High-frequency trading in an electronic market. The Journal of Finance 72, 967–998.

Kovalenkov, A., Vives, X., 2014. Competitive rational expectations equilibria without apology. Journal of Economic Theory 149, 211–235.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica: Journal of the Econometric Society 53, 1315–1335.

Madhavan, A., 2012. Exchange-traded funds, market structure, and the flash crash. Financial Analysts Journal 68, 20–35.

Madhavan, A., Panchapagesan, V., 2000. Price discovery in auction markets: A look inside the black box. Review of Financial Studies 13, 627–658.

Medrano, L.A., Vives, X., 2001. Strategic behavior and price discovery. RAND Journal of Economics 32, 221–248.

Megarbane, N., Saliba, P., Lehalle, C.A., Rosenbaum, M., 2017. The behaviour of high-frequency traders under different market stress scenarios. Working Paper, 22 August. Available at SSRN 3021829.

Moshirian, F., Nguyen, H.G.L., Pham, P.K., 2012. Overnight public information, order placement, and price discovery during the pre-opening period. Journal of Banking & Finance 36, 2837–2851.

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

Pagano, M.S., Peng, L., Schwartz, R.A., 2013. A call auction's impact on price formation and order routing: Evidence from the NASDAQ stock market. Journal of Financial Markets 16, 331–361.

Pagano, M.S., Schwartz, R.A., 2003. A closing call's impact on market quality at Euronext Paris. Journal of Financial Economics 68, 439–484.

Rindi, B., 2008. Informed traders as liquidity providers: Anonymity, liquidity and price formation. Review of Finance 12, 497–532.

Serbera, J.P., Paumard, P., 2016. The fall of high-frequency trading: A survey of competition and profits. Research in International Business and Finance 36, 271–287.

Stoll, H.R., Whaley, R.E., 1990. Stock market structure and volatility. The Review of Financial Studies 3, 37–71.

Vives, X., 1995. The speed of information revelation in a financial market mechanism. Journal of Economic Theory 67, 178–204.

Tables

Table 1: Preopening order flow for the CAC 40 sample.

Time	07:15–07:30	07:30–07:45	07:45–08:00	08:00–08:15	08:15–08:30	08:30–08:45	08:45–Open
HFT							
Submit total	385.0	149.0	14,937.0	3,620.0	2,265.0	2,018,873.0	1,029,859.0
Submit average	0.0	0.0	1.6	0.4	0.2	220.8	112.6
Average size	1,807.0	813.3	2,173.6	673.9	823.3	192.1	667.9
Modify total	0.0	0.0	0.0	0.0	5.0	570.0	334,694.0
Modify average	0.0	0.0	0.0	0.0	0.0	0.1	36.6
Cancel total	6.0	10.0	6.0	145.0	1,166.0	3,620.0	279,430.0
Cancel average	0.0	0.0	0.0	0.0	0.1	0.4	30.6
MIXED							
Submit total	80.0	336.0	3,052.0	245,510.0	29,949.0	102,097.0	1,188,136.0
Submit average	0.0	0.0	0.3	26.8	3.3	11.2	129.9
Average size	2,776.0	3,729.1	1,486.7	2,095.0	2,432.8	1,824.3	5,938.6
Modify total	4.0	87.0	1,877.0	2,178.0	1,759.0	6,013.0	538,900.0
Modify average	0.0	0.0	0.2	0.2	0.2	0.7	58.9
Cancel total	4.0	9.0	812.0	24,847.0	14,088.0	9,346.0	304,282.0
Cancel average	0.0	0.0	0.1	2.7	1.5	1.0	33.3
NON HFT							
Submit total	546,131.0	475,968.0	198,644.0	147,416.0	77,681.0	958,569.0	295,765.0
Submit average	59.7	52.1	21.7	16.1	8.5	104.8	32.3
Average size	216.8	187.0	338.9	288.8	653.2	779.8	1,373.8
Modify total	587.0	599.0	280.0	374.0	278.0	2,299.0	160,165.0
Modify average	0.1	0.1	0.0	0.0	0.0	0.3	17.5
Cancel total	54,282.0	83,630.0	83,286.0	20,355.0	18,608.0	18,584.0	54,548.0
Cancel average	5.9	9.1	9.1	2.2	2.0	2.0	6.0

This Table reports the average and total (sum), across days and stocks, number of new submissions, revisions and cancellations during the preopening and in 15 minute intervals, for each type of trader in the data set; HFT, MIXED and NON HFT. Zero averages (0.0) represent very small values (< 0.1). For submissions, the, across days and stocks, average order size is also reported.

Table 2: Opening statistics for the 36 stock sample from the CAC 40 Index.

