News Trading and Speed*

Thierry Foucault† Johan Hombert‡ Ioanid Roşu§

February 15, 2013

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 news traders on liquidity, volatility, price discovery, short-run price dynamics 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, and Mao Ye for their suggestions. We are also grateful to finance seminar participants at Paris Dauphine, Copenhagen Business School, Univ. Carlos III in Madrid, ESSEC, Lugano, INSEAD and conference participants at the NYU Stern Microstructure Meeting, CNMV International Conference on Securities Markets, Central Bank Microstructure Workshop, Market Microstructure Many Viewpoints Conference in Paris, and Paris Hedge Fund Conference 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 Verrecchia (1991, 1994), Busse and Green (2001), Vega (2006), or Tetlock (2010)). A new breed of market participants, *high frequency traders on news* (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 (public information) 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?

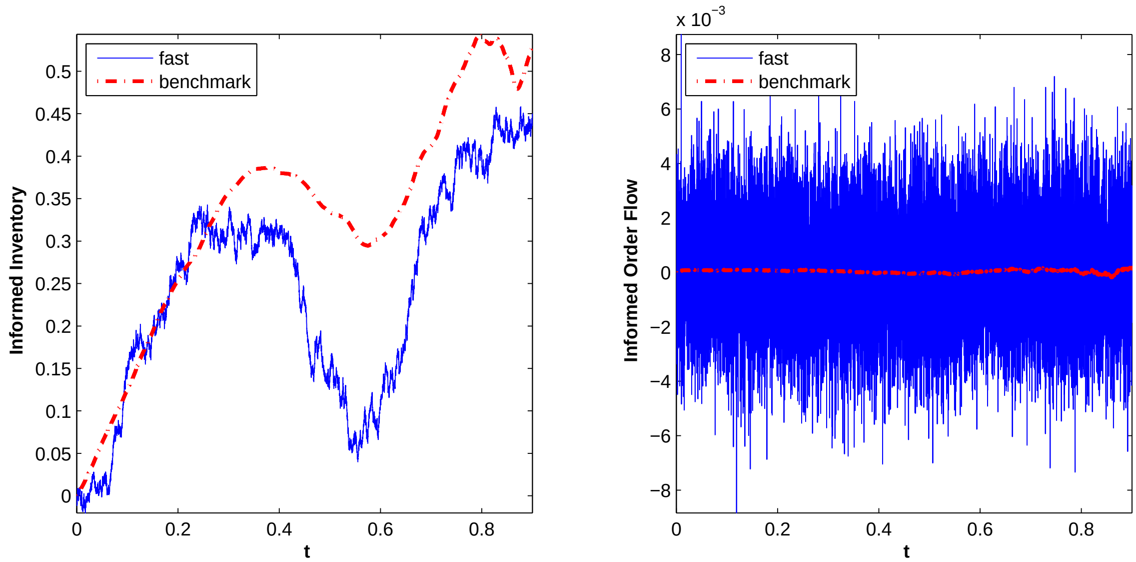
To address these questions, we consider a model in which an informed investor continuously receives news about the payoff of a risky security. He has both a greater information processing capacity *and* a higher speed of reaction to news than market makers. The information processing advantage enables 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 versus when he does not, holding the precision of his private information constant. In particular, a small speed advantage 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 market

[1]News exploited by these traders are very diverse and should not be construed only as scheduled corporate or macroeconomic announcements. In fact, they also 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 limit order book updates, and market-wide returns, and macroeconomic announcements. Jovanovic and Menkveld (2011) or Zhang (2012) show that high frequency traders also use index futures price information as a source information to establish positions in underlying stocks.

[2]This is also the case for models that specifically analyze informed trading around news releases. For instance, Kim and Verrecchia (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.

Figure 1: Informed participation rate at 1-second frequency. The figure plots the evolution of the informed investor's position (left graph) and the change in this position—the informed investor's trade—(right graph), when the informed investor has a speed advantage (plain line) and when he has no speed advantage (dot-dashed line), using the characterization of the optimal trading strategy for the investor in each case. The parameters used for the simulation are $\sigma_u = \sigma_v = \sigma_e = \Sigma_0 = 1$ (see Theorem 1). The liquidation date $t = 1$ corresponds to 1 month.



makers.

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 market makers. Second, by receiving news a split second before market makers, the informed investor can forecast market makers' 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 market makers' forecast error (the difference between the informed investor's and market makers' 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 point for one particular realization of news in our model.

This finding has several important and new implications. For instance, the informed

[3]In contrast, the drift of the investor's position is proportional to market makers' 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).

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 (i.e., over a very short time interval) has a positive correlation with subsequent returns, because the informed investor's trades are mainly driven by news arrivals, at high frequency. These features fit well with some stylized 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 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, but no speed advantage, cannot explain these facts.

Moreover, the effect of the precision of public information (that is, the news received by market makers) differs from that obtained in other models of trading around news, such as Kim and Verrecchia (1994). In these models, more precise public information is associated with greater market liquidity (lower price impact) but lower trading volume; see Kim and Verrecchia (1994) for instance. In contrast, in our model, it is associated with both an increase in liquidity (as market makers 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 market makers, 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 market makers 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]

[4]For instance, Kirilenko, Kyle, Samadi, and Tuzun (2011, p. 21) note 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."*

[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

We also 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.

