News Trading and Speed*

Thierry Foucault† Johan Hombert‡ Ioanid Roşu§

September 25, 2012

Abstract

Informed trading can take two forms: (i) trading on more accurate information or (ii) trading on public information faster than other investors. The latter is increasingly important due to technological advances. To disentangle the effects of accuracy and speed, we derive the optimal dynamic trading strategy of an informed investor when he reacts to news (i) at the same speed or (ii) faster than other market participants, holding information precision constant. With a speed advantage, the informed investor's order flow is much more volatile, accounts for a much bigger fraction of trading volume, and forecasts very short run price changes. We use the model to analyze the effects of high frequency traders on news (HFTNs) on liquidity, volatility, price discovery and provide empirical predictions about the determinants of their activity.

KEYWORDS: Informed trading, news, high frequency trading, liquidity, volatility, price discovery.

*We thank Terry Hendershott, Pete Kyle, Stefano Lovo, Victor Martinez, Dimitri Vayanos for their sugegstions. We are also grateful to finance seminar participants at Paris Dauphine, Copenhagen Business School, Univ. Carlos III in Madrid, ESSEC and conference participants at the NYU Stern Microstructure Meeting for valuable comments.

†HEC Paris, foucault@hec.fr.

‡HEC Paris, hombert@hec.fr.

§HEC Paris, rosu@hec.fr.

1 Introduction

The effect of news arrival on trades and prices in securities markets is of central interest. For instance, informational efficiency is often measured by the speed at which prices incorporate public information and many researchers have studied trading volume and prices around news (e.g., Patell and Wolfson (1984), Kim and Verrechia (1991, 1994), Busse and Green (2001), Vega (2006), or Tetlock (2010)). A new breed of market participants, “high frequency news traders” (HFTNs), now use the power of computers to collect, process and exploit news faster than other market participants (see “Computers that trade on the news”, the New-York Times, May 2012).[1] Hence, the impact of news in today’s securities markets depends on the behavior of these traders. Can we rely on traditional models of informed trading to understand this behavior and its effects? Is trading faster on public information the same thing as trading on more accurate private information?

To address this question, we consider a model in which an informed investor continuously receives news about the payoff of a risky security. He has both a greater information processing capacity *and* a higher speed of reaction to news than market-makers. The information processing advantage enables the informed investor to form a more precise forecast of the fundamental value of the asset while the speed advantage enables him to forecast quote updates due to news arrival. Models of informed trading focus on the former type of advantage (accuracy) but not on the latter (speed).[2]

Our central finding is that the optimal trading strategy of the informed investor is very different when he has a speed advantage and when he has not, holding the precision of his private information constant. In particular, a small speed advantage

[1] News exploited by these traders are very diverse and include market events (quote updates, trades, orders), blog posts, news headlines, discussions in social forums etc. For instance, Brogaard, Hendershott, and Riordan (2012) show that high frequency traders in their data react to information contained in macro-economic announcements, limit order book updates, and market-wide returns. Data vendors such as Bloomberg, Dow-Jones or Thomson Reuters have started providing pre-processed real-time news feed to high frequency traders. For instance, in their on-line advertisement for real-time data processing tools, Dow Jones states: “Timing is everything and to make lucrative, well-timed trades, institutional and electronic traders need accurate real-time news available, including company financials, earnings, economic indicators, taxation and regulation shifts. Dow Jones is the leader in providing high-frequency trading professionals with elementized news and ultra low-latency news feeds for algorithmic trading.” See http://www.dowjones.com/info/HighFrequencyTrading.asp.

[2] This is also the case for models that specifically analyze informed trading around news releases. For instance, Kim and Verrechia (1994) assume that when news are released about the payoff of an asset, some traders (”information processors”) are better able to interpret their informational content than market-makers. As a result these traders have more accurate forecasts than market-makers but receive news at the same time as other traders.

for the informed investor makes his optimal portfolio much more volatile, that is, the informed investor trades much more when he can react to news faster than dealers.

In our set-up, the informed investor has two motivations for trading. First, his forecast of the asset liquidation value is more precise than that of dealers. Second, by receiving news a split second before dealers, the informed investor can forecast dealers' quote updates due to public information arrival, that is, price changes in the very short run. The investor's optimal position in the risky asset reflects these two motivations: (i) its drift is proportional to dealers' forecast error (the difference between the informed investor's and dealers' estimates of the asset payoff) while (ii) its instantaneous variance is proportional to news. The second component (henceforth the "news trading component") arises only if the informed investor has a speed advantage.[3] The investor's position is therefore much more volatile in this case. Figure 1 illustrates this claim for one particular realization of news in our model.

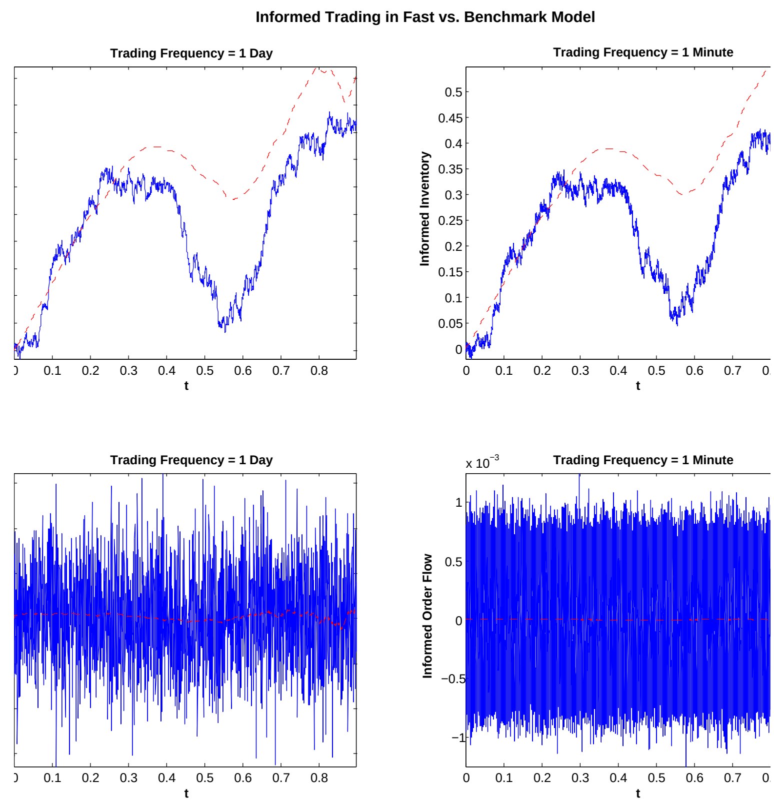
This finding has several important and new implications. For instance, the informed investor's share of trading volume is much higher when he has a speed advantage. Indeed, the volatility of his order flow is of the same order of magnitude as the volatility of noise traders' order flow. Moreover, with a speed advantage, the informed investor's order flow at the high frequency (over very short interval) has a very short positive correlation with subsequent returns because the informed investor's trade are mainly driven by news arrivals, at high frequency. These features fit well with some stylised facts about high frequency traders: (a) their trades account for a large fraction of the trading volume (see Hendershott, Jones, and Menkveld (2011), Brogaard (2011), Brogaard, Hendershott, and Riordan (2012) or Chaboud, Chiquoine, Hjalmarsson, and Vega (2009)) and (b) their aggressive orders (i.e., marketable orders) anticipate on very short run price changes (see Kirilenko, Kyle, Samadi, and Tuzun (2011) or Brogaard, Hendershott, and Riordan (2012)).[4] In contrast, we show that the model in which the informed investor has more accurate information only cannot explain these facts.

Moreover, the effect of the precision of public information (that is, the news received by dealers) differs from that obtained in other models of trading around news, such

[3] In contrast, the drift of the investor's position is proportional to dealers' forecast error even when the investor has no speed advantage, as in the continuous time version of Kyle (1985) or extensions of this model such as Back (1992), Back and Pedersen (1998), Back, Cao, and Willard (2000) or Chau and Vayanos (2007).

[4] For instance, Kirilenko, Kyle, Samadi, and Tuzun (2011) note (page 21) that "*possibly due to their speed advantage or superior ability to predict price changes, HFTs are able to buy right as the prices are about to increase.*"

Figure 1: Informed participation rate at various trading frequencies. The figure plots the evolution of the informed investor's position (upper panel) and the change in this position-the informed investor's trade- (lower panel) when the informed investor has a speed advantage (plain line) and when he has no speed advantage (dashed line) using the characterization of the optimal trading strategy for the investor in each case.



as Kim and Verrecchia (1994). Usually, more precise public information is associated with greater market liquidity (price impacts) but lower trading volume (see Kim and Verrechia (1994) for instance). In contrast, in our model, it is associated with both an increase in liquidity (as dealers are less exposed to adverse selection), more trading volume, and a greater participation rate of the informed investor. Indeed, an increase in the precision of public information enables the informed investor to better forecast short run quote updates by dealers, which induces him to trade more aggressively on news. As a result, the volatility of his position increases, which means that both the trading volume and the fraction of this trading volume due to the informed investor increases. These effects imply that dealers are more exposed to adverse selection due to news trading but this effect is second order relative to the fact that they can better forecast the final payoff of the asset, so that they are less at risk of accumulating a long position when the asset liquidation value is low or vice versa. As a result, liquidity improves when public news are more precise, even though informed trading is more intense.[5]

The informed investor's ability to forecast quote updates also implies that short run returns are positively related to his contemporaneous order flow and his lagged order flow. Indeed, quote updates reflect the arrival news on which the informed investor has already traded. Thus, short term returns and the informed investor's lagged order flow are positively correlated. An increase in the precision of the news received by dealers implies that short run returns are more correlated with contemporaneous the informed investor's order flow but less correlated with his lagged order flow. Again, these predictions arise only if the informed investor has a speed advantage.

Last, we use the model to analyze the effects of speed on liquidity, price discovery, and volatility. This is of interest since speed is often viewed as the distinctive advantage of high frequency traders and the debate on high frequency trading revolves around the question of what is the effect of speed on measures of market performance (see for instance SEC (2010) or Gai, Yao, and Ye (2012)). To speak to this debate, we compare standard measures of market performance when the informed investor has a speed advantage (the new environment with HFTNs) and when he has not (the old environment without HFTNs) in our model.

Illiquidity (price impact of trades) is higher when the informed investor has a speed advantage because the ability of the informed investor to react faster to news is an additional source of adverse selection for dealers.[6] Less obviously, this speed advantage also affects the nature of price discovery: price changes over short horizon are more corre-

[5] This finding suggests that controlling for the precision of public information is important in analyzing the impact of high frequency news trading activity on liquidity. Indeed, when public information is more precise, both the informed investor's share of trading volume and liquidity improves. Thus, variations in the precision of public information across stocks or over time should work to create a positive association between liquidity and measures of high frequency news traders activity. Yet, this association is spurious since as explained below granting a speed advantage to the informed investor always impairs liquidity in our model.

[6] In line with this prediction, Hendershott and Moulton (2011) find that a reduction in the speed of execution for market orders submitted to the NYSE in 2006 is associated with larger bid-ask spreads, due to an increase in adverse selection.

lated with innovations in the asset value (as found empirically in Brogaard et al.(2012) but less correlated with the long run estimate of this value by the informed investor. The first effect improves price discovery while the second impairs price discovery. In equilibrium, they exactly cancel out so that the average pricing error (the difference between the transaction price and the informed investor's estimate of the asset value) is the same whether the informed investor has a speed advantage or not.[7] Relatedly, high frequency news trading alters the relative influences of trades and news arrivals on short run volatility. Trades move dealers' price more when the investor has a speed advantage because they are more informative about imminent news. But precisely for this reason, dealers' quotes are less sensitive to news because news have been partly revealed through trading. Therefore, the magnitude of quote revisions after news is smaller when the informed investor has a speed advantage, which dampens volatility. These two effect exactly offset each other so that overall high frequency news trading has no effect on volatility.

High frequency traders' strategies are heterogeneous (see SEC (2010)). Accordingly, they do not necessarily have all the same effects on market quality. In particular, some HFTs implicitly act as market-makers (see Brogaard, Hendershott and Riordan (2012) or Menkveld (2012)). Market-makers may use speed to protect themselves against better informed traders (e.g., by cancelling their limit orders just before news arrival) and provide liquidity at lower cost (see Jovanovic and Menkveld (2011)). This type of strategy is not captured by our model, which restricts the informed investor to submit market orders, as in Kyle (1985). This assumption is reasonable since Brogaard, Hendershott and Riordan (2012) show empirically that only aggressive orders (i.e., market orders) submitted by high frequency traders are a source of adverse selection. However, it limits the scope of our implications. Accordingly, we do not claim that these implications are valid for all activities by high frequency traders.[8]

[7]Brogaard, Hendershott and Riordan (2012) find that aggressive trades by high frequency traders in their sample are negatively correlated with pricing errors. This is also the case in our model (the informed investor sells on average when the price is above his estimate of the fundamental value and buys otherwise). It does not follow however that the average pricing error would be higher if high frequency traders did not have a speed advantage. The crux for this result is that the optimal trading strategy of the informed investor is different in both cases.

[8]This caveat is important for the interpretation of empirical findings in light of our predictions. For instance, Hasbrouck and Saar (2012) find a negative effect of their proxy for high frequency trading on volatility and a positive effect on liquidity while our model predicts respectively no effect and a negative effect of HFTNs on these variables. However, Hasbrouck and Saar (2012)'s proxy does not specifically capture the high frequency trades triggered by the arrival of news. Thus, it may be a very noisy proxy for HFTNs' trades.

Our paper is related to the growing theoretical literature on high frequency trading[9]. Our analysis is most related to Biais, Foucault, Moinas (2011) and Jovanovic and Menkveld (2011) who also build upon the idea that high frequency traders have a speed advantage in getting access to information. These models are static. Therefore they do not analyze the optimal dynamic trading strategy of an investor with fast access to news while this analysis is central to our paper. Our approach is helpful to understand dynamic relationships between returns and high-frequency traders' flows. One drawback however is that it does not lend itself to welfare analysis since it relies on the existence of noise traders. Our paper is therefore silent on the social value of high frequency trading, which is the focus of Biais, Foucault, and Moinas (2011).

Technically, our model is related to Back and Perdersen (1998) (BP(1998)), Chau and Vayanos (2008) (CV(2008)), and Martinez and Rosu (2012) (MR (2012)). As in BP(1998), one investor receives a continuous flow of information ("news") on the final payoff of an asset (its fundamental value) and optimally trades with dealers. As in CV(2008), dealers receives news continuously as well, but not as precisely as the investor.[10] In contrast to both models, we assume that the informed investor observe news an infinitesimal amount of time before market-makers. This feature implies that the instantaneous variance of the informed investor's position becomes strictly positive. MR(2012) obtains a similar finding for a different reason. In their model, dealers receive no news. In this particular case, the news trading component would disappear in our model. This is not the case in Martinez and Rosu (2012) because the informed investor dislikes speculating on the long run value of the asset because of ambiguity aversion.

The paper is organized as follows. Section 2 describes our two models: the benchmark model, and the fast model. The models are set in continuous time, but in Appendix B we present the corresponding discrete versions. Section 3.1 describes the resulting equilibrium price process and trading strategies, and compares the various coefficients involved. Section 4 discusses empirical implications of the model. Section 5 concludes. All proofs are in Appendix. Appendix B analyzes the discrete time version of the model. The goal of this analysis is to show that that the continuous time model captures the effects obtained in a discrete time model in which news and trading decisions are very frequent.

[9] See, for instance, Cvitanic and Kirilenko (2011), Jovanovic and Menkveld (2011), Biais, Foucault, and Moinas (2011), Pagnotta and Philippon (2011), Cartea and Penalva (2012), or Hoffman (2012)

[10] We take the greater precision of information for the investor as given. As in Kim and Verrechia (1994), it could stem from greater processing ability for the informed investor.

2 Model

Trading for a risky asset occurs over the time interval $[0, 1]$. The liquidation value of the asset at time 1 is v_1. The risk-free rate is taken to be zero. Over the time interval $[0, 1]$, a single informed trader ("he") and uninformed noise traders submit market orders to a competitive market maker ("she"), who sets the price at which the trading takes place. The informed trader learns about the asset liquidation value, v_1, over time. His expectation of v_1 conditional on his information available until time t is denoted v_t. We refer to this estimate as the fundamental value of the asset at date t. This value follows a Gaussian process given by

$$v_t \;=\; v_0 + \int_0^t \mathrm{d}v_\tau, \quad \text{with} \quad \mathrm{d}v_t \;=\; \sigma_v \, \mathrm{d}B_t^v, \tag{1}$$

where v_0 is normally distributed with mean 0 and variance Σ_0, and B_t^v is a Brownian motion.[11] The informed trader observes v_0 at time 0 and, at each time $t + \mathrm{d}t \in [0, 1]$ observes $\mathrm{d}v_t$. We refer to this innovation in asset value as the news received by the informed trader at t.

The position of the informed trader in the risky asset at t is denoted by x_t. As the informed trader is risk-neutral, he chooses x_t (his "trading strategy") to maximize his expected profit at $t = 0$ given by

$$U_0 \;=\; \mathsf{E}\left[\int_0^1 (v_1 - p_{t+\mathrm{d}t}) \, \mathrm{d}x_t\right] \;=\; \mathsf{E}\left[\int_0^1 (v_1 - p_t - \; \mathrm{d}p_t) \, \mathrm{d}x_t\right], \tag{2}$$

where $p_{t+\mathrm{d}t} = p_t + \mathrm{d}p_t$ is the price at which the informed trader's order $\mathrm{d}x_t$ is executed.[12]

The aggregate position of the noise traders at t is denoted by u_t. It follows an

[11] This assumption can be justified as follows. First, define the asset value v_t as the full information price of the asset, i.e., the price that would prevail at t if all information until t were to become public. Then, v_t moves any time there is news, which should be interpreted not just as information from newswires, but more broadly as changes in other correlated prices or economic variables such as trades in other securities etc. For example, Brogaard et al.(2012), Jovanovic and Menkveld (2011) and Zhang (2012) show that the order flow of HFTs is correlated with changes in market-wide prices. Under this interpretation, v_t changes at a very high frequency, and can be assumed to be a continuous martingale, thus can be represented as an integral with respect to a Brownian motion (see the martingale representation theorem 3.4.2 in Karatzas and Shreve (1991)). Our representation (1) is then a simple particular case, with zero drift and constant volatility.