Security	BUY HFT (%)	BUY MIXED (%)	BUY NON HFT (%)	SELL HFT (%)	SELL MIXED (%)	SELL NON HFT (%)	Trades (%)	Trades average	Volume average	Value average
Accor	4.3	60.8	35.0	4.4	48.2	47.4	1.2	46.9	8,953.6	262,413.0
Air Liquide	5.9	39.1	55.1	3.1	34.7	62.2	1.9	124.1	13,806.2	1,343,993.7
Axa	4.4	45.3	50.3	4.0	36.0	60.0	1.5	145.5	96,130.2	1,533 514.0
BNP Paribas	4.0	42.4	53.6	2.9	34.8	62.3	1.1	179.3	57,957.6	2,728,559.1
Cap Gemini	6.9	46.0	47.1	6.2	40.9	52.9	1.2	58.5	8,940.1	356,822.5
Carrefour	4.6	49.8	45.6	5.2	45.8	49.0	1.3	89.4	30,825.1	709,855.6
Credit Agricole	6.8	20.3	72.9	5.5	20.7	73.8	2.1	190.0	126,248.8	959,861.9
Danone	3.0	66.9	30.0	3.0	50.7	46.4	1.3	92.5	27,963.2	1,529,767.4
Lvmh	4.5	49.2	46.3	4.4	48.4	47.2	1.6	106.8	13,617.0	1,838,125.0
Michelin Nom.	6.8	49.7	43.6	3.1	31.8	65.1	1.6	103.9	10,707.6	775,199.2
L'Oreal	5.4	54.4	40.2	4.5	45.8	49.8	1.5	80.6	9,845.7	1,209,364.8
Pernod Ricard	7.2	56.7	36.1	4.0	52.6	43.4	1.2	61.9	7,246.9	654,519.6
Kering	4.9	54.1	41.0	4.9	45.9	49.2	1.4	47.8	2,992.8	494,872.3
Saint Gobain	6.2	54.7	39.1	4.5	43.4	52.2	1.1	79.3	22,456.8	748,897.3
Schneider Electric	5.0	57.0	38.0	5.1	53.3	41.7	0.9	78.9	19,869.7	1,176,372.9
Total	3.2	44.3	52.5	3.7	36.0	60.3	1.4	189.6	82,860.6	3,339,932.5
Veolia Environn.	6.1	37.8	56.1	6.3	37.0	56.7	1.4	79.4	34,383.7	374,904.4
Vinci	3.6	45.5	50.9	4.0	34.3	61.8	1.6	115.9	25,525.7	1,016,483.0
Vivendi	6.4	45.6	47.9	5.1	41.2	53.7	1.8	120.8	73,438.7	1,211,928.8
Essilor Intl	5.1	49.0	45.9	3.8	52.7	43.5	1.4	75.7	8,571.6	689,196.2
SAFRAN	3.2	40.5	56.3	2.6	41.2	56.2	1.8	78.8	11,744.4	485,001.1
Bouygues	6.7	44.7	48.6	6.1	43.6	50.3	1.2	54.8	11,908.5	278,165.5
Publicis Groupe	5.6	58.7	35.7	4.9	53.8	41.3	0.9	44.0	6,749.9	389,673.0
Societe Generale	3.8	35.6	60.5	3.9	30.7	65.4	1.0	170.5	63,474.1	2,063,893.2
Technip	7.6	40.7	51.7	4.9	48.7	46.4	1.5	82.4	8,392.0	664,720.3
Renault Group	4.6	51.8	43.7	5.8	42.0	52.2	1.0	77.9	15,921.1	888,674.5
Airbus Group	2.6	41.0	56.4	1.9	24.1	74.0	1.8	194.4	40,984.7	1,772,246.1
Lafarge	6.2	38.9	54.8	6.0	33.3	60.7	1.5	84.8	8,687.1	433,775.4
Stmicroelectronics	9.5	41.7	48.8	8.5	46.9	44.7	1.0	41.9	32,015.2	203,453.4
EDF	3.0	40.2	56.8	3.4	21.8	74.8	2.4	117.4	24,131.0	465,484.1
Alstom	5.6	31.8	62.6	6.4	37.2	56.4	1.7	103.0	21,355.4	599,144.3
Vallourec	7.2	38.6	54.2	6.3	38.7	55.0	1.4	67.9	8,679.3	365,233.8
Engie	3.9	46.9	49.3	2.9	30.2	66.9	2.4	169.8	71,119.6	1,164,081.3
Legrand	3.6	61.7	34.7	2.7	53.1	44.1	1.3	41.3	7,831.7	290,675.1
Orange	4.3	35.6	60.1	3.4	32.5	64.1	1.6	168.0	127,862.4	1,069,817.0
Sanofi	4.4	50.5	45.1	3.5	48.0	48.5	1.1	147.2	52,700.8	4,042,025.7
Average across stocks	5.2	46.3	48.5	4.5	40.5	55.0	1.4	103.1	33,219.4	1,059,184.6

Out of sample	Platform negotiated
Solvay	Belgium
Gemalto	the Netherlands
Unibail-Rodamco	the Netherlands
ArcelorMittal Reg	Luxembourg

Clearing at the open: HFT, MIXED and NON HFT percentage on the buy and sell side (the percentage for each of the three categories and for each market side is calculated as the total number of flagged trade messages relative to the total number of opening trade messages), percentage of opening trades relative to total daily trades, average number of trades, average clearing volume (shares) and average trading value (Euros). Four stocks from the CAC 40 Index are excluded from our analysis as they are not negotiated on the Paris platform.

Figures

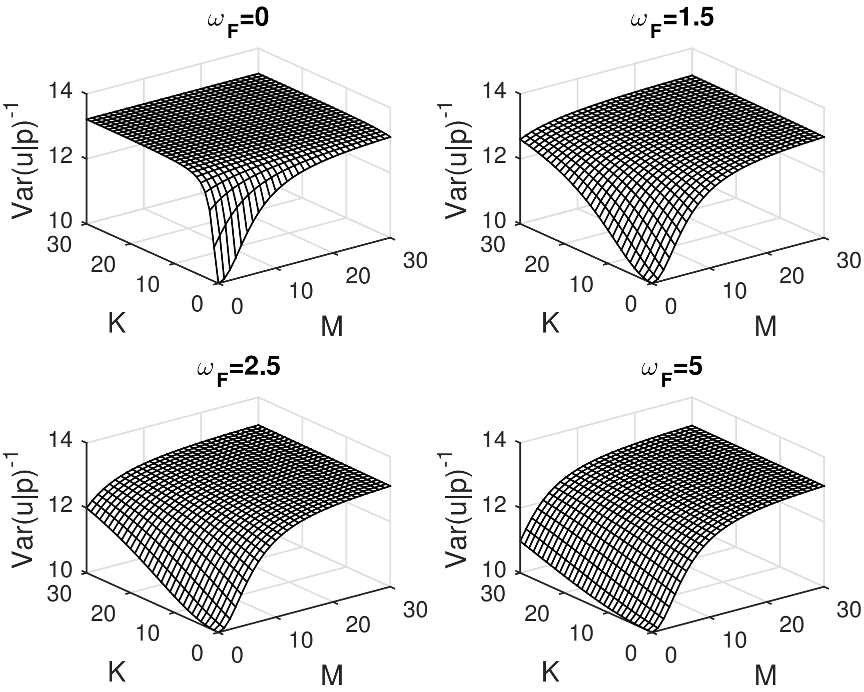

Figure 1: Price efficiency with respect to ω_F. The following parameters are fixed: $A = 3$, $Z = 5$, $\sigma_u^2 = 0.1$, $\sigma_S^2 = 0.5$, $\sigma_F^2 = 0.12$, $\sigma_Q^2 = 0.2$, $\omega_S = 0$, $\lambda = 0.5$.