Not surprisingly, 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.[6] Less obviously, this speed advantage also affects the nature of price discovery: price changes over short horizon are more correlated with innovations in the asset value (as found empirically in Brogaard, Hendershott, and Riordan (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. Similarly, high frequency news trading alters the relative influences of trades and news arrivals on short run volatility. Trades move market makers' price more when the investor has a speed advantage because they are more informative about imminent news. But precisely for this reason, market makers' 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.

A prerequisite for high frequency traders to contribute to price discovery is that their trades contain information, that is, permanently affect prices. Two empirical methods

is more precise, both the informed investor's share of trading volume and liquidity improve. 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.

have been used to test this: (i) the Vector Autoregressive Model (VAR) approach (e.g., Zhang (2012) or Hirschey (2012)) and (ii) the state space model approach (e.g., Brogaard, Hendershott, and Riordan (2012)).[7] However, the lack of dynamic structural models of high frequency trading hinders the interpretation of these models. To fill this gap, we also provide the VAR and state space model representations for the midquotes and order flow implied by our model and we express the coefficients of these representations in terms of the parameters of our model. Interestingly, the state space representation implied by our model has the same structure as the one estimated by Brogaard, Hendershott, and Riordan (2012), and our predictions for the signs of the coefficients of this representation are in line with their estimates.

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, in which the informed investor must 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, and some HFTs mainly use aggressive orders (see Baron, Brogaard, and Kirilenko (2012)). 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]

Our model encompasses the special case in which the market makers and the informed investor have exactly the same forecast of the asset value after news are observed by market makers because both the investor and market makers have the same information processing capacity. In this case, speed is the only way in which the investor can trade on public information. That is, the drift component of the investors' position

[7] These techniques have been used more generally to measure the informational content of trades, not just trades from HFTs (see Hasbrouck (1991a) and Menkveld, Koopman, and Lucas (2007)).

[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, the proxy of Hasbrouck and Saar (2012) does not specifically capture the high frequency trades triggered by the arrival of news. Thus, it may be a noisy proxy for the trades of HFTNs.

becomes zero but otherwise results are unchanged. This particular case of the model might be appropriate to describe trading on particular types of news, such as scheduled macroeconomic announcements, for which the interpretation of news is relatively easy.[9] In general, however, high frequency news traders might be able to react faster to news *and* to process news more efficiently, as allowed in our model. Ultimately, which specification of the model better matches the behavior of HFTNs is an empirical question. We provide very precise predictions on the dynamics of HFTNs' trades precisely for this reason: they offer a way for empiricists with data on HFTNs to reject the model (or one of its specifications) and thereby identify how it should be modified to obtain a more realistic description of their behavior.

Our paper is related to the growing theoretical literature on high frequency trading.[10] Our analysis is most related to Biais, Foucault, and 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 helps to understand dynamic relationships between returns and the order flow of high frequency traders. This is important since these relationships can be used to make inferences about the effects and role of high frequency traders (as, for instance, in Brogaard, Hendershott, and Riordan (2012)).

Technically, our model is related to Back and Pedersen (1998) (BP(1998)), Chau and Vayanos (2008) (CV(2008)), and Martinez and Roşu (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 market makers. As in CV(2008), market makers receive news continuously as well, but not as precisely as the investor.[11] In contrast to both models, we assume that the informed investor observes 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.

[9] Scholtus et al. (2012) show that speed is critical for investors trading on macroeconomic announcements. For a sample of 707 announcements, they find that a delay of 300 milliseconds in reacting to the announcement reduces the returns on a strategy exploiting the informational content of the announcement by about 0.44 bps (3.08% on average across all announcements).

[10] 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 Hoffmann (2012).

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

MR(2012) obtain a similar finding for a different reason. In their model, market makers receive no news. In this particular case, the news trading component would disappear in our model. This is not the case in MR(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 the model when the informed investor has no speed advantage (the *benchmark* model) and when he has one (the *fast* model). Section 3 describes the resulting equilibrium price process and trading strategies in each case. Section 4 analyzes the empirical implications of the model and the effects of HFTNs on measures of market quality. Section 5 concludes. All proofs are in Appendices A and B. The model is set in continuous time. In the Internet Appendix C we 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.

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 $\Sigma_0 > 0$, and B_t^v is a Brownian motion.[12] The informed trader observes v_0 at time 0, and observes $\mathrm{d}v_t$ during each time

[12]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, Hendershott, and Riordan (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.

interval $(t, t+\mathrm{d}t]$, $t \in (0,1)$. 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.[13]

The aggregate position of the noise traders at t is denoted by u_t. It follows an 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. During $(t, t+\mathrm{d}t]$, she receives a noisy signal of the innovation in asset value:

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

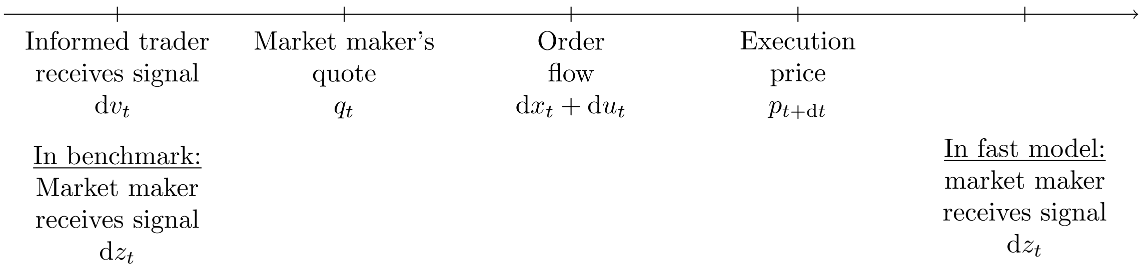
where B_t^e is a Brownian motion independent from all the others. We refer to $\mathrm{d}z_t$ as the flow of news received by the market maker at date t. Furthermore, the market maker learns information from the aggregate order flow:

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

because $\mathrm{d}x_t$ will reflect the information possessed by the informed trader (see below). We denote by q_t the market maker's expectation of the asset liquidation value just before she observes the aggregate order flow $\mathrm{d}y_t$. As the market maker is competitive and risk-neutral, she executes the order flow at a price equal to her expectation of the asset value just after she receives the order flow (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

[13] 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.

Figure 2: Timing of events during $(t, t + \mathrm{d}t]$ in the benchmark and the fast model



bid-ask midpoint just before the transaction at $t + \mathrm{d}t$.[14]

As explained in the introduction, our goal is to analyze the effects of granting to the informed investor the possibility to react to news faster than market makers, holding the accuracy of his signals (that is, $\Sigma_0 > 0$ and $\sigma_e > 0$) constant. To this end, we consider two different models: the *benchmark model* and the *fast model.* They differ in the timing with which the informed investor and the market maker receive news. The sequence of information arrival, quotes and trades in each model is summarized in Figure 2.

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 $\mathcal{I}_t \equiv \{z_\tau\}_{\tau \leq t} \cup \{y_\tau\}_{\tau \leq t}$, which comprises the order flow and the market maker's signals until time t, and the news just received in the interval $(t, t + dt]$. 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 over $(t, t + dt]$.

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 $\mathrm{d}u_t$

[14] 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.

and the market maker executes the aggregate order flow 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 her 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 midquote $q_{t+\mathrm{d}t}$ 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,

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

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 market maker 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 (1991a)). By contrast, in the fast model, the informed trader observes news a split second before the market maker. Thus, he also has a speed advantage relative to the market maker. 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, holding the precision of information constant.

In both the benchmark model and the fast model, when $\Sigma_0 > 0$, the informed investor starts with long-lived private information. Moreover, when $\sigma_e > 0$, the informed investor receives additional long-lived private signals because the information he extracts from news is more precise than that of the market makers. As in Kim and Verrecchia (1994), this gain in accuracy could stem from the fact that the informed investor is better able to process news than market makers. Our results however hold even when $\sigma_e = 0$ and Σ_0 is very close to zero (for technical reasons, Σ_0 must be strictly positive; see the discussion at the the end of Section 3). In this case, the fast model can be seen as the case in which the informed investor has a faster access to only very *short-lived information* ($\mathrm{d}v_t$), that is, information that will almost immediately be observed perfectly by market makers (as in the discrete time model of Admati and Pfleiderer (1988)).

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 market makers' pricing policy and (b) market makers' 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 that uses the framework of Kyle (1985), 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 ($\mathrm{d}x_t$) 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 market maker's quote revision is linear in the public information she receives, while the price impact is linear in the order flow. That is,

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

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,

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

where $\rho_t^F \mathrm{d}y_t$ is the market maker's expectation of the public information arriving over $(t, t+\mathrm{d}t]$ 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, $\rho_t^F > 0$ 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 market maker 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

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

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 equilibria which are linear in past realizations of p and v. However, we show in the Internet Appendix C 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 date t has two components. The first component, $\beta_t(v_t - p_t)\mathrm{d}t$, is proportional to the market maker's forecast error, i.e., the difference between the forecast of the asset value by the informed investor and the forecast of this value by the market maker prior to the trade over $(t, t + \mathrm{d}t]$. Intuitively, the informed investor buys when the market maker underestimates the fundamental value and sells otherwise. This component is standard in models of trading with asymmetric information such as Kyle (1985), Back and Pedersen (1998), Back, Cao, and Willard (2000), etc. In what follows, we refer to this component as the *forecast error* 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 detail 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 (the coefficients μ_t^k, λ_t^k, ρ_t^F, β_t^k, γ_t^k) in both the benchmark and the fast cases.

Theorem 1. *In the benchmark model there is a unique linear equilibrium, of the form*

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

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

with coefficients given by

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

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

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

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

In the fast model there is a unique linear equilibrium, of the form:[15]

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

$$\mathrm{d}q_t = \lambda^F \mathrm{d}y_t + \mu^F(\mathrm{d}z_t - \rho^F \mathrm{d}y_t), \tag{18}$$

with coefficients given by

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

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

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

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

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

[15] Note that the forecast error component in (17) has q_t instead of p_t. This is the same formula, since (9) implies $(p_t - q_t)\mathrm{d}t = 0$. The reason we use q_t as a state variable instead of p_t is that p_t is not an Itô process in the fast model. Indeed, $\mathrm{d}p_t = p_{t+\mathrm{d}t} - p_t = (p_{t+\mathrm{d}t} - q_t) + (q_t - p_t)$, and, according to (9) $p_{t+\mathrm{d}t} - q_t$ and $q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t}$ are Itô processes, but $q_t - p_t$ is not.