[12] Because the optimal trading strategy of the informed trader might have a stochastic component, we cannot set $\mathsf{E}(\mathrm{d}p_t \mathrm{d}x_t) = 0$ as, e.g., in the Kyle (1985) model.

exogenous Gaussian process given by

$$u_t \;=\; u_0 + \int_0^t \mathrm{d}u_\tau, \quad \text{with} \quad \mathrm{d}u_t \;=\; \sigma_v \,\mathrm{d}B_t^u, \tag{3}$$

where B_t^u is a Brownian motion independent from B_t^v.

The market maker also learns about the asset value from (a) public information and (b) trades. At $t + \mathrm{d}t$, she receives a noisy signal of the innovation in asset value:

$$\mathrm{d}z_t \;=\; \mathrm{d}v_t + \mathrm{d}e_t, \quad \text{with} \quad \mathrm{d}e_t \;=\; \sigma_e \,\mathrm{d}B_t^e, \tag{4}$$

where B_t^e is a Brownian motion independent from all the others. Thus, $\mathrm{d}z_t$ is the flow of news received by the market-maker at date $t + \mathrm{d}t$. Furthermore, the market-maker learn information from the aggregate order flow:

$$\mathrm{d}y_t \;=\; \mathrm{d}u_t + \mathrm{d}x_t, \tag{5}$$

because $\mathrm{d}x_t$ will reflect the information possessed by the informed trader(see below). We denote the market-maker's expectation of the asset liquidation value just before the trade at date $t{+}\mathrm{d}t$ by q_t. As the market-maker is competitive and risk-neutral, the trade at date $t{+}\mathrm{d}t$ takes place at a price equal to his expectation of the asset value conditional on all information up to date $t + \mathrm{d}t$, including the information contained in the trade at this date (as in Kyle (1985), BP (1998) or CV(2007)). We denote this transaction price by $p_{t+\mathrm{d}t}$. As in Kyle (1985), one can interpret q_t as the bid-ask midpoint just before the transaction over $[t, t + dt]$.[13]

If $\sigma_e > 0$, the news received by the market-maker are less precise than those received by the informed trader. Thus, one advantage of the informed investor over the market-maker is that he can form a more precise forecast of the asset payoff than the market-maker, at any point in time. As in Kim and Verrechia (1994), this advantage could stem from the fact that the informed investor is better able to process news than the dealers.

Our focus here is on the second advantage for the informed investor: the possibility to trade on news faster than the market-maker. To analyze this speed advantage and isolate its effects, we consider two different models: the *benchmark model* and the *fast*

[13]This interpretation is correct if the price impact is increasing in the signed order flow and a zero order flow has zero price impact. These conditions are satisfied in the linear equilibrium we consider in Section 3.1.

model. They differ in the timing with which the informed investor and the dealer receive news. A simplified timing of each model is presented in Table 1. Figure 2 shows the exact sequence of quotes and prices in each model.

Table 1: Timing of events during $[t, t + \mathrm{d}t]$ in the benchmark model and in the fast model

Benchmark Model	Fast Model
1. Informed trader observes news ($\mathrm{d}v_t$)	1. Informed trader observes news ($\mathrm{d}v_t$)
2. Market maker observes news ($\mathrm{d}z_t = \mathrm{d}v_t + \mathrm{d}e_t$)	2. Trading
3. Trading	3. Market maker observes news ($\mathrm{d}z_t = \mathrm{d}v_t + \mathrm{d}e_t$)

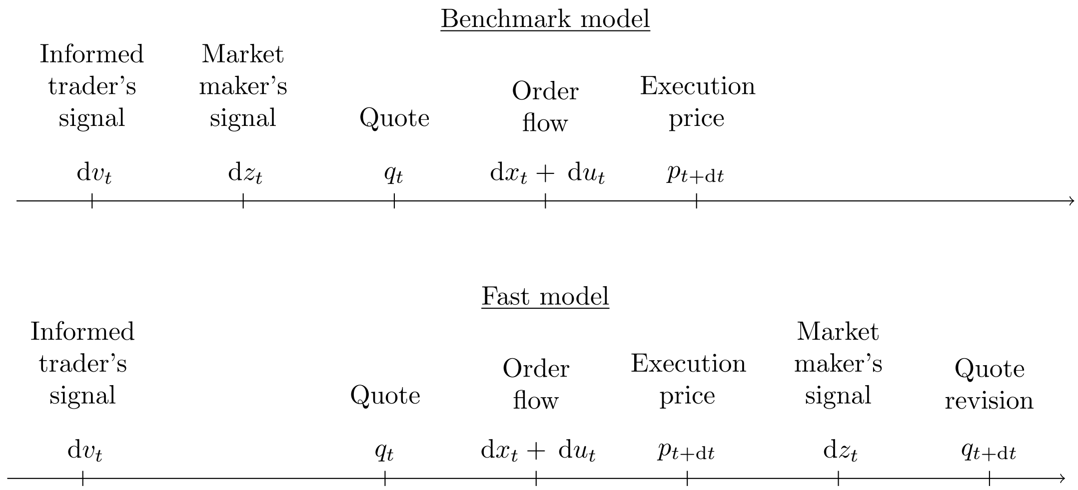
In the benchmark model, the order of events during the time interval $[t, t+\mathrm{d}t]$ is as follows. First, the informed trader observes $\mathrm{d}v_t$ and the market maker receives the signal $\mathrm{d}z_t$. The market maker sets her quote q_t based on her information set $\mathcal{I}_t \cup \mathrm{d}z_t$, where $\mathcal{I}_t \equiv \{z_\tau\}_{\tau \le t} \cup \{y_\tau\}_{\tau \le t}$, which comprises the order flow and the market maker's news until time t, and the news just received at date $t+dt$. Then, the informed trader and the noise traders submit their market orders and the aggregate order flow, $\mathrm{d}y_t = \mathrm{d}x_t + \mathrm{d}u_t$ is realized. The information set of the market maker when she sets the execution price $p_{t+\mathrm{d}t}$ is therefore $\mathcal{I}_t \cup \mathrm{d}z_t \cup \mathrm{d}y_t$. That is, $p_{t+\mathrm{d}t}$ differs from q_t because it reflects the information contained in the order flow at date $t + dt$.

In the fast model, the informed trader can trade on news faster than the market-maker. Namely, when the market-maker executes the order flow $\mathrm{d}y_t$, she does not yet observe the news $\mathrm{d}z_t$ while the informed investor has already observed the innovation in the asset value, $\mathrm{d}v_t$. More specifically, over the interval $[t, t+dt]$, the informed trader first observes $\mathrm{d}v_t$, submits his market order $\mathrm{d}x_t$ along with the noise traders' orders $\mathrm{d}u_t$ and the market-maker executes the aggregate order flow at at price $p_{t+\mathrm{d}t}$, which is her conditional expectation of the asset payoff on the information set $\mathcal{I}_t \cup \mathrm{d}y_t$. After trading has taken place and before the next trade, the market maker receives the signal $\mathrm{d}z_t$ and updates his estimate of the asset payoff based on the information set $\mathcal{I}_t \cup \mathrm{d}z_t \cup \mathrm{d}y_t$. Thus, the mid-quote $q_{t+\mathrm{d}t}$ at the beginning of the next trading round is the market-maker's expectation of the asset payoff conditional on $\mathcal{I}_t \cup \mathrm{d}z_t \cup \mathrm{d}y_t$.

To sum up, in the benchmark model:

$$q_t = \mathsf{E}\left[v_1 \mid \mathcal{I}_t \cup \mathrm{d}z_t\right] \quad \text{and} \quad p_{t+\mathrm{d}t} = \mathsf{E}\left[v_1 \mid \mathcal{I}_t \cup \mathrm{d}z_t \cup \mathrm{d}y_t\right] \tag{6}$$

Figure 2: Timing of events



while in the fast model:

$$q_t \;=\; \mathsf{E}\left[v_1 \mid \mathcal{I}_t\right] \quad \text{and} \quad p_{t+\mathrm{d}t} \;=\; \mathsf{E}\left[v_1 \mid \mathcal{I}_t \cup \mathrm{d}y_t\right] \tag{7}$$

Thus, in the benchmark model, the dealer and the informed investor observe news (innovations in the asset value) at the same speed but not with the same precision (unless $\sigma_e = 0$). This information structure is standard in models of informed trading following Kyle (1985) and also in empirical applications (see Hasbrouck (1991)). By contrast, in the fast model, the informed trader observes news a split second before the dealer. Thus, he also has a speed advantage relative to the dealer. Otherwise the benchmark model and the fast model are identical. Hence, by contrasting the properties of the benchmark model and the fast model, we can isolate the effects of high frequency traders' ability to react to news relatively faster than other market participants.

3 Optimal News Trading

In this section, we first derive the equilibrium of the benchmark model and the fast model. We then use the characterization of the equilibrium in each case to compare the properties of the informed investor's trades in each case.

3.1 Equilibrium

The equilibrium concept is similar to that of Kyle (1985) or Back and Pedersen (1998). That is, (a) the informed investor's trading strategy is optimal given dealers' pricing policy and (b) dealers' pricing policy follows equations (6) or (7) (depending on the model) with $\mathrm{d}y_t = \mathrm{d}u_t + \mathrm{d}x_t^*$ where $\mathrm{d}x_t^*$ is the optimal trading strategy for the informed investor. As usual in the literature using the Kyle (1985)'s framework, we look for equilibria in which prices are linear functions of the order flow and the informed investor's optimal trading strategy at date t (dx_t) is a linear function of his forecast of the asset value and the news he receives at date t.

More specifically, in the benchmark model, we look for an equilibrium in which the dealer's quote revision is linear in the public information he receives while the price impact is linear in the order flow. That is,

$$q_t = p_t + \mu_t^B \,\mathrm{d}z_t \quad and \quad p_{t+\mathrm{d}t} = q_t + \lambda_t^B \,\mathrm{d}y_t, \tag{8}$$

where index B denotes a coefficient in the Benchmark case. In the fast model, we look for an equilibrium in which the transaction price, $p_{t+\mathrm{d}t}$, is linear in the order flow as in equation (8) and the subsequent quote revision is linear in the *unexpected part* of the market maker's news. That is,

$$p_{t+\mathrm{d}t} = q_t + \lambda_t^F \,\mathrm{d}y_t \quad and \quad q_{t+\mathrm{d}t} = p_{t+\mathrm{d}t} + \mu_t^F(\,\mathrm{d}z_t - \rho_t \,\mathrm{d}y_t), \tag{9}$$

where $\rho_t \,\mathrm{d}y_t$ is the dealer's expectation of the public information arriving over the $[t, t+dt]$ conditional on the order flow over this period and index F refers to the value of a coefficient in the Fast model. In the fast model, $\rho_t > 0$ because, as shown below, the informed investor's optimal trade at date t depends on the news received at this date $(\,\mathrm{d}v_t)$. Thus, the dealer can forecast news from the order flow.

In both the benchmark and the fast model, we look for an equilibrium in which the informed investor's trading strategy is of the form

$$\mathrm{d}x_t = \beta_t^k(v_t - q_t)\,\mathrm{d}t + \gamma_t^k \,\mathrm{d}v_t. \quad for \quad k \in \{F, B\}. \tag{10}$$

That is, we solve for β_t^k and γ_t^k so that the strategy defined in equation (10) maximizes the informed trader's expected profit (2). More generally, one may look for linear equilibria in which $\mathrm{d}x_t = \int_0^t \beta_j^k dv_j + \alpha_t$. However, we show in In Appendix B that the optimal trading strategy for the informed investor in the discrete time version of our model is necessarily as in equation (10) when the market-maker's pricing rule is linear. It is therefore natural to restrict our attention to this type of strategy in the continuous time version of the model.

The trading strategy of the informed investor at, say, date t has two components. The first component ($\beta_t(v_t - p_t)\,\mathrm{d}t$) is proportional to the difference between the forecast of the asset value by the informed investor and the estimation of this value by the market-maker prior to the trade at date $t + dt$. Intuitively, the informed investor buys when the dealer underestimates the fundamental value and sells otherwise. This component is standard in models of trading with asymmetric information a la Kyle (1985) such as Back and Pedersen (1998), Back, Cao, and Willard (2000), etc. In what follows, we refer to this component as being the *level trading* component.

The second component of the informed investor's trading strategy is proportional to the news he receives at date t. We call it the news trading component. The next theorem shows that, in equilibrium, the news trading component is zero in the benchmark case ($\gamma_t^B = 0$) while it is strictly positive in the case in which the informed investor has a speed advantage in reacting to news ($\gamma_t^F = 0$). As explained in details below (see section 3.2), this difference implies that the informed investor's trades have very different properties when he is fast and when he is not. More generally, Theorem 1 provides a characterization of the equilibrium (coefficients μ_t^k, λ_t^k, ρ_t, β_t^k, and γ_t^k) in both the benchmark and the fast cases.

Theorem 1. *In the benchmark model there is a unique linear equilibrium:*

$$\mathrm{d}x_t = \beta_t^B (v_t - p_t)\,\mathrm{d}t + \gamma_t^B\,\mathrm{d}v_t, \tag{11}$$

$$\mathrm{d}p_t = \mu_t^B\,\mathrm{d}z_t + \lambda_t^B\,\mathrm{d}y_t, \tag{12}$$

with coefficients given by

$$\beta_t^B = \frac{1}{1-t}\,\frac{\sigma_u}{\Sigma_0^{1/2}}\left(1+\frac{\sigma_v^2\sigma_e^2}{\Sigma_0(\sigma_v^2+\sigma_e^2)}\right)^{1/2}, \tag{13}$$

$$\gamma_t^B = 0, \tag{14}$$

$$\lambda_t^B = \frac{\Sigma_0^{1/2}}{\sigma_u}\left(1+\frac{\sigma_v^2\sigma_e^2}{\Sigma_0(\sigma_v^2+\sigma_e^2)}\right)^{1/2}, \tag{15}$$

$$\mu_t^B = \frac{\sigma_v^2}{\sigma_v^2+\sigma_e^2}. \tag{16}$$

In the fast model there is a unique linear equilibrium:[14]

$$\mathrm{d}x_t = \beta_t^F(v_t-q_t)\,\mathrm{d}t+\gamma_t^F\,\mathrm{d}v_t, \tag{17}$$

$$\mathrm{d}q_t = \lambda_t^F\,\mathrm{d}y_t+\mu_t^F(\,\mathrm{d}z_t-\rho_t^F\,\mathrm{d}y_t), \tag{18}$$

with coefficients given by

$$\beta_t^F = \frac{1}{1-t}\,\frac{\sigma_u}{(\Sigma_0+\sigma_v^2)^{1/2}}\,\frac{1}{\left(1+\frac{\sigma_e^2}{\sigma_v^2}f\right)^{1/2}}\left(1+\frac{(1-f)\sigma_v^2}{\Sigma_0}\,\frac{1+\frac{\sigma_e^2}{\sigma_v^2}+\frac{\sigma_e^2}{\sigma_v^2}f}{2+\frac{\sigma_e^2}{\sigma_v^2}+\frac{\sigma_e^2}{\sigma_v^2}f}\right), \tag{19}$$

$$\gamma_t^F = \frac{\sigma_u}{\sigma_v}f^{1/2} = \frac{\sigma_u}{(\Sigma_0+\sigma_v^2)^{1/2}}\,\frac{\left(1+\frac{\sigma_e^2}{\sigma_v^2}f\right)^{1/2}(1+f)}{2+\frac{\sigma_e^2}{\sigma_v^2}+\frac{\sigma_e^2}{\sigma_v^2}f}, \tag{20}$$

$$\lambda_t^F = \frac{(\Sigma_0+\sigma_v^2)^{1/2}}{\sigma_u}\,\frac{1}{\left(1+\frac{\sigma_e^2}{\sigma_v^2}f\right)^{1/2}(1+f)}, \tag{21}$$

$$\mu_t^F = \frac{1+f}{2+\frac{\sigma_e^2}{\sigma_v^2}+\frac{\sigma_e^2}{\sigma_v^2}f}, \tag{22}$$

$$\rho_t^F = \frac{\sigma_v^2}{\sigma_u(\Sigma_0+\sigma_v^2)^{1/2}}\,\frac{(1+\frac{\sigma_e^2}{\sigma_v^2}f)^{1/2}}{2+\frac{\sigma_e^2}{\sigma_v^2}+\frac{\sigma_e^2}{\sigma_v^2}f}, \tag{23}$$

and f is the unique root in $(0,1)$ of the cubic equation

$$f = \frac{\left(1+\frac{\sigma_e^2}{\sigma_v^2}f\right)(1+f)^2}{\left(2+\frac{\sigma_e^2}{\sigma_v^2}+\frac{\sigma_e^2}{\sigma_v^2}f\right)^2}\,\frac{\sigma_v^2}{\sigma_v^2+\Sigma_0}. \tag{24}$$

In both models, when $\sigma_v \to 0$, the equilibrium converges to the unique linear equilibrium in the continuous time version of Kyle (1985).

[14] Note that the level trading component in (17) has q_t instead of p_t. This is the same formula, since (9) implies $(p_t-q_t)\,\mathrm{d}t=0$. We use q_t as a state variable, because it is a well defined Itô process.

The news trading component of the informed investor is non zero only if he has a speed advantage (and $\sigma_e < +\infty$ and $\sigma_v > 0$; see below). The reason for this important difference between the fast model and the benchmark model is as follows. In the fast model, the informed investor observes news an instant before the market-maker. Thus, as long as $\sigma_e < +\infty$, he can forecast how the dealer will adjust his quotes in the very short run (equation (9) describes this adjustment) and trades on this knowledge, that is, buy just before an increase in price due to good news ($dv_t > 0$) or sell just before a decrease in prices due to bad news ($dv_t < 0$). As a result, $\gamma_t^F > 0$ if $\sigma_e < +\infty$. In contrast, in the benchmark case, the dealer incorporates news in his quotes before executing the informed investor's trade. As a result, the latter cannot exploit any very short-run predictability in prices and, for this reason, $\gamma_t^B = 0$.