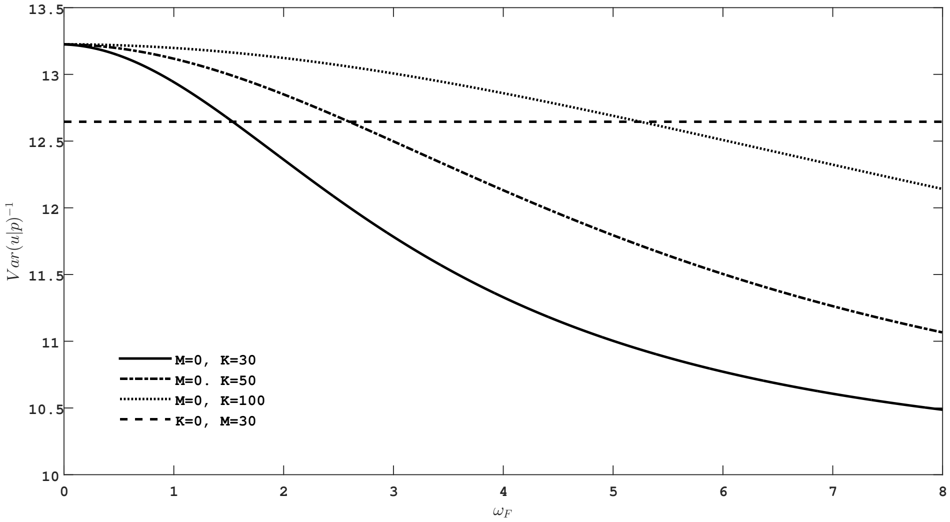

Figure 2: Price efficiency with respect to ω_F. The following parameters are fixed: $A = 1$, $Z = 5$, $\sigma_u^2 = 0.1$, $\sigma_S^2 = 0.5$, $\sigma_F^2 = 0.12$, $\sigma_Q^2 = 0.2$, $\omega_S = 0.5$, $\lambda = 0.5$.

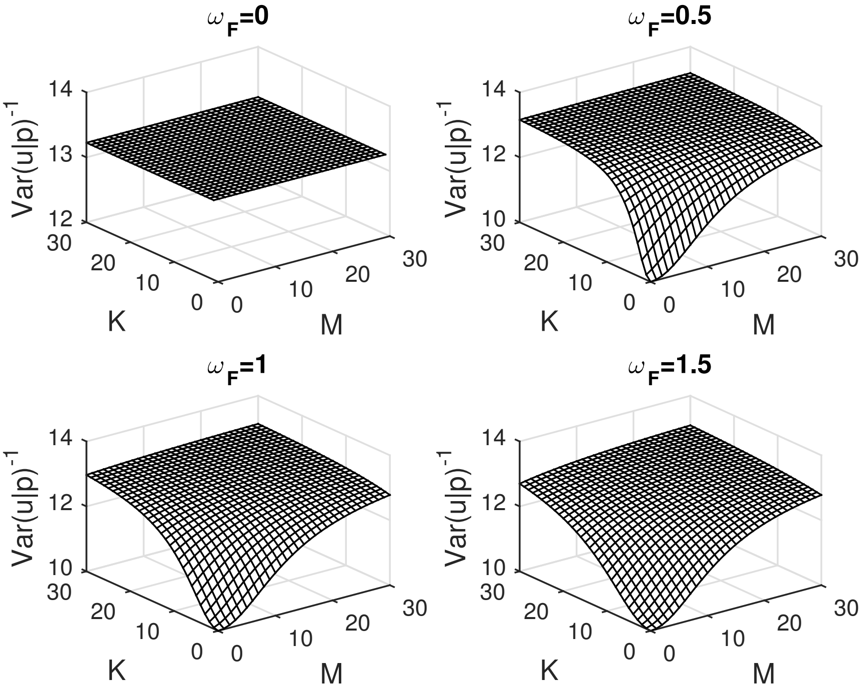

Figure 3: Price efficiency with respect to ω_F. The following parameters are fixed: $A = 0$, $Z = 5$, $\sigma_u^2 = 0.1$, $\sigma_S^2 = 0.5$, $\sigma_F^2 = 0.12$, $\sigma_Q^2 = 0.2$, $\omega_S = 0.5$, $\lambda = 0.5$.

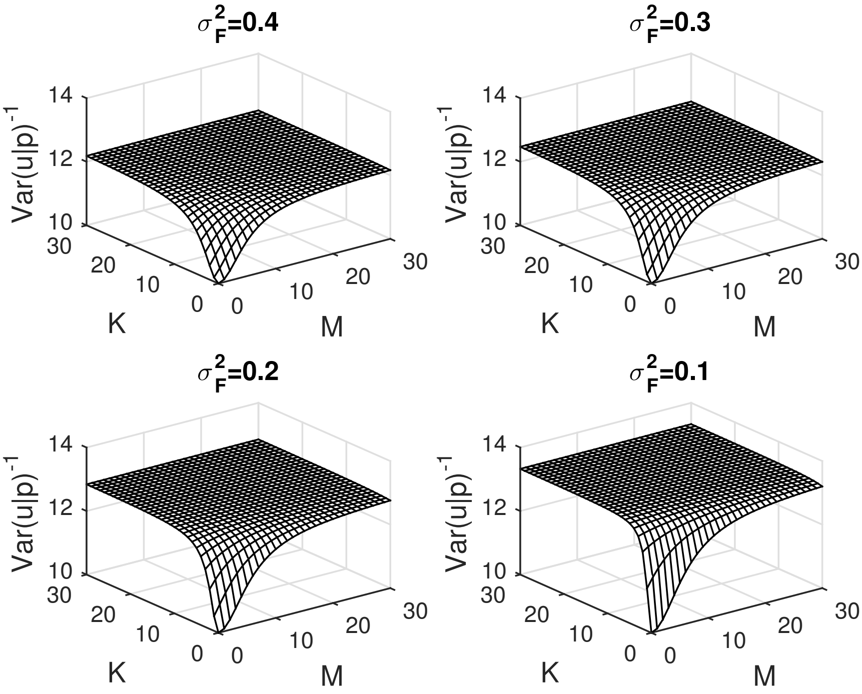

Figure 4: Price efficiency with respect to σ_F^2. The following parameters are fixed: $A = 3$, $Z = 5$, $\sigma_u^2 = 0.1$, $\sigma_S^2 = 0.5$, $\sigma_Q^2 = 0.2$, $\omega_F = 0$, $\omega_S = 0$, $\lambda = 0.5$.