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

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

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

The news trading component of the informed investor is nonzero 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 market maker will adjust her 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 ($\mathrm{d}v_t > 0$) or sell just before a decrease in prices due to bad news ($\mathrm{d}v_t < 0$). As a result, $\gamma^F > 0$ if $\sigma_e < +\infty$. In contrast, in the benchmark case the market maker incorporates news in her 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, $\gamma^B = 0$, even if market makers' information is less precise (i.e., $\sigma_e > 0$).

Whether he is fast or not, the informed investor can form a forecast (v_t) 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 forecast error component of the strategy is present whether the informed investor has a speed advantage or not ($\beta_t^k > 0$ for $k \in \{B, F\}$).

Proposition 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 forecast error trading 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 proposition 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.

Proposition 2. *In the benchmark equilibrium and the fast equilibrium, β_t^B and β_t^F are increasing in σ_v, σ_u, σ_e, and decreasing in Σ_0. Moreover, in the fast equilibrium, γ^F is increasing in σ_u, and decreasing 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 on the market maker's forecast error.[16]

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 converges to zero when σ_e approaches infinity. Thus, there is no news trading if the market makers do not receive news. Moreover, as the informed investor trades less aggressively on news, his trades become more sensitive to the market maker's forecast error (β_t^F increases) because of the substitution effect discussed after Proposition 1.[17]

When σ_e approaches $+\infty$, 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 therefore is identical to that obtained in Back and Pedersen (1998). If furthermore $\sigma_v = 0$, the informed investor receives no news beyond the initial v_0 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 market maker, his trading strategy will not feature a news trading component if the market maker does not receive news (γ^F

[16] The dependence of γ^F on σ_v is ambiguous, since when σ_v increases, γ^F first increases and then decreases, reflecting the fact that the price impact coefficient λ^F also increases with σ_v, which tempers the aggressiveness of the informed trader.

[17] Indeed, we show in Proposition 11 that the informed investor trades so that the total price informativeness is the same in both models.

converges zero when σ_e approaches $+\infty$). Another polar case is that in which $\sigma_e = 0$. In this case, the information contained in news is very short-lived for the informed investor. As implied by Proposition 2, the informed investor then trades very aggressively on news (γ^F is maximal when $\sigma_e = 0$).

Theorem 1 holds for all values of the parameters, except $\Sigma_0 = 0$. As Σ_0 approaches zero, market makers' forecast errors become very small, at least at date $t = 0$. However, the forecast error component of the investor's position remains finite, which implies that β_0^k approaches infinity when Σ_0 approaches zero. Thus, when $\Sigma_0 = 0$, there is no equilibrium where β_t^k increase over time as in Theorem 1. However, the theorem and all our results (which are all implications of the theorem) remain unchanged even when Σ_0 is very close to zero. When this is the case and in addition $\sigma_e = 0$, the informed investor has no long-lived information advantage.[18] However, as discussed previously, the speed advantage grants him a very short-lived information advantage that is reflected in the news trading component of his strategy.[19]

As soon as $\sigma_e > 0$, the informed investor has both a speed advantage and long-lived information. Interestingly, the two components of the strategy can dictate trades in opposite directions. For instance, the forecast error 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 $\mathrm{d}v_t < 0$). The net direction of the informed investor's trade is determined by the sum of these two desired trades. In reality, a proprietary trading firm might delegate the implementation of the two components of his trading strategy to two different agents (trading desks). In this case, one may see trades in opposite directions for these agents. Yet, they are part of the optimal trading strategy for the firm hiring them. If this is the case, then one needs data at the firm level (rather than trade level data) to make correct inferences about the behavior of high frequency traders.

3.2 The Trades of HFTNs

We now characterize further the properties of the informed investors' trades in our model. Our goal is twofold. First, we show that these properties better coincide with

[18] To see this formally, one can check that $\mathsf{Cov}_t(\mathrm{d}x_t^k, v_t - p_t)/\mathrm{d}t = \beta_t^k \Sigma_t = \beta_0^k \Sigma_0$ where $\Sigma_t = \mathsf{E}\big((v_t - p_t)^2\big)$. Then, using the expression for β_0^k, when $\sigma_e = 0$, it is readily obtained that $\mathsf{Cov}_t(\mathrm{d}x_t^k, v_t - p_t)/\mathrm{d}t$ converges to zero when Σ_0 approaches zero.

[19] Using the expression for γ^F in Theorem 1, it is immediate that when $\sigma_e = 0$ and Σ_0 approaches zero then γ^F remains strictly positive and converges to $\gamma^F = \frac{\sigma_u}{\sigma_v}$.

stylized facts about high frequency traders when the informed investor has a speed advantage than when he has not. Second, empirical analyses of high frequency traders have focused on short term dynamics of high frequency traders' order flows and their interactions with short term returns. However, so far, there is a lack of predictions from theory for specific high frequency trading strategies. The findings in this section provide such predictions for high frequency trading on news.

The position of the informed investor, x_t, is a stochastic process. The drift of this process is equal to the forecast error component, while the volatility component of this process is the news trading component. As the latter is zero in the benchmark case, the informed investor's trades at the high frequency ($\mathrm{d}x_t$) 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 $(\mathrm{d}t)^{1/2}$). In contrast, in the fast model, 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, the 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_t) be the contribution of the informed trader to total trading volume over an infinitesimal time interval $(t, t + \mathrm{d}t]$,

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

Proposition 3. *The informed participation rate is zero when the informed trader has no speed advantage, while it is strictly positive when he has a speed advantage:*

$$IPR^B \;=\; 0, \qquad IPR^F \;=\; \frac{g}{1+g} \;>\; 0, \tag{26}$$

where $g \in (0, 1)$ 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 (1998). As a result, the forecast error 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 the innovations in asset value 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 infinitesimal time intervals this autocorrelation is zero.