Whether he is fast or not, the informed investor can form a forecast of the long run value of the asset, v_1, that is more precise than that of the market-maker both because he starts with an informational advantage (he knows v_0) and because he receives more informative news (if $\sigma_e > 0$). The informed investor therefore also exploits the market-maker's pricing (or forecast) error, $v_t - q_t$. As usual, the trading strategy exploiting this advantage is to buy the asset when the market-maker's pricing error is positive: $v_t - q_t > 0$ and to sell it otherwise. For this reason, the level trading component of the strategy is strictly present whether the informed investor has a speed advantage or ($\beta_t^k > 0$ for $k \in \{F, B\}$).

Interestingly, the two components of the strategy can dictate trades in opposite directions. For instance, the level trading component may call for additional purchases of the asset (because $v_t - q_t > 0$) when the news trading component calls for selling it (because $dv_t < 0$). The net direction of the informed investor's trade is determined by the sum of these two desired trades. Interestingly, if the investor delegates the implementation of the two components of his trading strategy to two different agents (trading desks), one may see trades in opposite directions for these agents. Yet, they are part of an optimal trading strategy. Also, the two strategies cannot be considered independently in the sense that the sensitivity of the investor's trading strategy is optimally smaller when he has a speed advantage, as shown by the next corollary.

Corollary 1. *For all values of the parameters and at each date:* $\beta_t^F < \beta_t^B$.

Thus, in the fast model, the informed investor always exploits less aggressively the market-maker's pricing error than in the benchmark case. In a sense, he substitutes profits from this source with profits from trading on news. The intuition for this *substitution effect* is that trading more on news now reduces future profits from trading on the market-maker's forecast error. Therefore, the informed investor optimally reduces the size of the trade exploiting the market-maker's forecast error when he starts trading on news. As explained in Section 4, this substitution effect has an impact on the nature of price discovery. The next corollary describes how the sensitivities of the informed investor's trades to the market-maker's forecast error and news vary with the exogenous parameters of the model.

Corollary 2. *In both the benchmark equilibrium and the fast equilibrium, β_t^k increases in σ_v, σ_u, σ_e; Moreover, in the fast equilibrium, γ^F increases in σ_v, σ_u; and decreases in σ_e, Σ_0.*

An increase in σ_v or σ_u increases the informed investor's informational advantage. In the first case because news are more important (innovations in the asset value have a larger size) and in the second case because the order flow is noisier, other things equal. Thus, the informed investor reacts to an increase in these parameters by trading more aggressively both on the market-maker's forecast error and news.

An increase in σ_e implies that the market-maker receives noisier news. Accordingly, it becomes more difficult for the informed investor to forecast very short run price changes by the market-maker. Hence, γ^F decreases with σ_e and goes to zero when σ_e goes to infinity. Thus, there is no news trading if the dealers do not receive news. Furthermore, as the market-maker receives less accurate information, it becomes easier for the informed investor to exploit his more precise forecast of the final payoff of the asset. Accordingly, the informed investor's trades become more sensitive to the market-maker's forecast error (β_t^F increases).

When σ_e goes to $+\infty$, everything is as if the market-maker never receives public information, as in Back and Pedersen (1998), since news for the market-maker becomes uninformative. The equilibrium of the benchmark model in this case is identical to that obtained in Back and Pedersen (1998). If furthermore $\sigma_v = 0$, the informed investor receives no news and the benchmark case is then identical to the continuous time version of the Kyle (1985) model. In either case, the equilibrium of the fast model is identical to that of the benchmark case. In particular, even if the informed investor receives news faster than the dealer, his trading strategy will not feature a news trading component if the dealer does not receive news (γ_t^F goes to zero when σ_e goes to ∞). Another polar case is the case in which $\sigma_e = 0$. In this case, the information contained in news is very short-lived for the informed investor. As implied by Corollary 2, the informed investor then trades very aggressively on news (γ_t^F is maximal when $\sigma_e = 0$).

3.2 The Trades of High Frequency News' Traders

We now show that the behavior of the informed investor's order flow better coincides with stylised facts about high frequency traders when he has a speed advantage than when he has not.

The position of the informed investor, x_t is a stochastic process. The drift of this process is equal to the level trading component while the volatility component of this process is determined by the news trading component. As the latter is zero in the benchmark case, the informed investor's trades at the high frequency (that is, the instantaneous change in the informed investor's position) are negligible relative to those of noise traders (they are of the order of $\mathrm{d}t$ while noise traders' trades are of the order of $\sqrt{(}dt)$). In contrast, in the benchmark case, the informed investor's trades are of the same order of magnitude as those of noise traders, even at the high frequency. Thus, as shown on Figure 1, position of the informed investor is much more volatile than in the benchmark case. Accordingly, over a short time interval, the fraction of total trading volume due to the informed investor is much higher when he has a speed advantage. To see this formally, let *the Informed Participation Rate (IPR)* be the contribution of the informed trader to total trading volume over a short time interval. That is:

$$IPR_t = \frac{\mathsf{Var}(\,\mathrm{d}x_t)}{\mathsf{Var}(\,\mathrm{d}y_t)} = \frac{\mathsf{Var}(\,\mathrm{d}x_t)}{\mathsf{Var}(\,\mathrm{d}u_t) + \mathsf{Var}(\,\mathrm{d}x_t)} \tag{25}$$

Corollary 3. *The informed participation rate is zero when he has no speed advantage while it is equal to:*

$$IPR_t^B = 0, \qquad IPR_t^F = \frac{f}{1+f}, \tag{26}$$

when he has as speed advantage, where f is defined in Theorem 1.

The direction of the market-maker's forecast error persists over time because the informed investor slowly exploits his private information (as in Kyle (1985) or Back and Pedersen (2007)). As a result, the level trading component of the informed investor's trading strategy commands trades in the same direction for a relatively long period of time. This feature is a source of positive autocorrelation in the informed investor's order flow. However, when the informed investor has a speed advantage, over short time interval, trades exploiting the market-maker's forecast error are negligible relative to those exploiting the short-run predictability in prices due to news arrival. As these trades have no serial correlation (since news are not serially correlated), the autocorrelation of the informed order investor's order flow is smaller in the fast model. In fact the next result shows that, over very short time intervals, this autocorrelation is zero.

Corollary 4. *Over short time intervals, the autocorrelation of the informed order flow is strictly positive when the informed investor has no speed advantage and nil when the he has a speed advantage. Formally,for* $\tau > 0$*,*

$$\mathsf{Corr}(\,\mathrm{d}x_t^B,\,\mathrm{d}x_{t+\tau}^B) \;=\; \left(\frac{1-t-\tau}{1-t}\right)^{\frac{1}{2}+\lambda^B\beta_0^B} \;>\; 0, \tag{27}$$

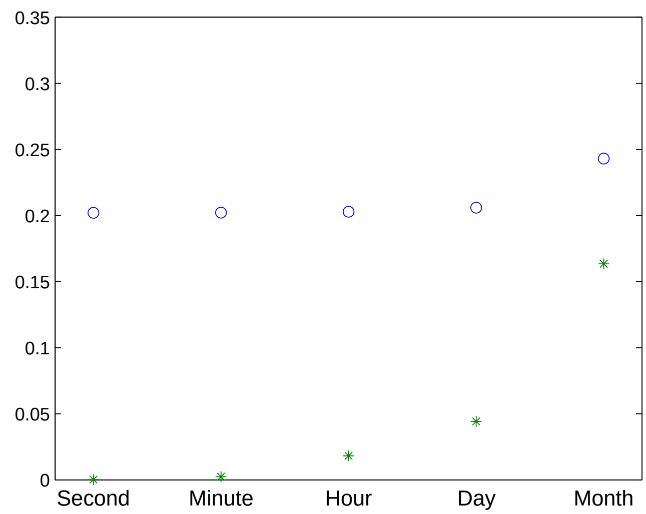
$$\mathsf{Corr}(\,\mathrm{d}x_t^F,\,\mathrm{d}x_{t+\tau}^F) \;=\; 0. \tag{28}$$

Corollaries 3 and 4 hold when the order flow of the informed investor is measured over an infinitesimal period of time. Econometricians often work with aggregated trades over some time interval (e.g., 10 seconds), due to limited data availability or by choice, to make data analysis more manageable.[15] In Appendix B, we show that the previous results are still qualitatively valid when the informed investor's trades are aggregated over time interval of arbitrary length (in this case, the informed investor's order flow over a given time interval is the sum of all of his trades over this time interval). In particular it is still the case that the informed investor's participation rate is higher while the autocorrelation of his order flow is smaller when he has a speed advantage. The only difference is that as flows are measured over longer time interval, the informed investor's participation rate and the autocorrelation of his trades increase, both in the fast model and in the benchmark case. Indeed, the trades that the informed investor conducts to exploit the market-maker's forecast error are positively autocorrelated and therefore account for an increasing fraction of his net order flow over longer time intervals. However, at relatively high frequencies (e.g. daily), the participation rate of the informed investor remains low when he has no speed advantage, as shown on Figure 3.

[15] For instance, Zhang (2012) aggregate the trades by HFTs in their sample over interval of 10 seconds. However, trades in her sample happen at a higher frequency.

Thus, the model in which the informed investor has no speed advantage does not explain well why high frequency traders account for a large fraction of the trading volume.

Figure 3: Informed participation rate at various trading frequencies. The figure plots the fraction of the trading volume due to the informed trader in a discrete time model for various lengths of time between trading periods (second, minute, hour, day, month) in (a) the benchmark model, marked with "$*$"; and (b) the fast model, marked with "$\circ$". The parameters used are $\sigma_u = \sigma_v = \sigma_e = \Sigma_0 = 1$ (see Theorem 1). The liquidation date $t = 1$ corresponds to 10 calendar years.



Using US stock trading data aggregated across twenty-six HFTs, Brogaard (2011) find a positive autocorrelation of the aggregate HFT order flow, which is consistent both with the benchmark model and the model in which the informed investor has a speed advantage, provided the sampling frequency is not too high. In addition, our model implies that this autocorrelation should decrease with the sampling frequency (see Corollary 12 in Appendix B). In contrast, Menkveld (2011) using data on a single HFT in the European stock market, and Kirilenko, Kyle, Samadi, and Tuzun (2011) using data on the Flash Crash of May 2010, find evidence of mean reverting positions for HFTs. One possibility is that the HFTs in their study do not trade on news as in our model. In fact, Menkveld (2011) shows that the HFT in his dataset behaves very much as a market-maker rather than an informed investor. Alternatively, HFTNs could face inventory constraints due to risk management concerns. This feature is absent from our model. Such constraints would naturally lead to mean reversion in the informed investor's trades.

Some empirical papers also find that aggressive orders by HFTs (that is, marketable orders) have a very short run positive correlation with subsequent returns (see Brogaard, Hendershott and Riordan (2012) and Kirilenko, Kyle, Samadi, and Tuzun (2011)). This finding is consistent with our model when the informed investor has a speed advantage but not otherwise. To see this, let AT_t (which stands for *Anticipatory Trading (AT)*) be the correlation between the informed order flow at a given date and the next instant return, that is:

$$AT_t = \mathsf{Corr}(\,\mathrm{d}x_t, q_{t+\,\mathrm{d}t} - p_{t+\,\mathrm{d}t}), \tag{29}$$

where we recall that $p_{t+\,\mathrm{d}t}$ is the price at which the order flow $\mathrm{d}x_t$ is executed, and $q_{t+\,\mathrm{d}t}$ is the quote posted by the dealer after observing news at date $t+\,\mathrm{d}t$.

Corollary 5. *Anticipatory trading is zero when the informed investor has no speed advantage while it is strictly positive when has a speed advantage:*

$$AT_t^B = 0, \qquad AT_t^F = \frac{(1-\rho^F\gamma^F)\sigma_v}{\sqrt{(1-\rho^F\gamma^F)^2\sigma_v^2 + \sigma_e^2 + (\rho^F)^2\sigma_u^2}} > 0. \tag{30}$$

When the informed investor observes news an instant before the market-maker, his order flow over a short period of time is mainly determined by the direction of incoming news. Thus, his trades anticipate on the adjustment of his quotes by the market-maker, which creates a short run positive correlation between the trades of the informed investor and subsequent returns, as observed in reality.[16]

In Appendix B, we analyze how this result generalizes when the sampling frequency used by the econometrician is lower than the frequency at which the informed investor trades on news. We show (see Corollary 13 in Appendix B) that the correlation between the aggregate order flow of the informed investor over an interval of time of fixed length and the asset return over the next time interval (of equal length) declines with the frequency at which the investor trades and goes to zero when the informed trading frequency goes to infinity (as in the continuous time model). Thus, the choice of a sampling frequency to study high frequency news trading is not innocuous can affect inferences. If this frequency is too low relative to the frequency at which trades take place (which by definition is very high for high frequency traders), it will be more difficult to detect the presence of anticipatory trading by the informed investor.

[16] Anticipatory trading in our model refers to the ability of the informed investor to trade ahead of incoming news. The term "anticipatory trading" is sometimes used to refer to trades ahead of of or alongside other investors, for instance institutional investors (see Hirschey (2011)). This form of anticipatory trading is not captured by our model.

4 Implications

4.1 High Frequency Trading on News and Market Quality

Controversies about high frequency traders focus on the effects of their speed advantage on liquidity, price discovery and price volatility. In this section, we study how the informed investor's ability to react to news faster than the market-maker (i.e., high frequency trading on news) affects measures of market quality. To this end, we compare these measures when the informed investor has a speed advantage and when he has not, holding the precision of the informed investor's information constant.

As in Kyle (1985), we measure market illiquidity by the immediate price impact of a trade, that is, λ^k.

Corollary 6. *Liquidity is lower when the informed investor has a speed advantage:* $\lambda^F > \lambda^B$.

Trades by the informed investor exposes the market-maker to adverse selection because the informed investor has a more accurate forecast of the asset liquidation value than the market-maker. Thus, the market-maker tends to accumulate a short position when she underestimates the asset value and a long position when she overestimates the asset value. This source of adverse selection is present both when the investor has a speed advantage and when he has not. However, adverse selection is stronger when the informed investor has a speed advantage because he can also buy just in advance of positive news and sell in advance of negative news. As a result, market illiquidity is higher when the informed investor has a speed advantage.[17].

Next, we consider the effect of HFTN on price discovery. We measure price discovery at any given point in time t by the average squared pricing error, that is,

$$\Sigma_t \;=\; \mathsf{E}\big((v_t - p_t)^2\big). \tag{31}$$

The smaller is Σ_t, the higher is informational efficiency.

Lemma 1. *The average pricing error at each date can be written:*

$$\Sigma_t \;=\; (1+t)\Sigma_0 + 2t\sigma_v^2 - 2\int_0^t \mathsf{Cov}(\,\mathrm{d}p_\tau, v_{\tau+\,\mathrm{d}\tau}). \tag{32}$$

[17] This implication is consistent with Hendershott and Moulton (2011).

As:[18]

$$\mathsf{Cov}(\,\mathrm{d}p_t, v_{t+\mathrm{d}t}) \;=\; \mathsf{Cov}(\,\mathrm{d}p_t, v_t - p_t) + \mathsf{Cov}(\,\mathrm{d}p_t,\,\mathrm{d}v_t). \tag{33}$$

we deduce that price discovery improves when short run changes in prices are more correlated with (a) news ($\mathsf{Cov}(\,\mathrm{d}p_t,\,\mathrm{d}v_t)$ increases) and (b) the direction of the market-maker's forecast error ($\mathsf{Cov}(\,\mathrm{d}p_t, v_t - p_t)$ increases). Interestingly, granting a speed advantage to the informed investor has opposite effects on these two dimensions of price discovery.

Corollary 7. *Short run changes in prices are more correlated with innovations in the asset value (*$\mathsf{Cov}(\mathrm{d}p_t, \mathrm{d}v_t)$ *is higher) but less correlated with the market-maker's pricing error (*$\mathsf{Cov}(\mathrm{d}p_t, v_t)$ *is smaller) when the informed investor has a speed advantage; Overall,* Σ_t *is identical whether or not the informed investor has a speed advantage.*

Hence, the speed advantage of the informed investor does not increase or reduce pricing errors on average. However, it changes the *nature* of price discovery in the short run. In a nutshell, returns are more informative about the level of the asset value in the benchmark model, while they are more informative about changes in the asset value in the fast model. The reason is as follows. In the benchmark, the contemporaneous correlation between changes in the price and innovations in the asset value comes from quote revisions only:

$$\mathsf{Cov}(\,\mathrm{d}p_t^B,\,\mathrm{d}v_t) \;=\; \mathsf{Cov}(\mu^B\,\mathrm{d}z_t,\,\mathrm{d}v_t) \;=\; \mu^B\sigma_v^2\,\mathrm{d}t. \tag{34}$$

In the fast model, this correlation is higher because the informed investor aggressively trades on the innovation in the asset value (the news trading component in his strategy is not zero). That is:

$$\mathsf{Cov}(\mathrm{d}p_t^F, \mathrm{d}v_t) \;=\; \mathsf{Cov}(\lambda^F \mathrm{d}x_t^F + \mu^F(\mathrm{d}z_t - \rho^F \mathrm{d}x_t), \mathrm{d}v_t) \;=\; \big(\mu^B + (\lambda^F - \mu^F\rho^F)\big)\sigma_v^2\mathrm{d}t. \tag{35}$$

As $\lambda^F > \mu^F\rho^F$, it implies that returns are more correlated with the innovations of the asset value in the fast model. By contrast, the covariance of returns with the market-maker's forecast error is higher in the benchmark model. The reason is that the informed investor trades less aggressively on the market-maker's forecast error when he has a speed advantage ($\beta_t^F < \beta_t^B$), as shown in Corollary 1.

[18] In this equation, $\mathrm{d}p_t$ denotes $p_{t+\mathrm{d}t} - p_t$ in the benchmark model, and $q_{t+\mathrm{d}t} - q_t$ in the fast model.