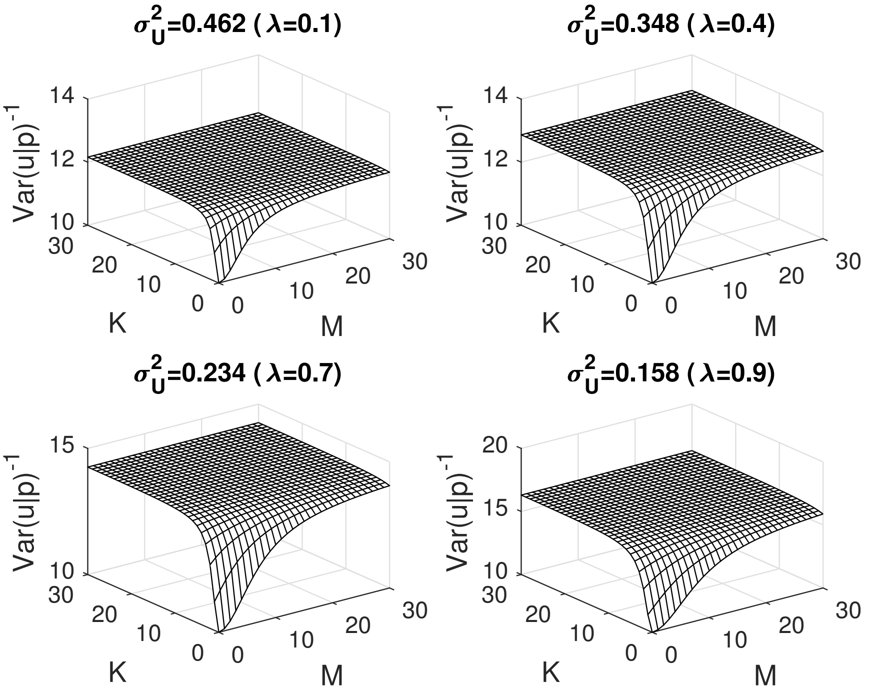

Figure 5: Price efficiency with respect to λ. The following parameters are fixed: $A = 3$, $Z = 5$, $\sigma_u^2 = 0.1$, $\sigma_S^2 = 0.5$, $\sigma_F^2 = 0.12$, $\sigma_Q^2 = 0.2$, $\omega_F = 0$, $\omega_S = 0$.

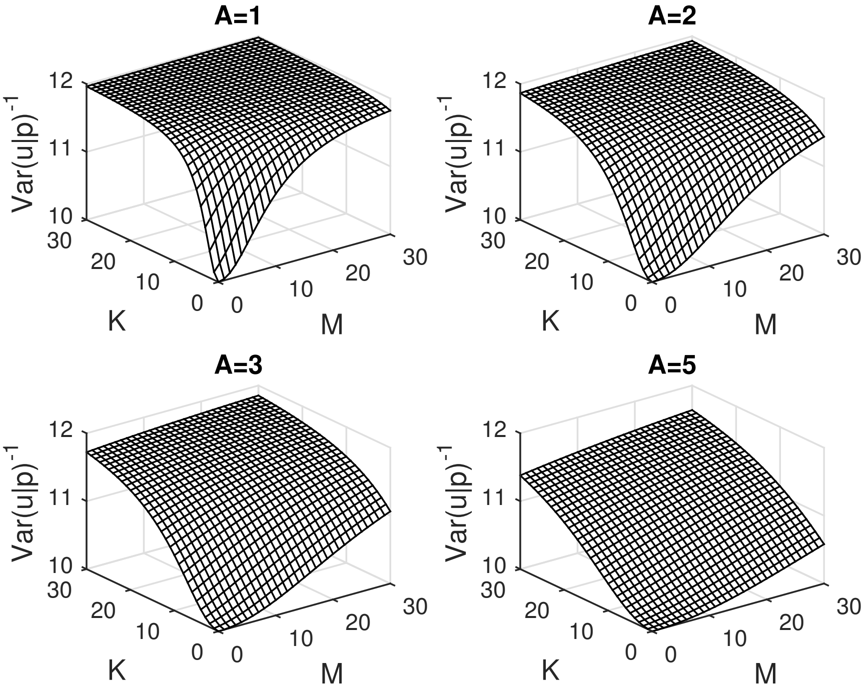

Figure 6: Price efficiency with respect to A. The following parameters are fixed: $Z = 15$, $\sigma_u^2 = 0.1$, $\sigma_S^2 = 0.7$, $\sigma_F^2 = 0.3$, $\sigma_Q^2 = 0.4$, $\omega_F = 0$, $\omega_S = 0$, $\lambda = 0.5$.

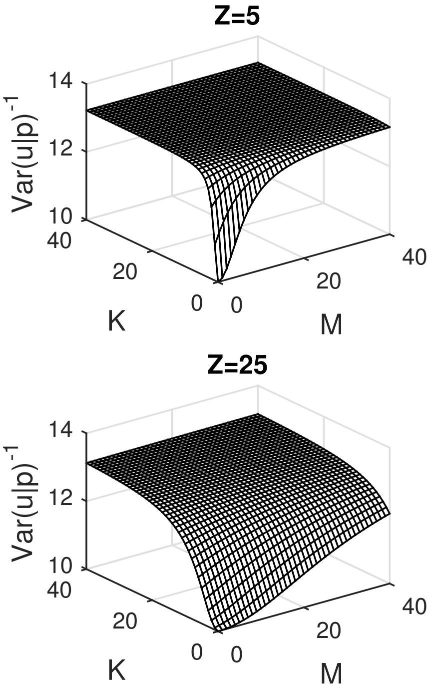

Figure 7: Price efficiency with respect to Z. The following parameters are fixed: $A = 3$, $\sigma_u^2 = 0.1$, $\sigma_S^2 = 0.5$, $\sigma_F^2 = 0.12$, $\sigma_Q^2 = 0.2$, $\omega_F = 0$, $\omega_S = 0$, $\lambda = 0.5$.

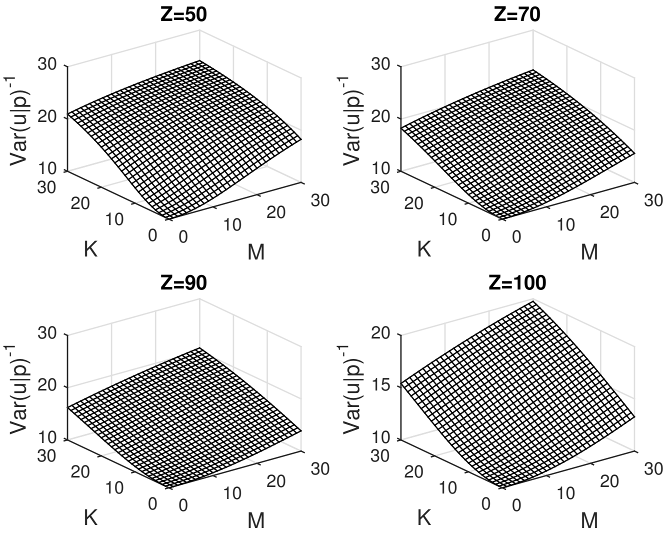

Figure 8: Price efficiency with respect to Z. The following parameters are fixed: $A = 3$, $\sigma_u^2 = 0.1$, $\sigma_S^2 = 0.07$, $\sigma_F^2 = 0.05$, $\sigma_Q^2 = 0.5$, $\omega_F = 0$, $\omega_S = 0$, $\lambda = 0.5$.