Proposition 4. *Over short time intervals, the autocorrelation of the informed order flow is strictly positive when the informed investor has no speed advantage, and zero when the he has a speed advantage: for* $\tau \in (0, 1-t)$*,*

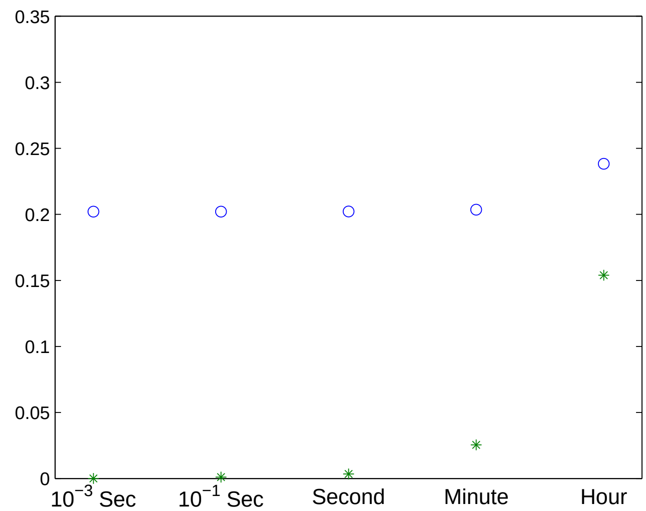
$$\begin{aligned} \mathsf{Corr}(\mathrm{d}x_t^B, \mathrm{d}x_{t+\tau}^B) &= \left(\frac{1-t-\tau}{1-t}\right)^{\lambda^B \beta_0^B - \frac{1}{2}} > 0, \\ \mathsf{Corr}(\mathrm{d}x_t^F, \mathrm{d}x_{t+\tau}^F) &= 0. \end{aligned} \tag{27}$$

Proposition 3 and 4 hold when the order flow of the informed investor is measured over an infinitesimal time interval. But 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.[20] 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.[21] 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 intervals, the informed investor's participation rate in the benchmark as well as the autocorrelation of his trades in the fast model both increase above zero. 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 sampling frequencies (e.g. daily), the participation rate of the informed investor remains low when he has no speed advantage, as shown on Figure 3. 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.

[20]For instance, Zhang (2012) aggregates the trades by HFTs in her sample over intervals of 10 seconds, although trades in her sample occur at a higher frequency.

[21]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.

Figure 3: Informed participation rate at various sampling frequencies. The figure plots the fraction of the trading volume due to the informed trader when data are sampled over time intervals of various lengths (10^{-3} seconds, 10^{-1} seconds, 1 second, 1 minute, 1 hour) in (a) the benchmark model, marked with "$*$"; and (b) the fast model, marked with "$\circ$". The parameters used for the simulation are $\sigma_u = \sigma_v = \sigma_e = \Sigma_0 = 1$ (see Theorem 1). The liquidation date $t = 1$ corresponds to 1 month.



Using U.S. stock trading data aggregated across twenty-six HFTs, Brogaard (2011) finds 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 in the fast model (see Proposition B.1 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 HFTNs face inventory constraints due to risk management concerns. While this feature is absent from our model, such constraints would naturally lead to mean reversion in the informed investor's trades. Alternatively, these empirical studies may describe the behavior of a different category of high frequency traders we do not model, namely the high frequency market makers.[22] Menkveld (2011) shows that the high frequency trader in his dataset behaves very much as a market maker rather than as an informed investor.

[22]Baron, Brogaard, and Kirilenko (2012) find that HFTs that mainly use limit orders tend to have lower overnight inventories than HFTs mainly using aggressive orders. This difference suggests that not all HFTs manage their inventories in the same way.

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*) be the correlation between the informed order flow at a given date and the next instant return, that is:

$$AT_t \;=\; \mathsf{Corr}(\mathrm{d}x_t, q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t}), \tag{28}$$

where we recall that $p_{t+\mathrm{d}t}$ is the price at which the trade $\mathrm{d}x_t$ is executed, and $q_{t+\mathrm{d}t}$ is the next quote posted by the market maker after she receives additional news (see Figure 2).

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

$$\begin{aligned} AT^B &= 0, \\ AT^F &= \frac{1}{\sqrt{(1+g)(1+\frac{\sigma_e^2}{\sigma_v^2})}} > 0, \end{aligned} \tag{29}$$

where $g \in (0,1)$ is as in Theorem 1.

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.[23]

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 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 when the frequency at which data are sampled decreases relative to the frequency at which the investor trades, and approaches

[23] 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 or alongside other investors, for instance institutional investors (see Hirschey (2011)). This form of anticipatory trading is not captured by our model.

zero when the ratio of sampling frequency to trading frequency converges to zero (as in the continuous time model). Thus, the choice of a sampling frequency to study high frequency news trading is not innocuous and 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 would be more difficult to detect the presence of anticipatory trading by the informed investor.