In equilibrium, these two effects exactly cancel out so that eventually the pricing error is the same in both models. In the fast model, new information is incorporated more quickly into the price while older information is incorporated less quickly, leaving the total pricing error equal in both models.[19]

We now consider the effect of HFTN on the volatility of short run returns, $\mathsf{Var}(\mathrm{d}p_t)$. This volatility has two sources in our model: trades and news for the market-maker. The second source volatility is reflected into the quote adjustments of the market-maker in-between trades. Thus, following Hasbrouck (1991a, 1991b), we decompose price volatility into the volatility coming from trades and the volatility coming from quotes:

$$\mathsf{Var}(\mathrm{d}p_t) \;=\; \mathsf{Var}(p_{t+\,\mathrm{d}t} - q_t) + \mathsf{Var}(q_t - p_t). \tag{36}$$

The first term in this variance decomposition capture the effect of trades on volatility while the second term captures the effect of quote revisions due to news arrivals for the market-maker.

Corollary 8. *The volatility of returns is equal to* $\mathsf{Var}(\,\mathrm{d}p_t) \;=\; \sigma_v^2 + \Sigma_0$. *whether the informed investor has a speed advantage or not. However, trades contribute to a larger fraction of this volatility when the informed investor has a speed advantage.*

Thus, the speed advantage of the informed investor alters the contribution of each source of volatility. In the fast model, trades contribute more to volatility since trades are more informative on impending news. The flip side is that the market-maker's quote is less sensitive to news. Thus, the contribution of quote revisions the short run return volatility is lower in the fast model. These two effects cancel each other in equilibrium so that volatility is the same in both models.[20]

[19]The more formal reason why these two effects exactly offset each other is that, in both the benchmark and the fast models, the informed trader finds it optimal to release information at a constant rate to minimize price impact. Therefore, Σ_t decreases linearly over time in both models. Moreover, the transversality condition for optimization requires that no money is left on the table at $t = 1$, i.e., $\Sigma_1 = 0$. Since the initial value Σ_0 is exogenously given, the evolution of Σ_t is the same in both models.

[20]This must be the case when changes in price are not autocorrelated. Actually, in this case, the short-term price variance per unit of time must be equal to the long-term price variance per unit of time, which is itself equal to the variance per unit of time of the exogenous asset value. In our model, changes in prices are not autocorrelated because prices are equal to the conditional expected value of the asset for the market-maker, as in Kyle (1985).

4.2 The Determinants of High Frequency News Traders' Activity

Empirical findings suggest that the activity of high frequency traders vary across stocks (e.g., Brogaard et al. (2012) find that HFTs are more active in large cap stocks than small cap stocks). Our model suggests two possible important determinants of the activity of high frequency news traders (measured by their participation rate as defined in Equation (25)): (i) the precision of the public information received by market-makers and (ii) the informational content of the news received by the informed investor.

Following Kim and Verrechia (1994), we measure the precision of public information by σ_e since a smaller σ_e^2 means that the news received by the market-maker provide a more precise signal about innovations in the asset value.[21] Moreover we measure the volume of trading by $Var(dy_t)$, a measure of the average absolute order imbalance in each transaction.

Corollary 9. *In the fast model, an increase in the precision of public news, i.e., a decrease in σ_e, results in a (i) higher participation of the informed investor (IPR_t^F), (b) higher trading volume, and (iii) higher liquidity. It has no effect when the informed investor has no speed advantage.*

When public information is more precise, the informed investor trades more aggressively on the news he receives as shown by Corollary 2. Indeed, the market-maker's quotes are then more sensitive to news (μ^F decreases in σ_e) and, as a result, the informed investor can better exploits his foreknowledge of news when he receives news faster than the dealer. As a result, he trades more over short-time interval so that his participation rate and trading volume increase.

An increase in the precision of public information has an ambiguous effect on the exposure to adverse selection for the market-maker. On the one hand, it increases the sensitivity of the informed investor's trade to news, which increases the exposure to adverse selection for the dealer. On the other hand, it helps the dealer to better forecast the asset liquidation value, which reduces his exposure to adverse selection. As shown by Corollary 9, the second effect always dominates so that illiquidity is reduced when the market-maker receives more precise news.

These findings are in sharp contrast with other models analyzing the effects of public information or corporate disclosures, such as Kim and Verrechia (1994). Indeed, in these

[21]Holding constant the variance of the innovation of the asset value σ_v^2, more precise public news about the changes in asset value amounts to a lower $\sigma_z^2 = \sigma_v^2 + \sigma_e^2$, or, equivalently, a lower σ_e^2.

models, an increase in the precision of public information leads to *less* trading volume (as informed investors trade less) and greater market liquidity. Furthermore, they suggest that controlling for the precision of public information is important to analyze the effect of high frequency trading on liquidity. Indeed, in our model, variations in the precision of public information can lead to a positive association between liquidity and the activity of high frequency news traders, despite the fact that, other things equal, high frequency news trading causes illiquitity to be higher in equilibrium (see Corollary 6).

To test the implications of Corollary 9, one needs one proxy for the precision of public news received by dealers. One possibility is to use "news sentiment scores" that are now provided by data vendors such as Reuters, Bloomberg, Dow Jones etc (see for instance GroB-KluBmann and Hautsch (2011)). These vendors report firm-specific news in real-time and assign a direction to the news (a proxy for the sign of dz_t) and a relevance score to news. The average relevance score of news about a firm (or a portfolio of firms) could be used as proxy for σ_e: firms with more relevant news should be firms for which public information is more precise.

In our model, the informed investor has two sources of information: (i) his initial forecast v_0 and (ii) news about the asset value. His initial forecast is never disclosed to dealers and can be seen as "private information" in the traditional sense. In contrast, the news that the informed investor receives are partially revealed to dealers and his corresponding information advantage is very short lived if σ_e is small. When σ_v^2 increases, news become relatively more informative than private information for the informed investor since news account for a greater fraction of the total volatility of the liquidation value for the asset ($\frac{\sigma_v^2}{\sigma_v^2+\Sigma_0}$ increases in σ_v^2). At the same time, for a fixed value of σ_e^2, the news received by the market-maker are less informative.[22]. Thus, to analyze the effect of increasing the informational content of news for the informed investor *ceteris paribus*, the next corollary considers the effect of a change in σ_v^2, while holding constant the ratio σ_e^2/σ_v^2.

Corollary 10. *For a fixed σ_e/σ_v, an increase in the informational content of news for the informed investor (a higher σ_v) results in a (i) higher participation rate for the informed investor (increases IPR_t^F), (ii) higher trading volume, and (c) reduces liquidity when the informed investor has a speed advantage. It has no effect when the informed*

[22]Indeed, $Var(dv_t \mid dz_t) = \frac{\sigma_e^2}{\sigma_v^2+\sigma_e^2}$, which decreases with σ_e^2

investor has no speed advantage.

A greater value of σ_v^2 raises γ^F, the sensitivity of the informed investor's trades to news when he has a speed advantage. As a result the participation rate of the informed investor and trading volume increase. Simultaneously, the exposure to adverse selection for the market-maker increases and therefore illiquidity increases. One possible proxy for σ_v^2 is the number of times the news sentiment score for a given firm exceeds a certain threshold over a fixed period of time (say the day). Intuitively, firms for which this number is high are firms for which more information is released over time (that is firms for which $\frac{\sigma_v^2}{\sigma_v^2+\Sigma_0}$ is higher).

Observe that an increase in σ_v^2 is associated with a higher price volatility since $\mathsf{Var}(\mathrm{d}p_t) = \sigma_v^2 + \Sigma_0$ (see Corollary 8). Thus, an increase in the informational content of news received by HFTNs may lead to a positive association between the volatility of short-run return and the activity of high frequency news traders even though HFTNs have no causal effect on volatility in our model (Corollary 8).[23]

4.3 Measuring the information content of HFTNs trades

As shown by Corollary 8, trades contributes more to total volatility relative to public information when the informed investor has a speed advantage. Intuitively, the reason is that trades by the informed investor are more informative about news. To analyze this point more formally, we use $CPI_t(\tau)$, the covariance between the informed order flow per unit of time at t and the subsequent price change over the time interval $[t, t+\tau]$ for $\tau > 0$, as a measure of the informational content of trades by the informed investor. That is:[24]

$$CPI_t(\tau) \;=\; \mathsf{Cov}\left(\frac{\mathrm{d}x_t}{\mathrm{d}t}, p_{t+\tau} - p_t\right). \tag{37}$$

This method to measure trade informativeness is similar to that used by Koski and Michaely (2000) or Green (2004). They measure the cumulative return following a trade of a given size and infer that the informativeness of the trade is higher if the cumulative return is stronger and in the same direction as the trade. The same method can be used

[23] In line with this prediction, Chaboud et al.(2010) find that high frequency traders in their data are more active on days with high volatility but do not appear to cause higher volatility.

[24] The optimal strategy of the informed trader is of the type $\mathrm{d}x_t = \beta_t(v_t - p_t)\,\mathrm{d}t + \gamma_t\,\mathrm{d}v_t$. Thus, $CPI_t(\tau) \;=\; \beta_t\,\mathsf{Cov}(v_t - p_t, p_{t+\tau} - p_t) + \frac{1}{\mathrm{d}t}\gamma_t\,\mathsf{Cov}(\,\mathrm{d}v_t, p_{t+\tau} - p_t),.$ Note that the second term is well defined, because $\mathsf{Cov}(\,\mathrm{d}v_t, p_{t+\tau} - p_t)$ is of the order of $\mathrm{d}t$, since the asset value, v_t, is a Gaussian process. Using p_t or q_t in the definition of $CPI_t(\tau)$ is equivalent because the difference between the two is smaller than $p_{t+\tau} - p_t$ by an order of magnitude.

by empiricists to analyze the information content of trades by high frequency traders when they observe these trades (or have a proxy for these trades).

Corollary 11. *For all values of τ, the cumulative price impact of the informed investor's trade per unit of time is higher in the fast model than in the slow model ($CPI_t^F(\tau) > CPI_t^B(\tau)$). However, the difference in cumulative price impacts vanish as τ increases. In the benchmark model, the cumulative price impact is*

$$CPI_t^B(\tau) \;=\; k_1^B \left[1 - \left(1 - \frac{\tau}{1-t}\right)^{\lambda^B \beta_0^B}\right], \tag{38}$$

while in the fast model it is

$$CPI_t^F(\tau) \;=\; k_0^F + k_1^F \left[1 - \left(1 - \frac{\tau}{1-t}\right)^{(\lambda^F - \mu^F \rho^F)\beta_0^F}\right], \tag{39}$$

where

$$k_1^B \;=\; \beta_0^B \Sigma_0, \tag{40}$$

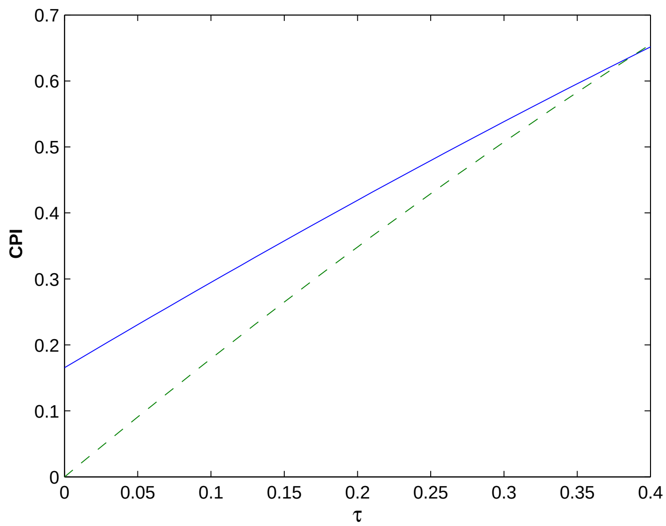
$$k_0^F \;=\; \gamma^F((\lambda^F - \mu^F \rho^F)\gamma^F + \mu^F)\sigma_v^2, \tag{41}$$

$$k_1^F \;=\; \beta_0^F \Sigma_0 + \gamma^F(1 - (\lambda^F - \mu^F \rho^F)\gamma^F - \mu^F)\sigma_v^2. \tag{42}$$

As shown in Figure 4, over short horizon, the informativeness of the investor's trades is markedly higher when he has a speed advantage. Indeed, when the informed investor has a speed advantage, his trades over short interval are driven by news ($\gamma^F > 0$). Thus, their informational content is not infinitesimal in contrast to the benchmark case. In both the fast model and the benchmark case, the cumulative price impact of a trade by the informed investor is gradually increasing over time. The reason is that the informed investor exploits the market-maker's forecast error by slowly accumulating a long (short) position if the market-maker underestimates (overestimates) the liquidation value of the asset. As a result, the cumulative return gradually increases over time following any trade by the informed investor. Thus, the intercept in Figure 4 is evidence of news trading, while the positive slope is evidence of level trading.

This gradual price adjustment is not a symptom of market inefficiency. Indeed, our measure of trade informativeness assumes that the econometrician can observe ex-post the trades of the informed investor. This information is not available to the market-

Figure 4: Cumulative Price Impact at Different Horizons. The figure plots the cumulative price impact at $t = 0$, $\mathsf{Cov}\left(\frac{\mathrm{d}x_0}{\mathrm{d}t}, p_\tau - p_0\right)$ against the horizon $\tau \in (0, 1]$ in (a) the benchmark model, with a dotted line; and (b) the fast model, with a solid line. The parameters used are $\sigma_u = \sigma_v = \sigma_e = \Sigma_0 = 1$ (see Theorem 1). The liquidation date $t = 1$ corresponds to 10 calendar years.



maker who just observes the aggregate order flow (dy_t). This point is important for the interpretation of empirical analyses relating returns to high frequency traders' order flows. Indeed, these flows are available to the empiricists but were not observed by market participants at the time of the trade.

This discussion suggests one drawback of using $CPI_t^k(\tau)$ as measure of trade informativeness. The gradual adjustment in prices following a trade at, say, date t by the informed investor in our model is due to additional trades following the trade at date t. Thus, the cumulative price impact really reflects the informational content of a string of orders and certainly exaggerates the informational content of each trade taken separately.

To better isolate the informational content of a trade while accounting for autocorrelation in trades, Hasbrouck (1991) advocates the use of Vector Autoregressive models. Some researchers (e.g., Zhang (2012) or Hirschley (2012)) have therefore used this approach to measure the informational content of high frequency traders' order flows. Following these papers, suppose again that the econometrician has data on Δx_t, the informed investor's order flow over a small time interval Δt and let $r_t = p_{t+1} - p_t$ be the return over this time interval. In the fast model, Theorem 1 implies that:

$$r_t = \lambda^F \Delta y_t + \zeta_t$$

where $\zeta_t = \mu^F(\Delta z_{t-\Delta t} - \rho_{t-\Delta t}\Delta y_{t-\Delta t})$. Now, using the fact that $\Delta y_t = \Delta x_t + \Delta u_t$ and the fact that $\Delta x_{t-\Delta t} \approx \gamma^F \Delta v_{t-\Delta t}$ when $\Delta t \to 0$, we deduce that the model in which the informed investor has a speed advantage implies the following VAR model for short run returns and trades by the informed investor:

$$r_t \approx \lambda^F \Delta x_t + \mu^F (\gamma^F)^{(-1)}(1 - \gamma^F \rho)\Delta x_{t-\Delta t} + \varepsilon_t.$$

$$\Delta x_{t-\Delta t} = \gamma^F \Delta v_{t-\Delta t}$$

where $\varepsilon_t = \lambda^F \Delta u_t - \mu\rho\Delta u_{t-\Delta t} + \mu\Delta e_{t-\Delta t}$.

Hasbrouck (1991) proposes to measure the informational content of a trade at date t by the sum of predicted quote revisions through a fixed number of steps, say m, after a trade innovation of a fixed size (this sum measures the "permanent impact of a trade"). In our case, when $m \geq 2$, this measure of informational content is $\lambda^F + \mu^F(\gamma^F)^{(-1)}(1 - \gamma^F\rho)$. It is always higher than λ^F because $\gamma^F\rho = \frac{f}{1+f} < 1$. Thus, it overestimates the true permanent impact of the informed investor's trade, that is, λ^F. The reason is that the lagged order flow of the informed investor appears in equation (4.3) because it is correlated with the news received by the dealer at date t and therefore the update in the dealer's quote at this date, not because these quotes slowly adjusts to information contained in past trades. This observation suggests that one must be careful in interpreting measures of the informational content of trades using the Hasbrouck (1991)'s approach when informed investors trade on public information. This problem is in fact discussed by Hasbrouck (1991) (see Section III in his paper). It may have become more severe in recent years with the development of high frequency trading on news.