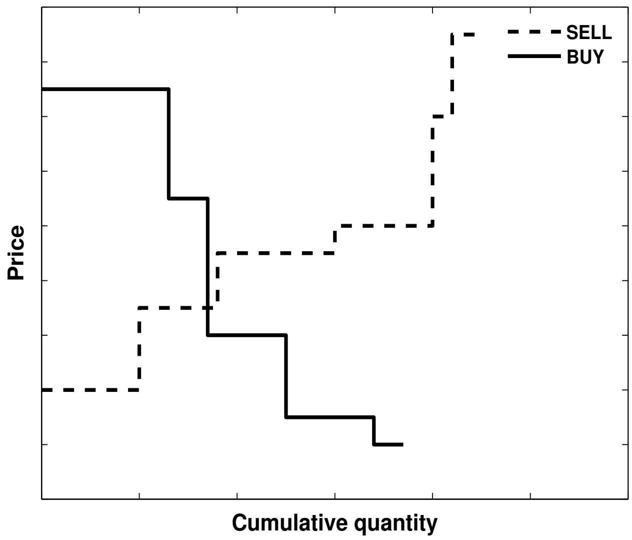

Figure 9: A hypothetical supply–demand schedule at the opening.

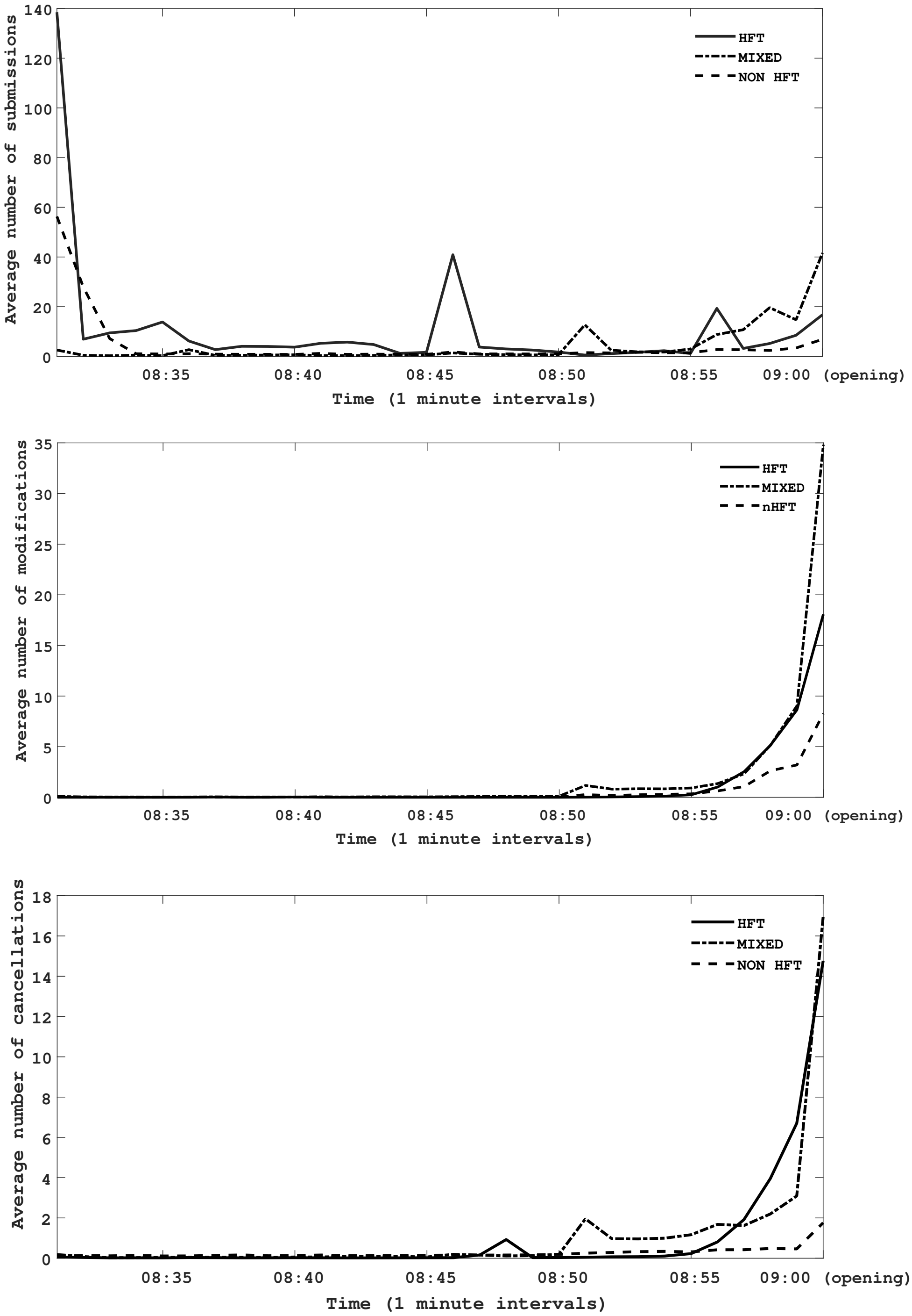

Figure 10: The across stocks and days average number of submissions, modifications and cancellations during the last 30 minutes of the preopening, in 1 minute intervals.