4 Empirical Implications

4.1 News Informativeness, Volume and Liquidity

Empirical findings suggest that the activity of high frequency traders vary across stocks (e.g., Brogaard, Hendershott, and Riordan (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 traders on news, 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 Verrecchia (1994), we measure the precision of public information by σ_e, since a smaller σ_e means that the news received by the market maker provide a more precise signal about innovations in the asset value.[24] Moreover we measure the volume of trading by $\mathsf{Var}(\mathrm{d}y_t)$, a measure of the average absolute order imbalance in each transaction.

Proposition 6. *In the fast model, an increase in the precision of public news, i.e., a decrease in σ_e, results in (i) higher participation of the informed investor (IPR^F), (ii) higher trading volume ($\sqrt{\mathsf{Var}(\mathrm{d}y)}$), and (iii) higher liquidity (lower λ^F).*

When public information is more precise, the informed investor trades more aggressively on the news he receives as shown by Proposition 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 exploit his foreknowledge of news when he receives news faster than the market maker. As a result, he trades more over short-time interval so that his participation rate and trading volume increase.

[24]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.

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 market maker. On the other hand, it helps the market maker to better forecast the asset liquidation value, which reduces his exposure to adverse selection. As shown by Proposition 6, 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 Verrecchia (1994), who find that an increase in the precision of public information leads to *lower* trading volume, as informed investors trade less. Furthermore, our results 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 lead to a positive association between liquidity and the activity of high frequency news traders, but this does not imply that high frequency news trading causes the market to be more liquid. Instead, we show in Proposition 10 that the opposite is true.

To test the implications of Proposition 6, one needs a proxy for the precision of public news received by market makers. For this, one can consider various "news sentiment scores" that are provided by data vendors such as Reuters, Bloomberg, Dow Jones (see for instance Gross-Klussmann 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 $\mathrm{d}z_t$) and a relevance score to news. Thus, as proxy for σ_e, we suggest the average relevance score of news about a firm (or a portfolio of firms). Indeed, 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 market makers and can be seen as private information in the traditional sense. In contrast, the news that the informed investor receives are partially revealed to market makers and his corresponding information advantage is very short lived if σ_e is small. When σ_v increases, news become relatively more informative than private information for the informed investor since news account for a greater fraction of the total variance of the liquidation value for the asset, i.e., $\frac{\sigma_v^2}{\sigma_v^2+\Sigma_0}$ increases in σ_v.[25] At the same time, for a

[25] The liquidation value v_1 decomposes as $v_1 = v_0 + (v_1 - v_0)$, thus $\mathsf{Var}(v_1) = \Sigma_0 + \sigma_v^2$.

fixed value of σ_e, the news received by the market maker are more informative.[26] Thus, to analyze the effect of increasing the informational content of news for the informed investor *ceteris paribus*, the next proposition considers the effect of a change in σ_v, while holding constant the ratio $\frac{\sigma_e}{\sigma_v}$.

Proposition 7. *In the fast model, an increase in the informational content of news for the informed investor, i.e., an increase in σ_v holding fixed σ_e/σ_v, results in (i) higher participation rate for the informed investor (IPR^F), (ii) higher trading volume, and (iii) lower liquidity.*

A greater value of σ_v increases the informed trader's profit from the 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. To test our results, we suggest as proxy for σ_v the number of times the news sentiment score for a given firm exceeds a certain threshold over a fixed period of time (say, over a 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.

Last, an increase in σ_v is associated with a higher price volatility since, according to Proposition 12, $\frac{1}{\mathrm{d}t}\mathsf{Var}(\mathrm{d}p_t) = \sigma_v^2 + \Sigma_0$. Thus, an increase in the informational content of news received by HFTNs leads to a positive association between the volatility of short-run return and the activity of high frequency traders on news, although this does not imply that HFTNs have a causal effect on volatility. In fact, as shown in Proposition 12, the speed advantage of the informed investor has no effect on volatility in our model.[27]

4.2 Measuring the informational content of HFTNs' trades

As explained in the introduction, it is of interest to test whether the trades of high frequency traders contain information. The market microstructure literature suggests two approaches for this purpose: (i) the Vector Autoregressive Model (VAR) approach, developed by Hasbrouck (1991a), and (ii) the state space model approach, initially developed by Menkveld, Koopman, and Lucas (2007). Naturally, these methods have

[26] The fraction of the uncertainty about $\mathrm{d}v_t$ that is resolved upon observing $\mathrm{d}z_t$ is equal to $1 - \frac{\mathsf{Var}(\mathrm{d}v_t|\mathrm{d}z_t)}{\mathsf{Var}(\mathrm{d}v_t)} = \frac{\sigma_v^2}{\sigma_v^2+\sigma_e^2}$, which increases with σ_v.

[27] 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.

been used to measure the informativeness of HFTs trades; see, for instance, Zhang (2012) for an implementation of the VAR approach and Brogaard, Hendershott, and Riordan (2012) (BHR(2012)) for the state space model approach. In this section, we provide the VAR and state space model representations of our model and we express the coefficients of these representations in terms of the parameters of our model. In this way, we offer a structural foundation for the use of the VAR and state space model approaches to gauge the informational content of high frequency news traders' flows.[28]

In order to make our model more comparable to econometric models, we consider a discrete time version of our fast model, as described in the Internet Appendix. 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 consider Δt small, so that 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. Except for β_t, all other coefficients are constant. Since econometric models are stationary, we also assume that β_t is constant.[29] For simplicity, we also write $t+1$ instead of $t+\Delta t$ and we omit the superscript F since in this section we consider only the fast model. Moreover, we assume that the econometrician can observe the trades of high frequency news traders (the informed investor in our model), since empirical papers on HFTs often relate returns to measures of net trades by groups of high frequency traders over some time interval (e.g., BHR(2012) or Kirilenko et al. (2012)).