5 Conclusions

To be written

A Proofs of Results

A.1 Proof of Theorem 1

Benchmark model: We compute the optimal strategy of the informed trader at $t+\mathrm{d}t$. As we have seen in the discrete version of the model, in Appendix B, we need to consider only strategies $\mathrm{d}x_\tau$ of the type $\mathrm{d}x_\tau = \beta_\tau(v_\tau - p_\tau)\,\mathrm{d}\tau + \gamma_\tau\,\mathrm{d}v_\tau$. Recall that $\mathcal{I}_t^p$ is the market maker's information set immediately after trading at t. If we denote by $\mathcal{J}_t^p = \mathcal{I}_t^p \cup \{v_\tau\}_{\tau \le t+\mathrm{d}t}$ the trader's information set before trading at $t+\mathrm{d}t$, the expected profit from trading after t is

$$\pi_t \;=\; \mathsf{E}\left(\int_t^1 (v_1 - p_{\tau+\mathrm{d}\tau})\,\mathrm{d}x_\tau \mid \mathcal{J}_t^p\right). \tag{43}$$

From (12), $p_{\tau+\mathrm{d}\tau} = p_\tau + \mu_\tau(\mathrm{d}v_\tau + \mathrm{d}e_\tau) + \lambda_\tau(\mathrm{d}x_\tau + \mathrm{d}u_\tau)$. For $\tau \ge t$, denote by

$$V_\tau \;=\; \mathsf{E}\big((v_\tau - p_\tau)^2 \mid \mathcal{J}_t^p\big). \tag{44}$$

Then the expected profit is

$$\begin{aligned}
\pi_t \;&=\; \mathsf{E}\left(\int_t^1 (v_\tau + \mathrm{d}v_\tau - p_\tau - \mu_\tau\,\mathrm{d}v_\tau - \lambda_\tau\,\mathrm{d}x_\tau)\,\mathrm{d}x_\tau \mid \mathcal{J}_t^p\right) &(45)\\
&=\; \int_t^1 \Big(\beta_\tau V_\tau + (1-\mu_\tau-\lambda_\tau\gamma_\tau)\gamma_\tau\sigma_v^2\Big)\,\mathrm{d}\tau. &(46)
\end{aligned}$$

V_τ can be computed recursively:

$$\begin{aligned}
V_{\tau+\mathrm{d}\tau} \;&=\; \mathsf{E}\big((v_{\tau+\mathrm{d}\tau} - p_{\tau+\mathrm{d}\tau})^2 \mid \mathcal{J}_t^p\big)\\
&=\; \mathsf{E}\big((v_\tau + \mathrm{d}v_\tau - p_\tau - \mu_\tau\,\mathrm{d}v_\tau - \mu_\tau\,\mathrm{d}e_\tau - \lambda_\tau\,\mathrm{d}x_\tau - \lambda_\tau\,\mathrm{d}u_\tau)^2 \mid \mathcal{J}_t^p\big) &(47)\\
&=\; V_\tau + (1-\mu_\tau-\lambda_\tau\gamma_\tau)^2\sigma_v^2\,\mathrm{d}\tau + \mu_\tau^2\sigma_e^2\,\mathrm{d}\tau + \lambda_\tau^2\sigma_u^2\,\mathrm{d}\tau - 2\lambda_t\beta_t V_\tau\,\mathrm{d}\tau.
\end{aligned}$$

therefore the law of motion of V_τ is a first order differential equation

$$V_\tau' \;=\; -2\lambda_t\beta_t V_\tau + (1-\mu_\tau-\lambda_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + \lambda_\tau^2\sigma_u^2, \tag{48}$$

or equivalently $\beta_\tau V_\tau = \frac{-V'_\tau + (1-\mu_\tau-\lambda_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + \lambda_\tau^2\sigma_u^2}{2\lambda_\tau}$. Substitute this into (43) and integrate by parts. Since $V_t = 0$, we get

$$\begin{aligned} \pi_t \;=\; & -\frac{V_1}{2\lambda_1} + \int_t^1 V_\tau \Big(\frac{1}{2\lambda_\tau}\Big)' \mathrm{d}\tau \\ & + \int_t^1 \left(\frac{(1-\mu_\tau-\lambda_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + \lambda_\tau^2\sigma_u^2}{2\lambda_\tau} + (1-\mu_\tau-\lambda_\tau\gamma_\tau)\gamma_\tau\sigma_v^2 \right) \mathrm{d}\tau. \end{aligned} \tag{49}$$

This is essentially the argument of Kyle (1985): we have eliminated the choice variable β_τ and replaced it by V_τ. Since $V_\tau > 0$ can be arbitrarily chosen, in order to get an optimum we must have $\big(\frac{1}{2\lambda_\tau}\big)' = 0$, which is equivalent to

$$\lambda_\tau \;=\; \text{constant}. \tag{50}$$

For a maximum, the transversality condition $V_1 = 0$ must be also satisfied.

We next turn to the choice of γ_τ. The first order condition is

$$-(1-\mu_\tau-\lambda_\tau\gamma_\tau) + (1-\mu_\tau-\lambda_\tau\gamma_\tau) - \lambda_\tau\gamma_\tau \;=\; 0 \quad\Longrightarrow\quad \gamma_\tau \;=\; 0. \tag{51}$$

Thus, there is no flow trading in the benchmark model. Note also that the second order condition is $\lambda_\tau > 0$.[25]

Next, we derive the pricing rules from the market maker's zero profit conditions. The equations $p_t = \mathsf{E}(v_1|\mathcal{I}_t^p)$ and $q_t = \mathsf{E}(v_1|\mathcal{I}_t^p, \mathrm{d}z_t)$ imply that $q_t = p_t + \mu_t\,\mathrm{d}z_t$, where

$$\mu_t \;=\; \frac{\mathsf{Cov}(v_1,\, \mathrm{d}z_t \mid \mathcal{I}_t^p)}{\mathsf{Var}(\,\mathrm{d}z_t \mid \mathcal{I}_t^p)} \;=\; \frac{\mathsf{Cov}(v_0 + \int_0^1 \mathrm{d}v_\tau,\, \mathrm{d}v_t + \,\mathrm{d}e_t \mid \mathcal{I}_t^p)}{\mathsf{Var}(\,\mathrm{d}v_t + \,\mathrm{d}e_t \mid \mathcal{I}_t^p)} \;=\; \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2}. \tag{52}$$

The equations $q_t = \mathsf{E}(v_1|\mathcal{I}_{t+\mathrm{d}t}^q)$ and $p_{t+\mathrm{d}t} = \mathsf{E}(v_1|\mathcal{I}_{t+\mathrm{d}t}^q, \mathrm{d}y_t)$ imply that $p_{t+\mathrm{d}t} = q_t + \lambda_t \mathrm{d}y_t$, where

$$\lambda_t \;=\; \frac{\mathsf{Cov}(v_1,\, \mathrm{d}y_t \mid \mathcal{I}_{t+\mathrm{d}t}^q)}{\mathsf{Var}(\,\mathrm{d}y_t \mid \mathcal{I}_{t+\mathrm{d}t}^q)} \;=\; \frac{\mathsf{Cov}(v_1, \beta_t(v_t - p_t)\,\mathrm{d}t + \,\mathrm{d}u_t \mid \mathcal{I}_{t+\mathrm{d}t}^q)}{\mathsf{Var}(\beta_t(v_t - p_t)\,\mathrm{d}t + \,\mathrm{d}u_t \mid \mathcal{I}_{t+\mathrm{d}t}^q)} \;=\; \frac{\beta_t\Sigma_t}{\sigma_u^2}, \tag{53}$$

where $\Sigma_t = \mathsf{E}\big((v_t - p_t)^2|\mathcal{I}_t^p\big)$.[26] The information set of the informed trader, $\mathcal{J}_t^p$, is a refinement of the market maker's information set, $\mathcal{I}_t^p$. Therefore, by the law of iterated

[25] The condition $\lambda_\tau > 0$ is also a second order condition with respect to the choice of β_τ. To see this, suppose $\lambda_\tau < 0$. Then if $\beta_\tau > 0$ is chosen very large, Equation (48) shows that V_τ is very large as well, and thus $\beta_\tau V_\tau$ can be made arbitrarily large. Thus, there would be no maximum.

[26] Because $\mathcal{I}_{t+\mathrm{d}t}^q = \mathcal{I}_t^p \cup \{\mathrm{d}z_t\}$, the two information sets differ only by the infinitesimal quantity $\mathrm{d}z_t$, and thus we can also write $\Sigma_t = \mathsf{E}\big((v_t - p_t)^2|\mathcal{I}_{t+\mathrm{d}t}^q\big) = \mathsf{E}\big((v_t - p_t)^2|\mathcal{I}_t^p\big)$.

expectations, Σ_t satisfies the same equation as V_t:

$$\Sigma_t' \;=\; -2\lambda_t\beta_t\Sigma_t + (1-\mu_t-\lambda_t\gamma_t)^2\sigma_v^2 + \mu_t^2\sigma_e^2 + \lambda_t^2\sigma_u^2, \tag{54}$$

except that it has a different initial condition. One can solve this differential equation explicitly and show that the transversality condition $V_1 = 0$ is equivalent to $\int_0^1 \beta_t \,\mathrm{d}t = +\infty$, and in turn this is equivalent to $\Sigma_1 = 0$. Since λ_t, μ_t and $\gamma_t = 0$ are constant, by (53) $\beta_t\Sigma_t$ is also constant. Equation (54) then implies that Σ_t' is constant. Since $\Sigma_1 = 0$, $\Sigma_t = (1-t)\Sigma_0$, and $\beta_t = \frac{\beta_0}{1-t}$. Finally, we integrate (54) between 0 and 1, and substituting for $\lambda = \frac{\beta_0\Sigma_0}{\sigma_u^2}$, $\mu = \frac{\sigma_v^2}{\sigma_v^2+\sigma_e^2}$, and $\gamma = 0$, we obtain β_0 and λ as stated in the Theorem.

Fast model: The informed trader has the same objective function as in (43):

$$\pi_t \;=\; \mathsf{E}\left(\int_t^1 (v_1 - p_{\tau+\,\mathrm{d}\tau})\,\mathrm{d}x_\tau \mid \mathcal{J}_t^p\right). \tag{55}$$

but here we use q_t instead of p_t as a state variable. From (8), $p_{t+\mathrm{d}t} = q_t + \lambda_t \mathrm{d}y_t$. Also, from (18), $q_{\tau+\,\mathrm{d}\tau} = q_\tau + \mu_\tau(\mathrm{d}z_\tau - \rho_\tau\mathrm{d}y_\tau) + \lambda_\tau(\mathrm{d}y_\tau)$, and we obtain

$$q_{\tau+\,\mathrm{d}\tau} \;=\; \mu_\tau\,\mathrm{d}z_\tau + m_\tau\,\mathrm{d}y_\tau, \quad \text{with} \tag{56}$$

$$m_\tau \;=\; \lambda_\tau - \mu_\tau\rho_\tau. \tag{57}$$

As we have seen in the discrete version of the model, in Appendix B, we need to consider only strategies $\mathrm{d}x_\tau$ of the type (17), $\mathrm{d}x_\tau = \beta_\tau(v_\tau - q_\tau)\,\mathrm{d}\tau + \gamma_\tau\,\mathrm{d}v_\tau$. For $\tau \geq t$, denote by

$$V_\tau \;=\; \mathsf{E}\big((v_\tau - q_\tau)^2 \mid \mathcal{J}_t^p\big). \tag{58}$$

The expected profit is

$$\pi_t \;=\; \mathsf{E}\left(\int_t^1 (v_\tau + \,\mathrm{d}v_\tau - q_\tau - \lambda_\tau\,\mathrm{d}x_\tau)\,\mathrm{d}x_\tau \mid \mathcal{J}_t^p\right) \tag{59}$$

$$\;=\; \int_t^1 \Big(\beta_\tau V_\tau + (1-\lambda_\tau\gamma_\tau)\gamma_\tau\sigma_v^2\Big)\,\mathrm{d}\tau. \tag{60}$$

V_τ is computed as in the benchmark model, except that λ_τ is replaced by m_τ:

$$\begin{aligned} V_{\tau+\mathrm{d}\tau} &= \mathsf{E}\big((v_{\tau+\mathrm{d}\tau} - q_{\tau+\mathrm{d}\tau})^2 \mid \mathcal{J}_t^p\big) \\ &= V_\tau + (1-\mu_\tau - m_\tau\gamma_\tau)^2\sigma_v^2\,\mathrm{d}\tau + \mu_\tau^2\sigma_e^2\,\mathrm{d}\tau + m_\tau^2\sigma_u^2\,\mathrm{d}\tau - 2m_t\beta_t V_\tau\,\mathrm{d}\tau. \end{aligned} \tag{61}$$

therefore the law of motion of V_τ is a first order differential equation

$$V_\tau' = -2m_t\beta_t V_\tau + (1-\mu_\tau - m_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + m_\tau^2\sigma_u^2, \tag{62}$$

or equivalently $\beta_\tau V_\tau = \frac{-V_\tau'+(1-\mu_\tau-m_\tau\gamma_\tau)^2\sigma_v^2+\mu_\tau^2\sigma_e^2+m_\tau^2\sigma_u^2}{2m_\tau}$. Substitute this into (43) and integrate by parts. Since $V_t = 0$, we get

$$\begin{aligned} \pi_t &= -\frac{V_1}{2m_1} + \int_t^1 V_\tau\Big(\frac{1}{2m_\tau}\Big)'\,\mathrm{d}\tau \\ &\quad + \int_t^1\left(\frac{(1-\mu_\tau - m_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + m_\tau^2\sigma_u^2}{2m_\tau} + (1-\lambda_\tau\gamma_\tau)\gamma_\tau\sigma_v^2\Big)\right)\,\mathrm{d}\tau. \end{aligned} \tag{63}$$

Since $V_\tau > 0$ can be arbitrarily chosen, in order to get an optimum we must have $\big(\frac{1}{2m_\tau}\big)' = 0$, which is equivalent to $m_\tau =$ constant. For a maximum, the transversality condition $V_1 = 0$ must be also satisfied.

We next turn to the choice of γ_τ. The first order condition is

$$-(1-\mu_\tau - m_\tau\gamma_\tau) + (1-\lambda_\tau\gamma_\tau) - \lambda_\tau\gamma_\tau = 0 \quad\Longrightarrow\quad \gamma_\tau = \frac{\mu_\tau}{2\lambda_\tau - m_\tau} = \frac{\mu_\tau}{\lambda_\tau + \mu_\tau\rho_\tau}. \tag{64}$$

Thus, we obtain a nonzero flow trading component. The second order condition is $\lambda_\tau + \mu_\tau\rho_\tau > 0$. There is also a second order condition with respect to β: $m_\tau > 0$: see Footnote 25.

Next, we derive the pricing rules from the market maker's zero profit conditions. As in the benchmark model, we compute

$$\lambda_t = \frac{\mathsf{Cov}_t(v_1,\,\mathrm{d}y_t)}{\mathsf{Var}_t(\,\mathrm{d}y_t)} = \frac{\mathsf{Cov}_t(v_1, \beta_t(v_t - p_t)\,\mathrm{d}t + \gamma_t\,\mathrm{d}v_t + \,\mathrm{d}u_t)}{\mathsf{Var}(\beta_t(v_t-p_t)\,\mathrm{d}t + \gamma_t\,\mathrm{d}v_t + \,\mathrm{d}u_t)} = \frac{\beta_t\Sigma_t + \gamma_t\sigma_v^2}{\gamma_t^2\sigma_v^2 + \sigma_u^2}, \tag{65}$$

$$\rho_t = \frac{\mathsf{Cov}_t(\,\mathrm{d}z_t,\,\mathrm{d}y_t)}{\mathsf{Var}_t(\,\mathrm{d}y_t)} = \frac{\gamma_t\sigma_v^2}{\gamma_t^2\sigma_v^2 + \sigma_u^2}, \tag{66}$$

$$\mu_t = \frac{\mathsf{Cov}_t(v_1,\,\mathrm{d}z_t - \rho_t\,\mathrm{d}y_t)}{\mathsf{Var}_t(\,\mathrm{d}z_t - \rho_t\,\mathrm{d}y_t)} = \frac{-\rho_t\beta_t\Sigma_t + (1-\rho_t\gamma_t)\sigma_v^2}{(1-\rho_t\gamma_t)^2\sigma_v^2 + \rho_t^2\sigma_u^2 + \sigma_e^2}. \tag{67}$$

By the same arguments as for the benchmark model, $\Sigma_t = (1-t)\Sigma_0$, $\beta_t = \frac{\beta_0}{1-t}$, and $\beta_t\Sigma_t$,

λ_t, ρ_t, μ_t are constant. Since Σ_t satisfies the same Equation (62) as V_t, and $\Sigma_t' = -\Sigma_0$, we obtain

$$-\Sigma_0 \;=\; -2m_t\beta_t\Sigma_t + (1-\mu_\tau - m_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + m_\tau^2\sigma_u^2. \tag{68}$$

We now define the following constants:

$$a \;=\; \frac{\sigma_u^2}{\sigma_v^2}, \quad b \;=\; \frac{\sigma_e^2}{\sigma_v^2}, \quad c \;=\; \frac{\Sigma_0}{\sigma_v^2}, \tag{69}$$

$$f \;=\; \frac{\gamma^2}{a}, \quad \tilde{\lambda} \;=\; \lambda\gamma, \quad \tilde{\rho} \;=\; \rho\gamma, \quad \nu = \frac{\beta_0\Sigma_0}{\sigma_u^2}\gamma, \quad \tilde{m} = m\gamma. \tag{70}$$

With these notations, Equations (64)–(68) become

$$\begin{aligned} \tilde{\lambda} \;&=\; \mu(1-\tilde{\rho}), \quad \tilde{\lambda} \;=\; \frac{\nu+f}{1+f}, \quad \tilde{\rho} \;=\; \frac{f}{1+f}, \quad \mu \;=\; \frac{1-\nu}{1+b(1+f)} \\ c \;&=\; \frac{2\nu}{f} - (1-\mu-\tilde{m})^2 - \mu^2 b - \frac{\tilde{m}^2}{f}. \end{aligned} \tag{71}$$

Substitute $\tilde{\lambda}$, $\tilde{\rho}$, μ in $\tilde{\lambda} = \mu(1-\tilde{\rho})$ and solve for ν:

$$\nu \;=\; \frac{1-(1+b)f - bf^2}{2+b+bf} \;=\; \frac{1+f}{2+b+bf} - f. \tag{72}$$

The other equations, together with $\tilde{m} = \tilde{\lambda} - \mu\tilde{\rho}$, imply

$$\begin{aligned} \tilde{\lambda} \;&=\; \frac{1}{2+b+bf}, \quad \tilde{\rho} \;=\; \frac{f}{1+f}, \quad \mu \;=\; \frac{1+f}{2+b+bf}, \quad \tilde{m} \;=\; \frac{1-f}{2+b+bf}, \\ 1+c \;&=\; \frac{(1+bf)(1+f)^2}{f(2+b+bf)^2}. \end{aligned} \tag{73}$$

Putting together (72) and the last equation in (73), we compute

$$\beta_0 \;=\; \frac{a}{c\gamma}\,\nu \;=\; \frac{a^{1/2}}{cf^{1/2}}\,\nu \;=\; \frac{a^{1/2}}{cf^{1/2}(1+c)}\,\frac{1}{(1+f)^{1/2}}\left(c + (1-f)\frac{1+b+bf}{2+b+bf}\right). \tag{74}$$

Now substitute a, b, c from (69) in Equations (73)–(74) and use $\gamma = a^{1/2}f^{1/2}$ to obtain Equations (19)–(24). One can also check that the second order conditions $\lambda+\mu\rho > 0$ and $m > 0$ are equivalent to $f \in (-1,1)$. Next, we show that the equation $1+c = \frac{(1+bf)(1+f)^2}{f(2+b+bf)^2}$ has a unique solution $f \in (-1,1)$, which in fact lies in $(0,1)$. This can be shown by noting that

$$F_b(f) \;=\; 1+c, \quad \text{with} \quad F_b(x) \;=\; \frac{(1+bx)(1+x)^2}{x(2+b+bx)^2}. \tag{75}$$

One verifies $F_b'(x) = \frac{(x+1)(x-1)(2+b+3bx)}{x^2(2+b+bx)^2}$, so $F_b(x)$ decreases on $(0,1)$. Since $F_b(0) = +\infty$ and $F_b(1) = \frac{1}{1+b} < 1$, there is a unique $f \in (0,1)$ so that $F_b(f) = 1 + c$.[27]

A.2 Proof of Corollary 1

To be written.