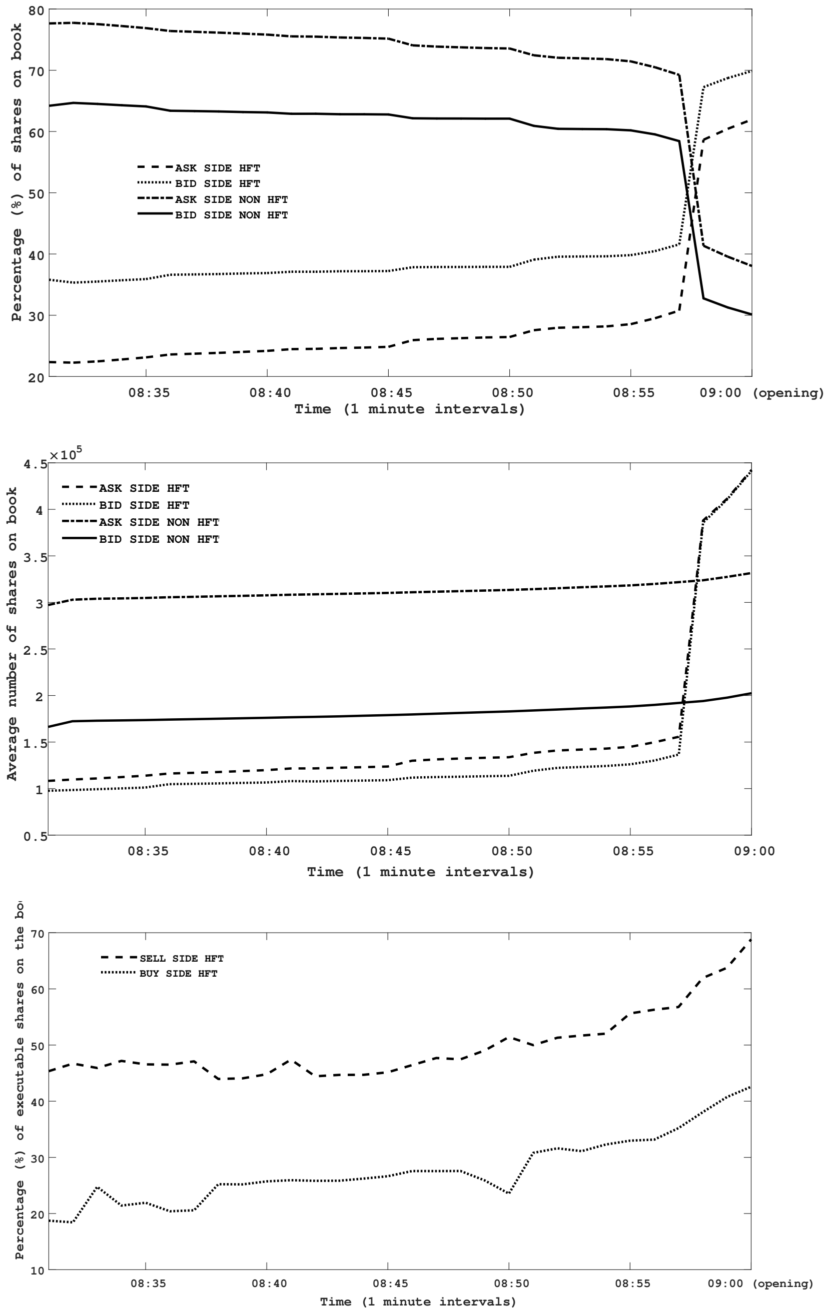

Figure 11: Top graph: The across stocks and days average percentage of total available shares on the order book due to HFT (HFT and MIXED in the data set) and NON HFT quotes. Middle graph: The across stocks and days average number of shares on the order book due to HFT and NON HFT quotes. Bottom graph: The across stocks and days average percentage of aggressive (i.e., executable) shares on the order book due to HFT quotes.

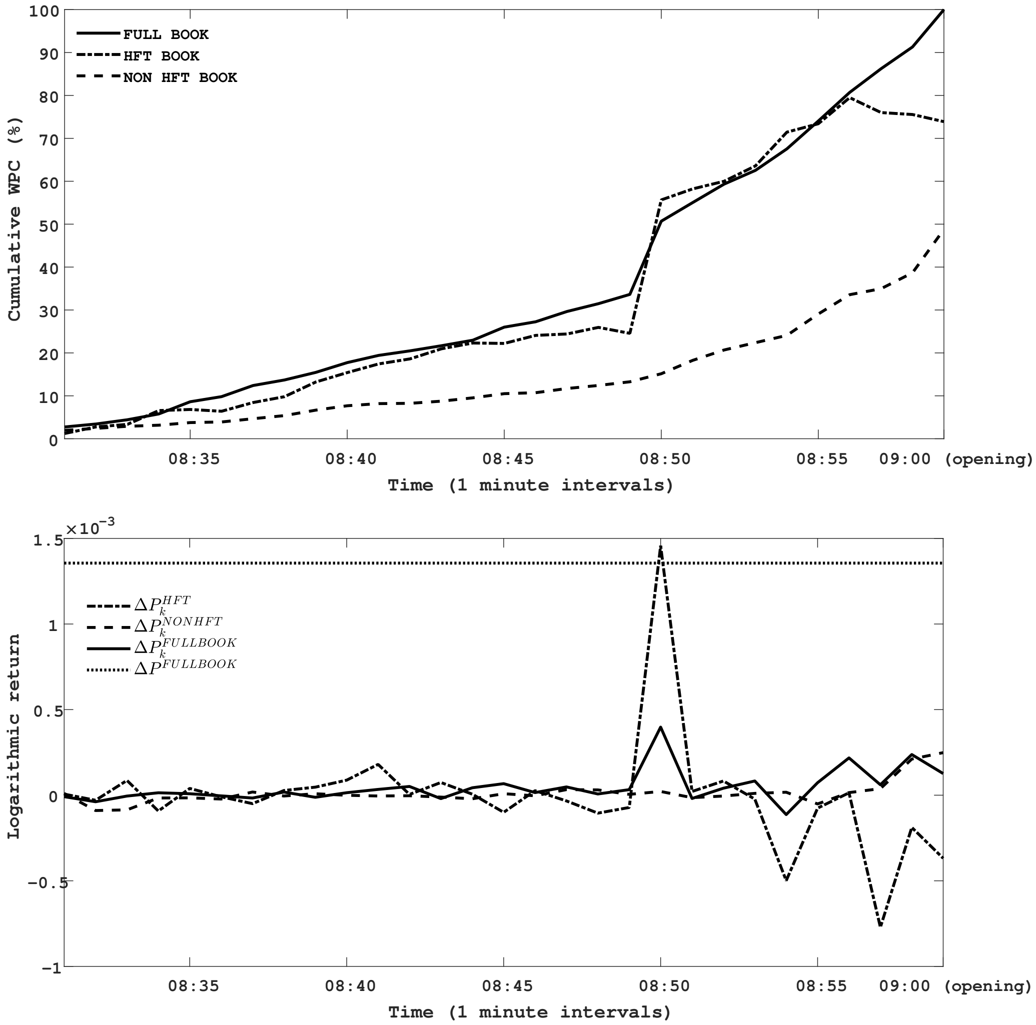

Figure 12: Top graph: Cumulative Weighted Price Contribution statistic during the last 30 minutes of the preopening for the three cases: HFT, NON HFT and FULL BOOK. WPC is calculated for each sample day as follows: $WPC_k = \sum_{i=1}^{N_{stocks}} \left(\frac{|\Delta P_i|}{\sum_{i=1}^{N_{stocks}} |\Delta P_i|} \right) \times \left(\frac{\Delta P_{k,i}}{\Delta P_i} \right)$, where ΔP_i is the total logarithmic price change from 08:30 to 09:00 (opening) for stock i ($i = 1, ..., N_{stocks}$) calculated on the basis of the full set of orders and $\Delta P_{k,i}$ is the logarithmic price change for interval k, $k = 1, ..., 29$, calculated for the three cases (HFT, NON HFT and FULL BOOK). Then, for each k interval WPC is averaged across days. Bottom graph: The across days and stocks average 1 min return, ΔP_k, $k = 1, ..., 29$, during the last 30 minutes of the preopening for the three cases: HFT, NON HFT and FULL BOOK. Together is plotted the across days and stocks total price change from 08:30 to 09:00, ΔP, calculated using the full book (horizontal line).