In the fast model, the informed order flow at $t+1$ is $\Delta x_t = x_{t+1} - x_t$:

$$\Delta x_t \;=\; \beta(v_t - q_t)\Delta t + \gamma \Delta v_t \;=\; \gamma \Delta v_t + O(\Delta t). \tag{30}$$

where q_t is the quote just before the trading at $t+1$. From Theorem 1, the order flow executes at $p_{t+1} = q_t + \lambda \Delta y_t = p_t + \mu(\Delta z_{t-1} - \rho \Delta y_{t-1}) + \lambda \Delta y_t$. If we denote by $r_t = p_{t+1} - p_t$ the return contemporaneous to Δx_t, we obtain the following result using the fact that $\Delta y_t = \Delta x_t + \Delta u_t$ and the fact that $\Delta x_{t-1} = \gamma \Delta v_{t-1} + O(\Delta t)$.

Proposition 8. *The model in which the informed investor has a speed advantage implies*

[28] Our approach here is similar in spirit to that of Boulatov, Hendershott, and Livdan (2012) who derive the VAR representation of a two-period multi-asset Kyle (1985) model.

[29] This is reasonable if we consider $\beta_t = \frac{\beta_0}{1-t}$ over a small time interval. For example, if $t = 1$ corresponds to 1 month, then β_t changes very little over one hour or one day, which is the horizon over which parameters are assumed constant in empirical studies of high frequency trading.

the following VAR model for short run returns and trades:

$$\begin{aligned} r_t &= \lambda \Delta x_t + \frac{\mu}{\gamma}(1-\gamma\rho)\Delta x_{t-1} + \lambda \Delta u_t - \mu\rho\Delta u_{t-1} + \mu \Delta e_{t-1} \\ \Delta x_t &= \gamma \Delta v_t \\ \Delta u_t &= \Delta u_t. \end{aligned} \tag{31}$$

In the VAR, a shock $\Delta x_t = 1$ has a cumulative price impact

$$\sum_{\tau \geq t} r_\tau = \lambda + \frac{\mu}{\gamma}(1-\gamma\rho). \tag{32}$$

Hasbrouck (1991a) proposes to measure the informational content of a trade by the *permanent impact of a trade*, which is the sum of the predicted quote revisions after a trade innovation of a fixed size. This sum is meant to isolate the effect of the trade on liquidity suppliers' beliefs. In our case, the permanent price impact of a unit trade $\Delta x_t = 1$ is

$$\lambda_{\text{perm}} = \lambda + \frac{\mu}{\gamma}(1-\gamma\rho) = \lambda + \frac{\mu}{\gamma}\frac{1}{1+g} > \lambda, \tag{33}$$

where $g \in (0,1)$ is as in Theorem 1. In our model, however, the effect of a trade on dealers' expectation of the asset value is given by λ (see equation (9)). Thus, λ_{perm} always overestimates λ, the true informational content of a trade. The reason is the following. When the informed investor has a speed advantage, the investor's trade at date t ($\Delta x_t = \gamma \Delta v_t + O(\Delta t)$) is correlated with the news received by market makers at date $t+1$ and therefore with the market makers' quote update after receiving news ($q_{t+1} - p_{t+1} = \mu(\Delta z_t - \rho \Delta y_t)$, from equation (9)). Thus, in equation (31) of the VAR representation of the model, the lagged informed investor's order flow (Δx_{t-1}) serves as proxy for the subsequent news received by market makers. Hence, the effect of the lagged order flow captures the effect of public information arrival on market makers' quotes rather than a correction for overreaction or underreaction to the informational content of the informed investor's trade.

This observation suggests that one must be careful in interpreting measures of the informational content of trades using the Hasbrouck (1991a) approach when informed investors trade on public information because they have a speed advantage.[30] This problem is in fact discussed by Hasbrouck (1991a, Section III). It may have become

[30]This problem does not arise in the benchmark case.

more severe in recent years with the development of high frequency trading on news since this form of trading precisely exploits market makers' delayed reaction to public information.

BHR(2012) use the state space model approach to study the informational content of HFTs' trades. In this approach, the midquote at a given point in time is decomposed into a permanent component m_t and a transitory component s_t. The permanent component, m_t, is a martingale and is interpreted as the efficient value of the security. Its innovation, w_t, is modeled as a function of the *innovation* in high frequency traders' order flow. The transitory component, s_t, is interpreted as a pricing error, since it is the difference between the midquote and the fair value of the security. BHR(2012) model s_t as a stationary autoregressive component whose innovation depends on high frequency traders' order flow, ΔHFT_t. More formally, the state space model is: (i) $w_t = \kappa \widetilde{\Delta HFT}_t + \varepsilon_t$, and (ii) $s_{t+1} = \phi s_t + \psi \Delta HFT_t + \eta_t$, where ΔHFT_t is the high frequency traders' order flow over a fixed time interval (10 seconds in BHR(2012)), and $\widetilde{\Delta HFT}_t$ is the innovation in this flow.

The theoretical counterpart of ΔHFT_t in our model is Δx_t. Its innovation, $\widetilde{\Delta HFT}_t$, is $\gamma \Delta v_t$, the news trading component of the informed trader's strategy. It is natural to search for a state space representation of the midquote in which $m_t = v_t$ and $w_t = \Delta v_t$. The next proposition shows that our fast model provides a representation similar to that proposed by BHR(2012).