A.3 Proof of Corollary 2

To be written.

A.4 Proof of Corollary 3

In the benchmark model, $Var(dx_t) = (\beta_t^B)^2 \Sigma_t \mathrm{d}t^2$ and $Var(du_t) = \sigma_u^2 \mathrm{d}t$. Therefore, $IPR_t^B = 0$.

In the fast model, $Var(dx_t) = (\gamma_t^B)^2 \sigma_v^2 \mathrm{d}t$. Therefore, $IPR_t^F = (\gamma_t^B)^2\sigma_v^2/((\gamma_t^B)^2\sigma_v^2 + \sigma_u^2) = f/(f+1)$, using the equation for γ_t^F in Theorem 1.

A.5 Proof of Corollary 4

We start with a useful preliminary result:

Lemma 2. *In the benchmark model and in the fast model, for all $s < u$, we have*

$$Cov(v_s - p_s, v_u - p_u) = \Sigma_s \left(\frac{1-u}{1-s} \right)^{m\beta_0}, \tag{76}$$

$$Cov(dv_s, v_u - p_u) = (1 - m\gamma - \mu)\sigma_v^2 \left(\frac{1-u}{1-s} \right)^{m\beta_0} ds, \tag{77}$$

where $m \equiv \lambda - \mu\rho$.

Proof. We start from

$$\begin{aligned} Cov(v_s - p_s, v_u - p_u) &= Cov(v_s - p_s, v_s - p_s) - \int_s^u Cov(v_s - p_s, dp_h)\, dh \\ &= \Sigma_s - \int_s^u Cov(v_s - p_s, m\beta_h(v_h - p_h))\, dh \end{aligned}$$

[27] One can check that $F_b(x) = 1 + c$ has no solution on $(-1,0)$: When $b \leq 1$, $F_b(x) < 0$ on $(-1,0)$. When $b > 1$, $F_b(x)$ attains its maximum on $(-1,0)$ at $x^* = -\frac{2+b}{3b}$, for which $F_b(x^*) = \frac{(b-1)^3}{b(b+2)^3} < 1$.

Differentiating with respect to u we obtain

$$\frac{\partial}{\partial u} Cov(v_s - p_s, v_u - p_u) = -m\beta_u Cov(v_s - p_s, v_u - p_u),$$

which rewrites as

$$\frac{\partial}{\partial \tau} \log Cov(v_s - p_s, v_u - p_u) = -m\beta_u = -m\beta_0 \frac{1}{1-u} = m\beta_0 \frac{\partial}{\partial u} \log(1-u).$$

Integrating between s and u and using $Cov(v_s - p_s, v_s - p_s) = \Sigma_s$, we obtain equation (76).

Similarly, we have

$$\begin{aligned} Cov(dv_s, v_u - p_u) &= Cov(dv_s, dv_s - dp_s) - \int_s^u Cov(dv_s, dp_h)\, dh \\ &= (1 - m\gamma - \mu)\sigma_v^2 ds - \int_s^u m\beta_h Cov(dv_s, v_h - p_h)\, dh. \end{aligned}$$

Proceeding as above we obtain (77). $\square$

We can now prove Corollary 4. The formula in the benchmark model follows immediately from Equation (76). In the fast model, the auto-covariance of the order flow is of order $\mathrm{d}t^2$ while the variance is of order $\mathrm{d}t$, therefore the autocorrelation is of order $\mathrm{d}t$, which is zero in continuous time.

A.6 Proof of Corollary 5

To be written

A.7 Proof of Corollary 6

We use the notations from the proof of Theorem 1. We start by showing that $\mu^F < \mu^B$; by computation, $\frac{1+f}{2+b+bf} < \frac{1}{1+b}$ is equivalent to $f < 1$, which is true since $f \in (0,1)$.

We show that $\lambda^F > \lambda^B$, i.e., $\frac{(c+1)^{1/2}}{a^{1/2}} \frac{1}{(1+bf)^{1/2}(1+f)} > \frac{c^{1/2}}{a^{1/2}} \big(1 + \frac{b}{c(1+b)}\big)^{1/2}$. After squaring the two sides, and using $1+c = \frac{(1+bf)(1+f)^2}{f(2+b+bf)^2}$, we need to prove that $\frac{1}{f(2+b+bf)^2} > c + 1 - \frac{1}{1+b}$, or equivalently $\frac{1}{1+b} > \frac{(1+bf)(1+f)^2}{f(2+b+bf)^2} - \frac{1}{f(2+b+bf)^2}$. This can be reduced to proving $1 + b + (1-f)(1+bf) > 0$, which is true.

A.8 Proof of Corollary 7

A.9 Proof of Lemma 1

We start from

$$\Sigma_t = Var(v_t - p_t) = Var(v_t) + Var(p_t) - 2Cov(v_t, \int_0^t dp_\tau).$$

We have $Var(v_t) = \Sigma_0 + t\sigma_v^2$. Since the price is a martingale and given that we prove in the proof of Proposition **??** that the volatility of the price is equal to the volatility of the asset value, we have $Var(p_t) = t(\Sigma_0 + \sigma_v^2)$. Finally, using that price change cannot be correlated with future innovation in asset value, we obtain equation (32).

A.10 Proof of Corollary 8

In the benchmark model, $Var(p_{t+\mathrm{d}t} - q_t) = (\lambda^B)^2\sigma_u^2 \mathrm{d}t$ and $Var(q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t}) = (\mu^B)^2\sigma_v^2 \mathrm{d}t$. Using the equilibrium parameter values of Theorem 1 we obtain $Var(dp_t) = \Sigma_0 + \sigma_v^2 t$.

Similarly, in the fast model, $Var(p_{t+\mathrm{d}t} - q_t) = (\lambda^F)^2((\gamma^F)^2\sigma_v^2 + \sigma_u^2)\mathrm{d}t$ and $Var(q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t}) = (\mu^B)^2((1 - \rho^F\gamma^F)^2\sigma_v^2 + \sigma_e^2 + (\rho^F)^2\sigma_u^2)\mathrm{d}t$. Using the equilibrium parameter values of Theorem 1, we obtain that $Var(p_{t+\mathrm{d}t} - q_t)$ is higher than in the benchmark, $Var(q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t})$ is lower than in the benchmark, and $Var(dp_t) = \Sigma_0 + \sigma_v^2 t$ is the same as in the benchmark.

A.11 Proof of Corollary 9

By Corollary 3, $IPR^F = \frac{f}{1+f}$, thus the informed participation rate in the fast model has the same dependence on σ_e as f. From (20), $\gamma^F = \frac{\sigma_u}{\sigma_v} f^{1/2}$, thus f has the same dependence on σ_e as γ^F. Therefore, IPR^F has the same dependence on σ_e as γ^F. But Corollary 2 shows that γ^F is decreasing in σ_e. Finally, we use again Corollary 2 to show that λ^F is increasing in σ_e.

A.12 Proof of Corollary 8

To be written.

A.13 Proof of Corollary 11

Follows immediately from Lemma 2.

B Models in Discrete Time

B.1 Discrete Time Fast Model

We divide the interval $[0,1]$ into T equally spaced intervals of length $\Delta t = \frac{1}{T}$. Trading takes place at equally spaced times, $t = 1, 2, \ldots, T-1$. The sequence of events is as follows. At $t = 0$, the informed trader observes v_0. At each $t = 1, \ldots, T-1$, the informed trader observes $\Delta v_t = v_t - v_{t-1}$; and the market maker observes $\Delta z_{t-1} = \Delta v_{t-1} + \Delta e_{t-1}$, except at $t = 1$. The error in the market maker's signal is normally distributed, $\Delta e_{t-1} \sim \mathcal{N}(0, \sigma_e^2 \Delta t)$. The market maker quotes the bid price = the ask price = q_t. The informed trader then submits Δx_t, and the liquidity traders submit in aggregate $\Delta u_t \sim \mathcal{N}(0, \sigma_u^2 \Delta t)$. The market maker observes only the aggregate order flow, $\Delta y_t = \Delta x_t + \Delta u_t$, and sets the price at which the trading takes place, p_t. The market maker is competitive, i.e., makes zero profit. This translates into the following formulas:

$$p_t = \mathsf{E}(v_t \mid \mathcal{I}_t^p), \qquad \mathcal{I}_t^p = \{\Delta y_1, \ldots, \Delta y_t, \Delta z_1, \ldots, \Delta z_{t-1}\}, \tag{78}$$

$$q_{t+1} = \mathsf{E}(v_t \mid \mathcal{I}_t^q), \qquad \mathcal{I}_t^q = \{\Delta y_1, \ldots, \Delta y_t, \Delta z_1, \ldots, \Delta z_t\}. \tag{79}$$

We also denote

$$\Omega_t = \mathsf{Var}(v_t \mid \mathcal{I}_t^p), \tag{80}$$

$$\Sigma_t = \mathsf{Var}(v_t \mid \mathcal{I}_t^q). \tag{81}$$

Definition 1. *A pricing rule p_t is called linear if it is of the form $p_t = q_t + \lambda_t \Delta y_t$, for all $t = 1, \ldots, T-1$.*[28] *An equilibrium is called linear if the pricing rule is linear, and the informed trader's strategy Δx_t is linear in $\{v_\tau\}_{\tau \le t}$ and $\{q_\tau\}_{\tau \le t}$.*

The next result shows that if the pricing rule is linear, the informed trader's strategy is also linear, and furthermore it can be decomposed into a level trading component,

[28] We could defined more generally, a pricing rule to be linear in the whole history $\{\Delta y_\tau\}_{\tau \le t}$, but as Kyle (1985) shows, this is equivalent to the pricing rule being linear only in Δy_t.

$\beta_t(v_{t-1} - q_t)$, and a flow trading component, $\gamma_t \Delta v_t$.

Theorem 2. *Any linear equilibrium must be of the form*

$$
\begin{aligned}
\Delta x_t &= \beta_t(v_{t-1} - q_t)\Delta t + \gamma_t \Delta v_t, && (82)\\
p_t &= q_t + \lambda_t \Delta y_t, && (83)\\
q_{t+1} &= p_t + \mu_t(\Delta z_t - \rho_t \Delta y_t), && (84)
\end{aligned}
$$

for $t = 1, \ldots, T-1$, where β_t, γ_t, λ_t, μ_t, ρ_t, Ω_t, and Σ_t are constants that satisfy

$$
\begin{aligned}
\lambda_t &= \frac{\beta_t \Sigma_{t-1} + \gamma_t \sigma_v^2}{\beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2}, && (85)\\
\mu_t &= \frac{\left(\sigma_u^2 + \beta_t^2 \Sigma_{t-1}\Delta t - \beta_t \gamma_t \Sigma_{t-1}\right)\sigma_v^2}{\left(\beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2\right)\sigma_e^2 + \left(\beta_t^2 \Sigma_{t-1}\Delta t + \sigma_u^2\right)\sigma_v^2}, && (86)\\
m_t &= \lambda_t - \rho_t \mu_t \;=\; \frac{\beta_t \Sigma_{t-1}(\sigma_v^2 + \sigma_e^2) + \gamma_t \sigma_v^2 \sigma_e^2}{\left(\beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2\right)\sigma_e^2 + \left(\beta_t^2 \Sigma_{t-1}\Delta t + \sigma_u^2\right)\sigma_v^2}, && (87)\\
\rho_t &= \frac{\gamma_t \sigma_v^2}{\beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2}, && (88)\\
\Omega_t &= \Sigma_{t-1} + \sigma_v^2 \Delta t - \frac{\beta_t^2 \Sigma_{t-1}^2 + 2\beta_t \gamma_t \Sigma_{t-1}\sigma_v^2 + \gamma_t^2 \sigma_v^4}{\beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2}\,\Delta t, && (89)\\
\Sigma_t &= \Sigma_{t-1} + \sigma_v^2 \Delta t - \frac{\beta_t^2 \Sigma_{t-1}^2(\sigma_v^2 + \sigma_e^2) + \beta_t^2 \Sigma_{t-1}\Delta t \sigma_v^4 + \sigma_v^4 \sigma_u^2 + \gamma_t^2 \sigma_v^4 \sigma_e^2 + 2\beta_t \gamma_t \Sigma_{t-1}\sigma_v^2 \sigma_e^2}{\left(\beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2\right)\sigma_e^2 + \left(\beta_t^2 \Sigma_{t-1}\Delta t + \sigma_u^2\right)\sigma_v^2}\,\Delta t. && (90)
\end{aligned}
$$

The value function of the informed trader is quadratic for all $t = 1, \ldots, T-1$:

$$
\pi_t \;=\; \alpha_{t-1}(v_{t-1} - q_t)^2 + \alpha'_{t-1}(\Delta v_t)^2 + \alpha''_{t-1}(v_{t-1} - q_t)\Delta v_t + \delta_{t-1}. \qquad (91)
$$

The coefficients of the optimal trading strategy and the value function satisfy

$$
\begin{aligned}
\beta_t \Delta t &= \frac{1 - 2\alpha_t m_t}{2(\lambda_t - \alpha_t m_t^2)}, && (92)\\
\gamma_t &= \frac{1 - 2\alpha_t m_t(1 - \mu_t)}{2(\lambda_t - \alpha_t m_t^2)}, && (93)\\
\alpha_{t-1} &= \beta_t \Delta t(1 - \lambda_t \beta_t \Delta t) + \alpha_t(1 - m_t \beta_t \Delta t)^2, && (94)\\
\alpha'_{t-1} &= \alpha_t(1 - \mu_t - m_t \gamma_t)^2 + \gamma_t(1 - \lambda_t \gamma_t), && (95)\\
\alpha''_{t-1} &= \beta_t \Delta t + \gamma_t(1 - 2\lambda_t \beta_t \Delta t) + 2\alpha_t(1 - m_t \beta_t \Delta t)(1 - \mu_t - m_t \gamma_t), && (96)\\
\delta_{t-1} &= \alpha_t\left(m_t^2 \sigma_u^2 + \mu_t^2 \sigma_e^2\right)\Delta t + \alpha'_t \sigma_v^2 \Delta t + \delta_t. && (97)
\end{aligned}
$$

The terminal conditions are

$$\alpha_T \;=\; \alpha'_T \;=\; \alpha''_T \;=\; \delta_T \;=\; 0. \tag{98}$$

The second order condition is

$$\lambda_t - \alpha_t m_t^2 \;>\; 0. \tag{99}$$

Given Σ_0, conditions (85)–(99) are necessary and sufficient for the existence of a linear equilibrium.

Proof. First, we show that Equations (85)–(90) are equivalent to the zero profit conditions of the market maker. Second, we show that Equations (92)–(99) are equivalent to the informed trader's strategy (82) being optimal.

Zero Profit of Market Maker: Let us start with with the market maker's update due to the order flow at $t = 1, \ldots, T-1$. Conditional on $\mathcal{I}^q_{t-1}$, the variables $v_{t-1} - q_t$ and Δv_t have a bivariate normal distribution:

$$\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \mid \mathcal{I}^q_{t-1} \;\sim\; \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{t-1} & 0 \\ 0 & \sigma_v^2 \end{bmatrix} \right). \tag{100}$$

The aggregate order flow at t is of the form

$$\Delta y_t \;=\; \beta_t (v_{t-1} - q_t)\Delta t + \gamma_t \Delta v_t + \Delta u_t. \tag{101}$$

Denote by

$$\Phi_t \;=\; \mathsf{Cov}\left(\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix}, \Delta y_t \right) \;=\; \begin{bmatrix} \beta_t \Sigma_{t-1} \\ \gamma_t \sigma_v^2 \end{bmatrix} \Delta t. \tag{102}$$

Then, conditional on $\mathcal{I}_t = \mathcal{I}^q_{t-1} \cup \{\Delta y_t\}$, the distribution of $v_{t-1} - q_t$ and Δv_t is bivariate normal:

$$\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \mid \mathcal{I}_t \;\sim\; \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right), \tag{103}$$

where

$$\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \;=\; \Phi_t \,\mathsf{Var}(\Delta y_t)^{-1} \Delta y_t \;=\; \begin{bmatrix} \beta_t \Sigma_{t-1} \\ \gamma_t \sigma_v^2 \end{bmatrix} \frac{1}{\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2} \,\Delta y_t, \tag{104}$$

and

$$
\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \mathsf{Var}\left(\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix}\right) - \Phi_t \,\mathsf{Var}(\Delta y_t)^{-1}\Phi_t' \tag{105}
$$

$$
= \begin{bmatrix} \Sigma_{t-1} & 0 \\ 0 & \sigma_v^2 \Delta t \end{bmatrix} - \frac{1}{\beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2} \begin{bmatrix} \beta_t^2 \Sigma_{t-1}^2 & \beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 \\ \beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 & \gamma_t^2 \sigma_v^4 \end{bmatrix} \Delta t.
$$

We compute

$$
p_t - q_t = \mathsf{E}(v_t - q_t \mid \mathcal{I}_t) = \mu_1 + \mu_2 = \frac{\beta_t \Sigma_{t-1} + \gamma_t \sigma_v^2}{\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2} \Delta y_t, \tag{106}
$$

which proves Equation (85) for λ_t. Also,

$$
\begin{aligned}
\Omega_t &= \mathsf{Var}(v_t - q_t \mid \mathcal{I}_t) = \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2 \\
&= \Sigma_{t-1} + \sigma_v^2 \Delta t - \frac{\beta_t^2 \Sigma_{t-1}^2 + 2\beta_t\gamma_t \Sigma_{t-1}\sigma_v^2 + \gamma_t^2 \sigma_v^4}{\beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2\sigma_v^2 + \sigma_u^2} \Delta t,
\end{aligned} \tag{107}
$$

which proves (89).