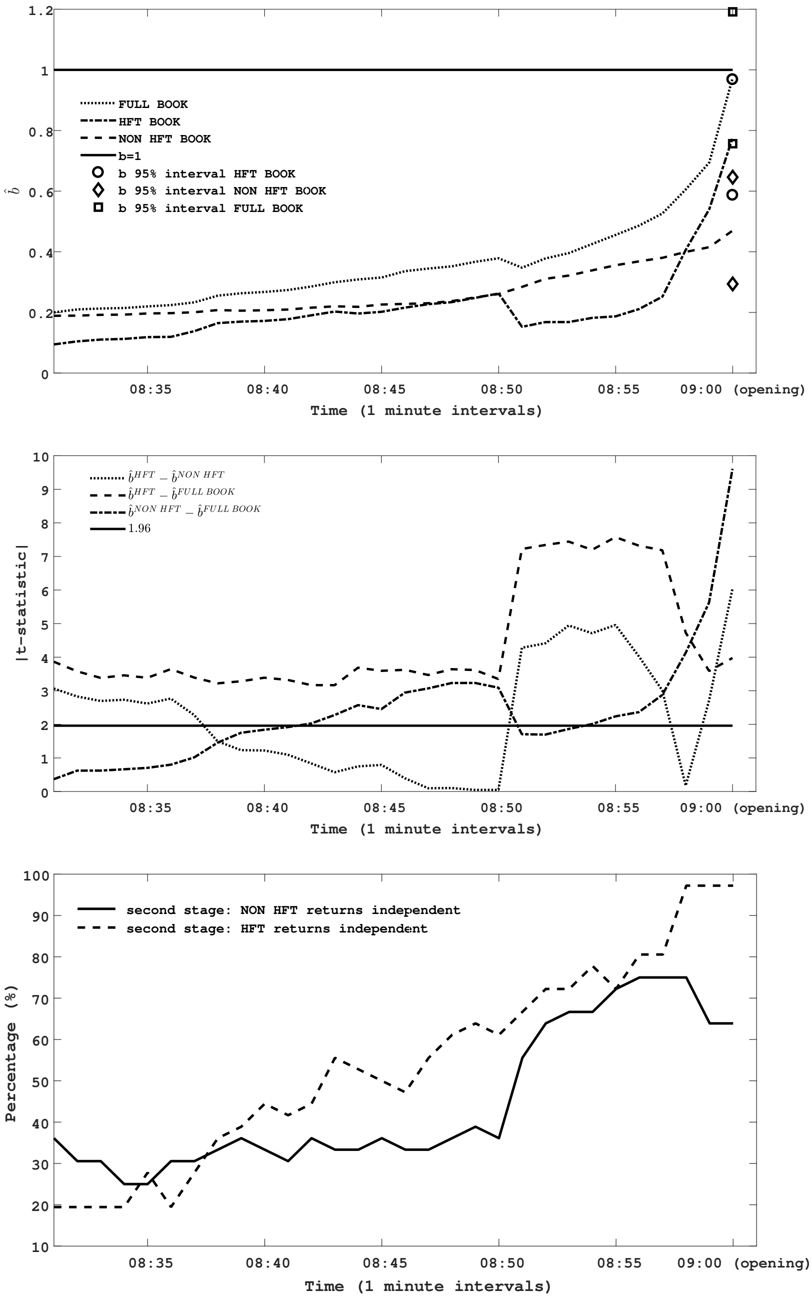

Figure 13: Upper graph: estimated unbiasedness regression b coefficients during the preopening along with the corresponding 95% confidence intervals at the opening, obtained by using i) the FULL BOOK, ii) the HFT BOOK and iii) the NON HFT BOOK indicative prices in equation (49). Middle graph: the absolute value of the t-statistic for the null hypothesis of across–stocks equal mean b for the following cases: i) $mean(\hat{b}^{HFT})$=$mean(\hat{b}^{NONHFT})$, ii) $mean(\hat{b}^{HFT})$=$mean(\hat{b}^{FULLBOOK})$ and iii) $mean(\hat{b}^{NONHFT})$=$mean(\hat{b}^{FULLBOOK})$. Bottom graph: The percentage of statistically significant d coefficient estimates from equations (50) and (51), for the following cases: i) NON HFT indicative returns (R_{ct}^{NONHFT}) in first stage as independent variable and HFT indicative returns (R_{ct}^{HFT}) as independent variable in second stage, ii) HFT indicative returns (R_{ct}^{HFT}) in first stage as independent variable and NON HFT indicative returns (R_{ct}^{NONHFT}) as independent variable in second stage.

Appendix A

This Appendix provides a list of the model parameters presented in Section 3:

u: ex–post liquidation value of the asset

σ_u^2: variance of u

K: number of Fast informed traders

M: number of Slow informed traders

L: number of uninformed traders

Z: number of noise traders

W: terminal wealth

A: risk aversion coefficient

Q_F: demand function for Fast informed traders

Q_S: demand function for Slow informed traders

Q_U: demand function for uninformed traders

Q: random demand for noise traders

U: utility function

ω_F: price elasticity factor for Fast informed traders

ω_S: price elasticity factor for Slow informed traders

ω_U: price elasticity factor for uninformed traders

s: the private signal

e: the error term of the private signal

σ_e^2: true variance of the error term of the private signal

σ_F^2: variance of the error term of the private signal for Fast informed traders

σ_S^2: variance of the error term of the private signal for Slow informed traders

σ_U^2: variance of the error term of the private signal conjectured by uninformed traders

λ: the weight uninformed traders place on σ_F^2 and σ_S^2 when conjecturing the linear price rule

p: the market price

α: coefficient of s in the linear pricing rule conjectured by uninformed traders

β: coefficient of Q in the linear pricing rule conjectured by uninformed traders

Internet Appendix

Price efficiency and High Frequency Trading in call auctions

(Supplementary material)

P. Anagnostidis, P. Fontaine, C. Varsakelis

This Internet Appendix provides supplementary material for the theoretical model developed in Section 3. In particular, Section I provides analytical formulas for the partial derivatives of price efficiency, as defined in equation (33), with respect to the model parameters: K (number of HFTs), M (number of slow informed traders), σ_F^2 (variance of HFTs' signal), σ_S^2 (variance of slow informed traders' signal), ω_F (price elasticity factor for HFTs), ω_S (price elasticity factor for slow informed traders), σ_U^2 (variance of informed traders' signal that uninformed traders conjecture), A (risk aversion coefficient) and Z (number of noise traders).