Next, to compute $q_{t+1} = \mathsf{E}(v_t \mid \mathcal{I}_t^q)$, we start from the same prior as in (100), but we consider the impact of both the order flow at t and the market maker's signal at $t+1$:

$$
\Delta y_t = \beta_t (v_{t-1} - q_t)\Delta t + \gamma_t \Delta v_t + \Delta u_t, \tag{108}
$$

$$
\Delta z_t = \Delta v_t + \Delta e_t. \tag{109}
$$

Denote by

$$
\Psi_t = \mathsf{Cov}\left(\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix}, \begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix}\right) = \begin{bmatrix} \beta_t \Sigma_{t-1} & 0 \\ \gamma_t \sigma_v^2 & \sigma_v^2 \end{bmatrix} \Delta t, \tag{110}
$$

$$
V_t^{yz} = \mathsf{Var}\left(\begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix}\right) = \begin{bmatrix} \beta_t^2 \Sigma_{t-1}\Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2 & \gamma_t \sigma_v^2 \\ \gamma_t \sigma_v^2 & \sigma_v^2 + \sigma_e^2 \end{bmatrix} \Delta t. \tag{111}
$$

Conditional on $\mathcal{I}_t^q = \mathcal{I}_{t-1}^q \cup \{\Delta y_t, \Delta z_t\}$, the distribution of $v_{t-1} - q_t$ and Δv_t is bivariate normal:

$$
\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \mid \mathcal{I}_t^q \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}\right), \tag{112}
$$

where

$$\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \Psi_t \, (V_t^{yz})^{-1} \begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix} = \frac{\begin{bmatrix} \beta_t \Sigma_{t-1}(\sigma_v^2 + \sigma_e^2)\Delta y_t \; - \; \beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 \Delta z_t \\ \gamma_t \sigma_v^2 \sigma_e^2 \Delta y_t \; + \; (\beta_t^2 \Sigma_{t-1} \Delta t + \sigma_u^2)\sigma_v^2 \Delta z_t \end{bmatrix}}{(\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2)\sigma_e^2 + (\beta_t^2 \Sigma_{t-1} \Delta t + \sigma_u^2)\sigma_v^2}, \quad (113)$$

and

$$\begin{aligned} \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} &= \mathsf{Var}\left(\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \right) - \Psi_t \, (V_t^{yz})^{-1} \, \Psi_t' \\ &= \begin{bmatrix} \Sigma_{t-1} & 0 \\ 0 & \sigma_v^2 \Delta t \end{bmatrix} - \frac{\begin{bmatrix} \beta_t^2 \Sigma_{t-1}^2 (\sigma_v^2 + \sigma_e^2) & \beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 \sigma_e^2 \\ \beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 \sigma_e^2 & (\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_e^2 + \sigma_u^2)\sigma_v^4 \end{bmatrix}}{(\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2)\sigma_e^2 + (\beta_t^2 \Sigma_{t-1} \Delta t + \sigma_u^2)\sigma_v^2} \Delta t. \end{aligned} \quad (114)$$

Therefore,

$$\begin{aligned} q_{t+1} - q_t &= \mu_1 + \mu_2 \\ &= \frac{\big(\beta_t \Sigma_{t-1}(\sigma_v^2 + \sigma_e^2) + \gamma_t \sigma_v^2 \sigma_e^2\big)\Delta y_t \; + \; \big(\sigma_u^2 + \beta_t^2 \Sigma_{t-1} \Delta t - \beta_t \gamma_t \Sigma_{t-1}\big)\sigma_v^2 \Delta z_t}{\big(\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2\big)\, \sigma_e^2 + \big(\beta_t^2 \Sigma_{t-1} \Delta t + \sigma_u^2\big)\, \sigma_v^2} \quad (115) \\ &= m_t \Delta y_t \; + \; \mu_t \Delta z_t \; = \; (\lambda_t - \rho_t \mu_t)\Delta y_t \; + \; \mu_t \Delta z_t, \quad (116) \end{aligned}$$

which proves Equations (86), (87), and (88). Also,

$$\begin{aligned} \Sigma_t &= \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2 \\ &= \Sigma_{t-1} + \sigma_v^2 \Delta t - \frac{\beta_t^2 \Sigma_{t-1}^2 (\sigma_v^2 + \sigma_e^2) + \beta_t^2 \Sigma_{t-1} \Delta t \sigma_v^4 + \sigma_v^4 \sigma_u^2 + \gamma_t^2 \sigma_v^4 \sigma_e^2 + 2\beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 \sigma_e^2}{\big(\beta_t^2 \Sigma_{t-1} + (\beta_t + \gamma_t)^2 \sigma_v^2 + \sigma_u^2\big)\, \sigma_e^2 + \big(\beta_t^2 \Sigma_{t-1} + \sigma_u^2\big)\, \sigma_v^2} \Delta t, \end{aligned} \quad (117)$$

which proves (90).

Optimal Strategy of Informed Trader: At each $t = 1, \ldots, T-1$, the informed trader maximizes the expected profit: $\pi_t = \max \sum_{\tau=t}^{T-1} \mathsf{E}\big((v_T - p_\tau)\Delta x_\tau\big)$. We prove by backward induction that the value function is quadratic and of the form given in (91): $\pi_t = \alpha_{t-1}(v_{t-1} - q_t)^2 + \alpha'_{t-1}(\Delta v_t)^2 + \alpha''_{t-1}(v_{t-1} - q_t)\Delta v_t + \delta_{t-1}$. At the last decision point ($t = T-1$) the next value function is zero, i.e., $\alpha_T = \alpha'_T = \alpha''_T = \delta_T = 0$, which are the terminal conditions (98). This is the transversality condition: no money is left on the table. In the induction step, if $t = 1, \ldots, T-1$, we assume that π_{t+1} is of the desired form. The Bellman principle of intertemporal optimization implies

$$\pi_t \;=\; \max_{\Delta x} \; \mathsf{E}\Big((v_t - p_t)\Delta x + \pi_{t+1} \mid \mathcal{I}_t^q, v_t, \Delta v_t\Big). \tag{118}$$

Equations (83) and (84) show that the quote q_t evolves by $q_{t+1} = q_t + m_t \Delta y_t + \mu_t \Delta z_t$, where $m_t = \lambda_t - \rho_t \mu_t$. This implies that the informed trader's choice of Δx affects the trading price and the next quote by

$$p_t \;=\; q_t + \lambda_t(\Delta x + \Delta u_t), \tag{119}$$

$$q_{t+1} \;=\; q_t + m_t(\Delta x + \Delta u_t) + \mu_t \Delta z_t. \tag{120}$$

Substituting these into the Bellman equation, we get

$$\begin{aligned}
\pi_t \;=\; & \max_{\Delta x} \; \mathsf{E}\Big(\Delta x(v_{t-1} + \Delta v_t - q_t - \lambda_t \Delta x - \lambda_t \Delta u_t) \\
& + \; \alpha_t(v_{t-1} + \Delta v_t - q_t - m_t \Delta x - m_t \Delta u_t - \mu_t \Delta v_t - \mu_t \Delta e_t)^2 \; + \; \alpha_t' \Delta v_{t+1}^2 \qquad (121) \\
& + \; \alpha_t''(v_{t-1} + \Delta v_t - q_t - m_t \Delta x - m_t \Delta u_t - \mu_t \Delta v_t - \mu_t \Delta e_t)\Delta v_{t+1} \; + \; \delta_t\Big) \\
= \; & \max_{\Delta x} \; \Delta x(v_{t-1} - q_t + \Delta v_t - \lambda_t \Delta x) \\
& + \; \alpha_t\Big((v_{t-1} - q_t - m_t \Delta x + (1-\mu_t)\Delta v_t)^2 + (m_t^2 \sigma_u^2 + \mu_t^2 \sigma_e^2)\Delta t\Big) \; + \; \alpha_t' \sigma_v^2 \Delta t \\
& + \; 0 \; + \; \delta_t.
\end{aligned} \tag{122}$$

The first order condition with respect to Δx is

$$\Delta x \;=\; \frac{1 - 2\alpha_t m_t}{2(\lambda_t - \alpha_t m_t^2)}(v_{t-1} - q_t) \;+\; \frac{1 - 2\alpha_t m_t(1-\mu_t)}{2(\lambda_t - \alpha_t m_t^2)}\Delta v_t, \tag{123}$$

and the second order condition for a maximum is $\lambda_t - \alpha_t m_t^2 > 0$, which is (99). Thus, the optimal Δx is indeed of the form $\Delta x_t = \beta_t(v_{t-1} - q_t)\Delta t + \gamma_t \Delta v_t$, where $\beta_t \Delta t$ and

γ_t are as in Equations (92) and (93). We substitute Δx_t in the formula for π_t to obtain

$$\begin{aligned}
\pi_t \;=\;& \Big(\beta_t \Delta t(1-\lambda_t\beta_t\Delta t) + \alpha_t(1-m_t\beta_t\Delta t)^2\Big)(v_{t-1}-q_t)^2 \\
&+ \Big(\alpha_t(1-\mu_t-m_t\gamma_t)^2 + \gamma_t(1-\lambda_t\gamma_t)\Big)\Delta v_t^2 \\
&+ \Big(\beta_t\Delta t + \gamma_t(1-2\lambda_t\beta_t\Delta t) + 2\alpha_t(1-m_t\beta_t\Delta t)(1-\mu_t-m_t\gamma_t)\Big)(v_{t-1}-q_t)\Delta v_t \\
&+ \alpha_t\big(m_t^2\sigma_u^2 + \mu_t^2\sigma_e^2\big)\Delta t \;+\; \alpha_t'\sigma_v^2\Delta t \;+\; \delta_t.
\end{aligned} \tag{124}$$

This proves that indeed π_t is of the form $\pi_t = \alpha_{t-1}(v_{t-1}-q_t)^2 + \alpha'_{t-1}(\Delta v_t)^2 + \alpha''_{t-1}(v_{t-1}-q_t)\Delta v_t + \delta_{t-1}$, with α_{t-1}, α'_{t-1}, α''_{t-1} and δ_{t-1} as in Equations (94)–(97). $\square$

We now briefly discuss the existence of a solution for the recursive system given in Theorem 2. The system of equations (85)–(97) can be numerically solved backwards, starting from the boundary conditions (98). We also start with an arbitrary value of $\Sigma_T > 0$.[29] By backward induction, suppose α_t and Σ_t are given. One verifies that Equation (90) implies

$$\Sigma_{t-1} \;=\; \frac{\Sigma_t\big(\sigma_v^2\sigma_u^2 + \sigma_v^2(\sigma_u^2+\gamma_t^2\sigma_e^2)\big) - \sigma_v^2\sigma_u^2\sigma_e^2\Delta t}{\big(\sigma_u^2\sigma_e^2 + \sigma_v^2(\sigma_u^2+\gamma_t^2\sigma_e^2) + \beta_t^2\Delta t^2\sigma_v^2\sigma_e^2 - 2\gamma_t\beta_t\Delta t\sigma_v^2\sigma_e^2\big) - \Sigma_t\beta_t^2\Delta t\big(\sigma_v^2+\sigma_e^2\big)}. \tag{125}$$

Then, Equations (85)–(87) can be rewritten to express λ_t, μ_t, m_t as functions of $(\Sigma_t, \beta_t, \gamma_t)$ instead of $(\Sigma_{t-1}, \beta_t, \gamma_t)$. Next, we use (92) and (93) to express λ_t, μ_t, m_t as functions of $(\lambda_t, \mu_t, m_t, \alpha_t, \Sigma_t)$. This gives a system of polynomial equations, whose solution λ_t, μ_t, m_t depends only on (α_t, Σ_t). Numerical simulations show that the solution is unique under the second order condition (99), but the authors have not been able to prove theoretically that this is true in all cases. Once the recursive system is computed for all $t = 1, \ldots, T-1$, the only condition left to do is to verify that the value obtained for Σ_0 is the correct one. However, unlike in Kyle (1985), the recursive equation for Σ_t is not linear, and therefore the parameters cannot be simply rescaled. Instead, one must numerically modify the initial choice of Σ_T until the correct value of Σ_0 is reached.

B.2 Discrete Time Benchmark Model

The setup is the same as for the fast model, except that the market maker gets the signal Δz at the same time as the informed trader observes Δv. The sequence of events

[29] Numerically, it should be of the order of Δt.

is as follows. At $t = 0$, the informed trader observes v_0. At each $t = 1, \ldots, T-1$, the informed trader observes $\Delta v_t = v_t - v_{t-1}$; and the market maker observes $\Delta z_t = \Delta v_t + \Delta e_t$, with $\Delta e_t \sim \mathcal{N}(0, \sigma_e^2 \Delta t)$. The market maker quotes the bid price = the ask price = q_t. The informed trader then submits Δx_t, and the liquidity traders submit in aggregate $\Delta u_t \sim \mathcal{N}(0, \sigma_u^2 \Delta t)$. The market maker observes only the aggregate order flow, $\Delta y_t = \Delta x_t + \Delta u_t$, and sets the price at which the trading takes place, p_t. The market maker is competitive, i.e., makes zero profit. This implies

$$p_t = \mathsf{E}(v_t \mid \mathcal{I}_t^p), \qquad \mathcal{I}_t^p = \{\Delta y_1, \ldots, \Delta y_t, \Delta z_1, \ldots, \Delta z_t\}, \tag{126}$$

$$q_t = \mathsf{E}(v_t \mid \mathcal{I}_t^q), \qquad \mathcal{I}_t^q = \{\Delta y_1, \ldots, \Delta y_{t-1}, \Delta z_1, \ldots, \Delta z_t\}. \tag{127}$$

We also denote

$$\Sigma_t = \mathsf{Var}(v_t \mid \mathcal{I}_t^p), \tag{128}$$

$$\Omega_t = \mathsf{Var}(v_t \mid \mathcal{I}_t^q). \tag{129}$$

The next result shows that if the pricing rule is linear, the informed trader's strategy is also linear, and furthermore it only has a level trading component, $\beta_t(v_t - q_t)$.

Theorem 3. *Any linear equilibrium must be of the form*

$$\Delta x_t = \beta_t(v_t - q_t)\Delta t, \tag{130}$$

$$p_t = q_t + \lambda_t \Delta y_t, \tag{131}$$

$$q_t = p_{t-1} + \mu_{t-1} \Delta z_t, \tag{132}$$

for $t = 1, \ldots, T-1$, where by convention $p_0 = 0$, and β_t, γ_t, λ_t, μ_t, Ω_t, and Σ_t are constants that satisfy

$$\lambda_t = \frac{\beta_t \Sigma_t}{\sigma_u^2}, \tag{133}$$

$$\mu_t = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2}, \tag{134}$$

$$\Omega_t = \frac{\Sigma_t \sigma_u^2}{\sigma_u^2 - \beta_t^2 \Sigma_t \Delta t}, \tag{135}$$

$$\Sigma_{t-1} = \Sigma_t + \frac{\beta_t^2 \Sigma_t^2}{\sigma_u^2 - \beta_t^2 \Sigma_t \Delta t} \Delta t - \frac{\sigma_v^2 \sigma_e^2}{\sigma_v^2 + \sigma_e^2} \Delta t. \tag{136}$$

The value function of the informed trader is quadratic for all $t = 1, \ldots, T-1$:

$$\pi_t \;=\; \alpha_{t-1}(v_t - q_t)^2 + \delta_{t-1}. \tag{137}$$

The coefficients of the optimal trading strategy and the value function satisfy

$$\begin{aligned}
\beta_t \Delta t &= \frac{1 - 2\alpha_t\lambda_t}{2\lambda_t(1-\alpha_t\lambda_t)}, &(138)\\
\alpha_{t-1} &= \beta_t\Delta t(1-\lambda_t\beta_t\Delta t) + \alpha_t(1-\lambda_t\beta_t\Delta t)^2, &(139)\\
\delta_{t-1} &= \alpha_t\big(\lambda_t^2\sigma_u^2 + \mu_t^2(\sigma_v^2+\sigma_e^2)\big)\Delta t + \delta_t. &(140)
\end{aligned}$$

The terminal conditions are

$$\alpha_T \;=\; \delta_T \;=\; 0. \tag{141}$$

The second order condition is

$$\lambda_t(1-\alpha_t\lambda_t) \;>\; 0. \tag{142}$$

Given Σ_0, conditions (133)–(142) are necessary and sufficient for the existence of a linear equilibrium.

Proof. First, we show that Equations (133)–(136) are equivalent to the zero profit conditions of the market maker. Second, we show that Equations (138)–(142) are equivalent to the informed trader's strategy (130) being optimal.

Zero Profit of Market Maker: Let us start with with the market maker's update due to the order flow at $t = 1, \ldots, T-1$. Conditional on $\mathcal{I}_t^q$, v_t has a normal distribution, $v_t|\mathcal{I}_t^q \;\sim\; \mathcal{N}(q_t, \Omega_t)$. The aggregate order flow at t is of the form $\Delta y_t = \beta_t(v_t - q_t)\Delta t + \Delta u_t$. Denote by

$$\Phi_t \;=\; \mathsf{Cov}(v_t - q_t, \Delta y_t) \;=\; \beta_t\Omega_t\Delta t. \tag{143}$$

Then, conditional on $\mathcal{I}_t^p = \mathcal{I}_t^q \cup \{\Delta y_t\}$, $v_t \sim \mathcal{N}(p_t, \Sigma_t)$, with

$$\begin{aligned}
p_t &= q_t + \lambda_t \Delta y_t, &(144)\\
\lambda_t &= \Phi_t \, \mathsf{Var}(\Delta y_t)^{-1} \;=\; \frac{\beta_t\Omega_t}{\beta_t^2\Omega_t\Delta t + \sigma_u^2}, &(145)\\
\Sigma_t &= \mathsf{Var}(v_t - q_t) - \Phi_t\,\mathsf{Var}(\Delta y_t)^{-1}\Phi_t' \\
&= \Omega_t - \frac{\beta_t^2\Omega_t^2}{\beta_t^2\Omega_t\Delta t + \sigma_u^2}\,\Delta t \;=\; \frac{\Omega_t\sigma_u^2}{\beta_t^2\Omega_t\Delta t + \sigma_u^2}. &(146)
\end{aligned}$$

To obtain Equation (133) for λ_t, note that the above equations for λ_t and Σ_t imply $\frac{\lambda_t}{\Sigma_t} = \frac{\beta_t}{\sigma_u^2}$. Equation (135) is obtained by solving for Σ_t in Equation (146).