I Partial derivatives for price efficiency

The following results are obtained from equation (33) by simple differentiation:

$$\frac{\partial Var(u|p)^{-1}}{\partial K} = \frac{2\,Z^2\,\sigma_Q^2}{\left(\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}}+\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}}\right)^3\left(A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}\right)\left(\sigma_U^2+\frac{Z^2\,\sigma_Q^2}{\left(\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}}+\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma^2{}_u}}\right)^2}\right)^2} > 0 \tag{I}$$

$$\frac{\partial Var(u|p)^{-1}}{\partial M} = \frac{2\,Z^2\,\sigma_Q^2}{\left(\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+{\sigma_u}^2\right)}{\sigma_u^2}}+\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}}\right)^3\left(A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}\right)\left(\sigma_U^2+\frac{Z^2\,\sigma_Q^2}{\left(\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}}+\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}}\right)^2}\right)^2} > 0 \tag{II}$$

$$\frac{\partial Var(u|p)^{-1}}{\partial \sigma_F^2} = -\frac{\lambda+\frac{2\,K\,Z^2\,\sigma_Q^2\,\left(A+\frac{\omega_F}{\sigma_u^2}\right)}{\left(\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}}+\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_u^2+\sigma_F^2\right)}{\sigma_u^2}}\right)^3\left(A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_u^2+\sigma_F^2\right)}{\sigma_u^2}\right)^2}}{\left(\lambda\,\sigma_F^2-\sigma_S^2\,(\lambda-1)+\frac{Z^2\,\sigma_Q^2}{\left(\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}}+\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_u^2+\sigma_F^2\right)}{\sigma_u^2}}\right)^2}\right)^2} < 0 \tag{III}$$

Symmetrically, we obtain that $\frac{\partial Var(u|p)^{-1}}{\partial \sigma_S^2} < 0$.

$$\frac{\partial Var(u|p)^{-1}}{\partial \omega_F} =$$

$$-\frac{2\,K\,Z^2\,\sigma_Q^2\,\left(\sigma_F^2+{\sigma_u^2}^2\right)}{\sigma_u^2\left(\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}}+\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}}\right)^3\left(A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}\right)^2\left(\sigma_U^2+\frac{Z^2\,\sigma_Q^2}{\left(\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}}+\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}}\right)^2}\right)^2} < 0 \tag{IV}$$

Symmetrically, we obtain that $\frac{\partial Var(u|p)^{-1}}{\partial \omega_S} < 0$.

$$\frac{\partial Var(u|p)^{-1}}{\partial \sigma_U^2} = -\frac{1}{\left(\sigma_U^2+\frac{Z^2\,\sigma_Q^2}{\left(\frac{K}{A\,\sigma_F^2+\frac{\omega_F\,\left(\sigma_F^2+\sigma_u^2\right)}{\sigma_u^2}}+\frac{M}{A\,\sigma_S^2+\frac{\omega_S\,\left(\sigma_S^2+\sigma_u^2\right)}{\sigma_u^2}}\right)^2}\right)^2} < 0 \tag{V}$$

$$\frac{\partial Var(u|p)^{-1}}{\partial A} =$$

$$-\frac{2\,Z^2\,\sigma_{\mathrm{Q}}^2\left(\frac{K\,\sigma_{\mathrm{F}}^2}{\left(A\,\sigma_{\mathrm{F}}^2+\frac{\omega_{\mathrm{F}}\,\left(\sigma_{\mathrm{F}}^2+\sigma_{\mathrm{u}}^2\right)}{\sigma_{\mathrm{u}}^2}\right)^2}+\frac{M\,\sigma_{\mathrm{S}}^2}{\left(A\,\sigma_{\mathrm{S}}^2+\frac{\omega_{\mathrm{S}}\,\left(\sigma_{\mathrm{S}}^2+\sigma_{\mathrm{u}}^2\right)}{\sigma_{\mathrm{u}}^2}\right)^2}\right)}{\left(\frac{K}{A\,\sigma_{\mathrm{F}}^2+\frac{\omega_{\mathrm{F}}\,\left(\sigma_{\mathrm{F}}^2+\sigma_{\mathrm{u}}^2\right)}{\sigma_{\mathrm{u}}^2}}+\frac{M}{A\,\sigma_{\mathrm{S}}^2+\frac{\omega_{\mathrm{S}}\,\left(\sigma_{\mathrm{S}}^2+\sigma_{\mathrm{u}}^2\right)}{\sigma_{\mathrm{u}}^2}}\right)^3\left(\sigma_{\mathrm{U}}^2+\frac{Z^2\,\sigma_{\mathrm{Q}}^2}{\left(\frac{K}{A\,\sigma_{\mathrm{F}}^2+\frac{\omega_{\mathrm{F}}\,\left(\sigma_{\mathrm{F}}^2+\sigma_{\mathrm{u}}^2\right)}{\sigma_{\mathrm{u}}^2}}+\frac{M}{A\,\sigma_{\mathrm{S}}^2+\frac{\omega_{\mathrm{S}}\,\left(\sigma_{\mathrm{S}}^2+\sigma_{\mathrm{u}}^2\right)}{\sigma_{\mathrm{u}}^2}}\right)^2}\right)^2} < 0 \quad \text{(VI)}$$

$$\frac{\partial Var(u|p)^{-1}}{\partial Z} =$$

$$-\frac{2\,Z\,\sigma_{\mathrm{Q}}^2}{\left(\frac{K}{A\,\sigma_{\mathrm{F}}^2+\frac{\omega_{\mathrm{F}}\,\left(\sigma_{\mathrm{F}}^2+{\sigma_{\mathrm{u}}^2}^2\right)}{\sigma_{\mathrm{u}}^2}}+\frac{M}{A\,\sigma_{\mathrm{S}}^2+\frac{\omega_{\mathrm{S}}\,\left(\sigma_{\mathrm{S}}^2+\sigma_{\mathrm{u}}^2\right)}{\sigma_{\mathrm{u}}^2}}\right)^2\left(\sigma_{\mathrm{U}}^2+\frac{Z^2\,\sigma_{\mathrm{Q}}^2}{\left(\frac{K}{A\,\sigma_{\mathrm{F}}^2+\frac{\omega_{\mathrm{F}}\,\left(\sigma_{\mathrm{F}}^2+\sigma_{\mathrm{u}}^2\right)}{\sigma_{\mathrm{u}}^2}}+\frac{M}{A\,\sigma_{\mathrm{S}}^2+\frac{\omega_{\mathrm{S}}\,\left(\sigma_{\mathrm{S}}^2+\sigma_{\mathrm{u}}^2\right)}{{\sigma_{\mathrm{u}}^2}^2}}\right)^2}\right)^2} < 0 \quad \text{(VII)}$$