Next, consider the market maker's update at $t = 1, \ldots, T-1$ due to the signal $\Delta z_t = \Delta v_t + \Delta e_t$. From $v_{t-1}|\mathcal{I}_{t-1}^p \sim \mathcal{N}(p_{t-1}, \Sigma_{t-1})$, we have $v_t|\mathcal{I}_{t-1}^p \sim \mathcal{N}(p_{t-1}, \Sigma_{t-1} + \sigma_v^2 \Delta t)$. Denote by

$$\Psi_t \;=\; \mathsf{Cov}(v_t - p_{t-1}, \Delta z_t) \;=\; \sigma_v^2 \Delta t. \tag{147}$$

Then, conditional on $\mathcal{I}_t^q = \mathcal{I}_{t-1}^p \cup \{\Delta z_t\}$, $v_t|\mathcal{I}_t^q \sim \mathcal{N}(q_t, \Omega_t)$, with

$$q_t \;=\; p_{t-1} + \mu_t \Delta z_t, \tag{148}$$

$$\mu_t \;=\; \Psi_t \,\mathsf{Var}(\Delta z_t)^{-1} \;=\; \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2}, \tag{149}$$

$$\begin{aligned} \Omega_t \;&=\; \mathsf{Var}(v_t - p_{t-1}) - \Psi_t \,\mathsf{Var}(\Delta z_t)^{-1} \Psi_t' \\ &=\; \Sigma_{t-1} + \sigma_v^2 \Delta t - \frac{\sigma_v^4}{\sigma_v^2 + \sigma_e^2}\, \Delta t \;=\; \Sigma_{t-1} + \frac{\sigma_v^2 \sigma_e^2}{\sigma_v^2 + \sigma_e^2}\, \Delta t. \end{aligned} \tag{150}$$

which proves Equation (134) for μ_t. Note that Equation (150) gives a formula for Σ_{t-1} as a function of Ω_t, and we already proved (135), which expresses Ω_t as a function of Σ_t. We therefore get Σ_{t-1} as a function of Σ_t, which is (136).

Optimal Strategy of Informed Trader: At each $t = 1, \ldots, T-1$, the informed trader maximizes the expected profit: $\pi_t = \max \sum_{\tau=t}^{T-1} \mathsf{E}\big((v_T - p_\tau)\Delta x_\tau\big)$. We prove by backward induction that the value function is quadratic and of the form given in (137): $\pi_t = \alpha_{t-1}(v_t - q_t)^2 + \delta_{t-1}$. At the last decision point ($t = T-1$) the next value function is zero, i.e., $\alpha_T = \delta_T = 0$, which are the terminal conditions (141). In the induction step, if $t = 1, \ldots, T-1$, we assume that π_{t+1} is of the desired form. The Bellman principle of intertemporal optimization implies

$$\pi_t \;=\; \max_{\Delta x} \; \mathsf{E}\Big((v_t - p_t)\Delta x + \pi_{t+1} \mid \mathcal{I}_t^q, v_t, \Delta v_t\Big). \tag{151}$$

Equations (131) and (132) show that the quote q_t evolves by $q_{t+1} = q_t + m_t \Delta y_t + \mu_t \Delta z_{t+1}$. This implies that the informed trader's choice of Δx affects the trading price and the next quote by

$$
\begin{aligned}
p_t &= q_t + \lambda_t(\Delta x + \Delta u_t), && (152)\\
q_{t+1} &= q_t + \lambda_t(\Delta x + \Delta u_t) + \mu_t \Delta z_{t+1}. && (153)
\end{aligned}
$$

Substituting these into the Bellman equation, we get

$$
\begin{aligned}
\pi_t \;=\; & \max_{\Delta x}\ \mathsf{E}\Big(\Delta x(v_t - q_t - \lambda_t\Delta x - \lambda_t\Delta u_t) \\
& +\ \alpha_t(v_t + \Delta v_{t+1} - q_t - \lambda_t\Delta x - \lambda_t\Delta u_t - \mu_t\Delta z_{t+1})^2 \ +\ \delta_t\Big) && (154)\\
=\; & \max_{\Delta x}\ \Delta x(v_t - q_t - \lambda_t\Delta x) \\
& +\ \alpha_t\Big((v_t - q_t - \lambda_t\Delta x)^2 + (\lambda_t^2\sigma_u^2 + \mu_t^2(\sigma_v^2 + \sigma_e^2))\Delta t\Big) \ +\ \delta_t. && (155)
\end{aligned}
$$

The first order condition with respect to Δx is

$$
\Delta x \;=\; \frac{1 - 2\alpha_t\lambda_t}{2\lambda_t(1 - \alpha_t\lambda_t)}(v_t - q_t), \tag{156}
$$

and the second order condition for a maximum is $\lambda_t(1 - \alpha_t\lambda_t) > 0$, which is (142). Thus, the optimal Δx is indeed of the form $\Delta x_t = \beta_t(v_t - q_t)\Delta t$, where $\beta_t\Delta t$ satisfies Equation (138). We substitute Δx_t in the formula for π_t to obtain

$$
\pi_t \;=\; \Big(\beta_t\Delta t(1 - \lambda_t\beta_t\Delta t) + \alpha_t(1 - \lambda_t\beta_t\Delta t)^2\Big)(v_t - q_t)^2 \ +\ \alpha_t\big(\lambda_t^2\sigma_u^2 + \mu_t^2(\sigma_v^2 + \sigma_e^2)\big)\Delta t \ +\ \delta_t. \tag{157}
$$

This proves that indeed π_t is of the form $\pi_t = \alpha_{t-1}(v_t - q_t)^2 + \delta_{t-1}$, with α_{t-1} and δ_{t-1} as in Equations (139) and (140). $\square$

Equations (133)–(136) and (138)–(140) form a system of equations. As before, it is solved backwards, starting from the boundary conditions (141), and so that $\Sigma_t = \Sigma_0$ at $t = 0$.

B.3 Sampling at a lower frequency than the trading frequency.

In order to make our model more comparable to econometric models, we consider the discrete time version of our continuous time model, as in Appendix B. It works very similarly to the continuous time model, the main difference being that the infinitesimal time interval $\mathrm{d}t$ is replaced by a real number $\Delta t > 0$. We also consider that Δt is small and we approximate the equilibrium variables (β_t, γ_t, λ_t, μ_t, ρ_t) in the discrete time model by their continuous time counterpart. Letting $T = \frac{1}{\Delta t}$ be the number of trading periods, time is indexed by $t = 0, 1, \ldots, T-1$. The informed order flow at time t is equal to

$$\Delta x_t \;=\; \beta_t(v_t - q_t)\Delta t + \gamma_t \Delta v_t, \tag{158}$$

where q_t is the quote just before the order flow arrives, and p_{t+1} is the execution price.

Now, assume that each observation in the data spans $n \geq 1$ trading rounds. In this case, the data are a time-series of length T/n. For $j = 1, \ldots, \frac{T}{n}$, the jth observation corresponds to trading during the n trading rounds starting at time $t = (j-1)n$. The order flow of the informed trader over this period of time is $\Delta x_j(n) \equiv \Delta x_t + \cdots + \Delta x_{t+n-1}$, and, assuming that prices are defined as post-trade quotes, the return is $r_j(n) \equiv p_{t+n} - p_t$. The empirical counterpart of the Informed Participation Rate and the autocorrelation of the informed investor's order flow when data are sampled every n trading rounds are respectively:

$$IPR_j(n) \;=\; \frac{\mathsf{Var}(\Delta x_j(n))}{\mathsf{Var}(\Delta x_j(n)) + \mathsf{Var}(\Delta u_j(n))}. \tag{159}$$

and

$$\mathsf{Corr}_j(n) = \mathsf{Corr}(\Delta x_j(n), \Delta x_{j+1}(n)) \;=\; \frac{\mathsf{Var}(\Delta x_j(n))}{\mathsf{Var}(\Delta x_j(n)) + \mathsf{Var}(\Delta u_j(n))}. \tag{160}$$

To obtain closed-form solutions for these variables, we consider the limit case where the trading frequency is large, holding fixed the time interval $\tau = n\Delta t$ at which data are aggregated. The informed participation rate and the autocorrelation of the informed investor's order flow as a function of τ are therefore: $IPR_j(\tau) = \lim_{\Delta t \to 0} IPR_j(\tau/\Delta t)$ and $\mathsf{Corr}(\tau) = \lim_{\Delta t \to 0} \mathsf{Corr}_j(\tau/\Delta t)$. We obtain the following result.

Corollary 12. *1. In the fast model and in the benchmark case, the empirical informed participation rate $IPR_j(\tau)$ increases with the sampling interval τ. Moreover, the informed investor's participation rate is higher in the fast model ($IPR_j^F(\tau) > IPR_j^B(\tau)$).*

2. In the fast model and in the benchmark case, the informed order flow autocorrelation, $\mathsf{Corr}(\tau)$, increases with the sampling interval τ. Moreover, this autocorrela-

tion is loewer in the fast model ($\mathsf{Corr}_j^F(\tau) > \mathsf{Corr}_j^B(\tau)$).

Finally, consider the empirical counterpart of our measure of anticipatory trading:

$$AT_j(n) = \mathsf{Corr}(\Delta x_j(n), r_{j+1}(n)). \tag{161}$$

The next result shows that that sampling data at a sufficiently high frequency (n) is important for detecting anticipatory trading.

Corollary 13. *The empirical measure of anticipatory trading $AT_j(n)$ decreases with n and converges to zero when $n \to +\infty$.*

The aggregated order flow spans n trading periods. Moreover, each trade anticipates news that is incorporated in the quotes in the next trading round. Therefore, only the last trade of the aggregated order flow $\Delta x_j(n)$ is correlated with the next aggregated return $r_{j+1}(n)$. As a result, when n increases, the correlation between $\Delta x_j(n)$ and $r_{j+1}(n)$ decreases. When n becomes too large, the correlation becomes almost zero.

B.4 Proof of Corollary 12

Part 1. We consider the limit $\Delta t \to 0$ and $n \to +\infty$ such that $n\Delta t = \tau$ is fixed. In this case, we have $Var(\Delta x_j) = Var(x_{t+\tau} - x_t)$, where $t = (j-1)\tau$. Then, we can write

$$\begin{aligned} Var(x_j) &= Var(\int_{s=t}^{t+\tau} \beta_s(v_s - p_s)ds + \gamma_s dv_s) \\ &= \int_{s=t}^{t+\tau} \gamma_s^2 Var(dv_s) \\ &\quad +2\int_{s=t}^{t+\tau}\int_{u=s}^{t+\tau} \beta_s\beta_u Cov\left(v_s - p_s, v_u - p_u\right) ds\, du \\ &\quad +2\int_{s=t}^{t+\tau}\int_{u=s}^{t+\tau} \gamma_s\beta_u Cov\left(dv_s, v_u - p_u\right) du. \end{aligned}$$

It then follows from Lemma 2 that

$$Var(x_j) = \gamma^2\sigma_v^2\tau + \beta_t(\beta_0\Sigma_0 + \gamma(1 - m\gamma - \mu)\sigma_v^2)\tau^2 + o(\tau^2)$$

when τ is small and where $m \equiv \lambda - \mu\rho$. Using that $Var(u_j) = \sigma_u^2\tau$, we obtain

$$IPR_j = \frac{\gamma^2\sigma_v^2}{\gamma^2\sigma_v^2 + \sigma_u^2} + \frac{\sigma_u^2}{(\gamma^2\sigma_v^2 + \sigma_u^2)^2}\beta_t(\beta_0\Sigma_0 + \gamma(1 - m\gamma - \mu)\sigma_v^2)\tau + o(\tau).$$

Part 2. Moreover, we have

$$Cov(\Delta x_j, \Delta x_{j+1}) = \beta_t(\beta_0 \Sigma_0 + \gamma(1 - m\gamma - \mu)\sigma_v^2)\tau^2 + o(\tau^2),$$
$$Var(\Delta x_j) = Var(\Delta x_{j+1}) = \gamma^2 \sigma_v^2 \tau + \beta_t(\beta_0 \Sigma_0 + \gamma(1 - m\gamma - \mu)\sigma_v^2)\tau^2 + o(\tau^2).$$

Therefore

$$Corr(\Delta x_j, \Delta x_{j+1}) = \frac{\beta_t^2 \Sigma_t + \beta_t \gamma_t (1 - m\gamma - \mu)\sigma_v^2}{\gamma^2 \sigma_v^2} \tau + o(\tau).$$

B.5 Proof of Corollary 13

In the limit $\Delta t \to 0$, we have

$$Cov(\Delta x_j(n), r_{j+1}(n)) = \mu\gamma(1 - \rho\gamma)\sigma_v^2 \Delta t,$$
$$Var(\Delta x_j) = n\gamma^2 \sigma_v^2 \Delta t,$$
$$Var(r_{j+1}) = (\sigma_v^2 + \Sigma_0) n \Delta t.$$

Therefore

$$Corr(\Delta x_j, r_{j+1}) = \frac{\mu(1 - \rho\gamma)\sigma_v}{n\sqrt{\sigma_v^2 + \Sigma_0}}$$

is decreasing in n and goes to 0 when n goes to infinity.

References

[1] Back, Kerry, Henry Cao, and Gregory Willard (2000): "Imperfect Competition among Informed Traders," *Journal of Finance*, 55, 2117–2155.

[2] Back, Kerry, and Hal Pedersen (1998): "Long-Lived Information and Intraday Patterns," *Journal of Financial Markets*, 1, 385–402.

[3] Biais, Bruno, Thierry Foucault, and Sophie Moinas (2011): "Equilibrium High Frequency Trading," Working Paper.

bibitemBFM Cvitanic Jaksa and Andrei Kirilenko (2011): "High Frequency Traders and Asset Prices," Working Paper.

[4] Brogaard, Jonathan (2011): "The Activity of High Frequency Traders," Working Paper.

[5] Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (2011): "High Frequency Trading and Price Discovery"," Working Paper.

[6] Busse Jeffrey and Clifton T. Green (2002): "Market efficiency in real time," *Journal of Financial Economics*, 65, 415–437.

[7] Cartea, Alvara and Jose Penalva (2012): "Where is the value in high frequency trading? forthcoming *Quarterly Journal of Finance.*

[8] Chaboud, Alain, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega (2009): "Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market," Working Paper, Board of Governors of the Federal Reserve System.

[9] Chau, Minh, and Dimitri Vayanos (2008): "Strong-Form Efficiency with Monopolistic Insiders," *Review of Financial Studies*, 21, 2275–2306.

[10] Gai, Jiading, Chen Yao, and Mao Ye (2012): "The Externalities of High-Frequency Trading," Working paper.

[11] Green, T. Clifton (2011): "Economic News and the Impact of Trading on Bond Prices," *Journal of Finance*, 59, 1201–1234.

[12] GroB-KluBmann, Alex and Nikolaus Hautsch (2011): "When Machines Read the News: Using Automated Text Analytics to Quantify High-Frequency News-Implied Market Reactions," *Journal of Empirical Finance*, 18, 321–340.

[13] Hasbrouck, Joel (1991a): "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46, 179–207.

[14] Hasbrouck, Joel (1991b): "The Summary Informativeness of Stock Trades: An Econometric Analysis," *Review of Financial Studies*, 4, 571–595.

[15] Hasbrouck, Joel, and Gideon Saar (2011): "Low-Latency Trading," Working Paper.

[16] Hendershott, Terrence, Charles Jones, and Albert Menkveld (2011): "Does Algorithmic Trading Improve Liquidity?," *Journal of Finance*, 66, 1–33.

[17] Hendershott, Terrence, and Pamela Moulton (2011): "Automation, Speed, and Stock Market Quality: The NYSE's Hybrid," *Journal of Financial Markets*, 14, 568–604.

[18] Hendershott, Terrence, and Ryan Riordan (2010): "Algorithmic Trading and Information," Working Paper.

[19] Hendershott, Terrence, and Ryan Riordan (2011): "High Frequency Trading and Price Discovery," Working Paper.

[20] Hirschey Nicholas (2011): "Do High Frequency Traders Anticipate Buying and Selling Pressure," Working Paper.

[21] Hoffman Peter (2011): "A Dynamic limit Order Market with Fast and Slow Traders," Working Paper, ECB.

[22] Jovanovic, Boyan, and Albert Menkveld (2011): "Middlemen in Limit-Order Markets," Working Paper.

[23] Karatzas, Ioannis, and Steven Shreve (1991): *Brownian Motion and Stochastic Calculus*, Springer Verlag.

[24] Kim, Oliver and Robert E. Verrecchia (1994): "Market Liquidity and Volume around Announcements," *Journal of Accounting and Economics*, 17, 41–67.

[25] Kim, Oliver and Robert E. Verrecchia (1991): "Trading Volume and Price Reactions to Public Announcements," *Journal of Accounting Research*, 29, 302–320.

[26] Kirilenko, Andrei, Albert Kyle, Mehrdad Samadi, and Tugkan Tuzun (2011): "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market," Working Paper.

[27] Kyle, Albert (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.

[28] Menkveld, Albert (2012): "High Frequency Trading and the New-Market Makers," Working Paper.

[29] Patell, J., Wolfson, M. (1984): "The intraday speed of adjustment of stock prices to earnings and dividend announcements," *Journal of Financial Economics*, 13, 223252.

[30] Pagnotta, Emiliano, and Thomas Philippon (2011): "Competing on Speed," Working Paper.

[31] SEC (2010): "Concept Release on Equity Market Structure," Release No. 34-61358; File No. S7-02-10.

[32] Tetlock, Paul C. (2008): "Does Public Financial News resolve Asymmetric Information," *Review of Financial Studies*, 23, 3520–2557.

[33] Vega Clara (2006): "Stock price reactions to public and private information," *Journal of Financial Economics*, 82, 103–133.

[34] Zhang, Sarah (2012): "Need for Speed: An Empirical Analysis of Hard and Soft Information in a High Frequency World," Working Paper.