Proposition 9. *The model in which the informed investor has a speed advantage implies the following state space model:*

$$\begin{aligned} w_t &= \kappa \widetilde{\Delta x}_t + \varepsilon_t, \qquad s_{t+1} = \phi s_t + \psi \Delta x_t + \eta_t, \quad \textit{with} \\ \kappa &= \frac{1}{\gamma} > 0, \quad \phi = 1 - \frac{1-\mu}{\gamma} \beta \Delta t, \quad \psi = -\frac{1}{\gamma}(1-\mu-l\gamma) < 0, \end{aligned} \tag{34}$$

where $\varepsilon_t = 0$, $\eta_t = l\Delta u_t + \mu \Delta e_t$, $l = \lambda - \mu\rho$, *and the other parameters are defined as in Theorem 1.*

In their baseline estimation of the state space model, BHR(2012) find a positive coefficient for κ, equal to 4.13 bps per \$10,000 traded (see their Table 2). This is consistent with Proposition 9, since the model implies that κ is the inverse of γ, the sensitivity of the informed investor's trade to news in our model. Moreover, they show that this effect really stems from the aggressive trades of high frequency traders; that is, as in the model, those are trades from HFTs using market orders that appear to contain information. BHR(2012) interpret the positive value for κ as HFTs contributing to the discovery of the efficient price, which is also consistent with our fast model. Indeed, in the model, the trades of the informed investor due to news trading strengthen the covariance between the changes in prices and the innovation in the efficient value ($\mathsf{Cov}(dp_t, dv_t)$), and we show in the next section that this effect does improve price discovery relative to the benchmark model.

Furthermore, for their baseline model, BHR(2012) find estimates for ψ and ϕ equal to -1.97 bps per $\$10,000$ traded and 0.47, respectively. Again, these signs are consistent with those implied for those coefficients by Proposition 9. Indeed, ψ is always negative, and ϕ is positive and less than 1. Moreover, their estimates always satisfy $\kappa > |\psi|$, both in the whole sample and also across all stock size categories, as predicted by Proposition 9.

BHR(2012) interpret the negative estimate for ψ as evidence that HFTs trade against pricing error, reducing the noise in prices. Our model suggests a different interpretation: Suppose that at date t the pricing error is zero ($s_t = q_t - v_t = 0$), and the informed investor receives positive news ($\Delta v_t > 0$). The investor then buys the asset ($\Delta x_t = \gamma \Delta v_t$). During the same time period, the market maker adjusts the midquote twice: she increases the midquote by $\lambda\gamma\Delta v_t$ after she observes the order flow, then she increases it again by $\mu(\Delta v_t - \rho\gamma\Delta v_t)$ after she receives her signal (assuming $\Delta u_t = \Delta e_t = 0$). The total increase in market makers' quotes is less than Δv_t.[31] The intuition is that the order flow and the market maker's signal are two noisy signals about Δv_t. Therefore, market makers only partially adjust the quotes to innovations in the fundamental. This partial adjustment creates a temporary pricing error, s_{t+1}, negatively correlated with Δv_t. Thus, in the model, the negative value of ψ is just a consequence of the fact that market makers tend to underreact to innovations in the permanent price component.

BHR(2012) find that the autoregressive coefficient ϕ in the pricing error equation is between 0 and 1, which is consistent with Proposition 9. In our model, there is mean reversion in the pricing error because the informed investor trades against the pricing error. The speed at which s_t decays is proportional to the intensity of the forecast error component β.

[31] This follows from the fact that $1 - \mu - \gamma(\lambda - \mu\rho) = \frac{b+bg}{2+b+bg} > 0$; see also equation (A.46).

4.3 The Effect of HFTN on 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., the presence of 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 λ, the immediate price impact of a trade.

Proposition 10. *Liquidity is lower when the informed investor has a speed advantage, i.e.,* $\lambda^F > \lambda^B$.

Trades by the informed investor expose 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. This is consistent with Hendershott and Moulton (2011), who find that a reduction in execution time for market orders on the NYSE in 2006 is associated with less liquidity.

Next, we consider the effect of HFTN on price discovery. We measure price discovery by the average squared pricing error at t, i.e.,

$$\Sigma_t^k \;=\; \mathsf{E}\big((v_t - p_t^k)^2\big), \qquad k \in \{B, F\}, \tag{35}$$

where $p_t^B = p_t$ in the benchmark model and $p_t^F = q_t$ in the fast model.[32] The smaller is Σ_t^k, the higher is the informational efficiency. The next result shows that while granting a speed advantage to the informed trader does not change the average pricing error, it

[32] This definition is the same whether we use q_t or p_t, since according to (9) their difference is infinitesimal. But for the subsequent analysis we want p_t^k to be a well defined Itô process, thus we choose $p_t^F = q_t$, as explained in Footnote 15.

changes the *nature* of price discovery.[33]

Proposition 11. *For $k \in \{B, F\}$, the change in Σ_t^k is given by*

$$\mathrm{d}\Sigma_t^k \;=\; -2\,\mathsf{Cov}(\mathrm{d}p_t^k, v_t - p_t^k) - 2\,\mathsf{Cov}(\mathrm{d}p_t^k, \mathrm{d}v_t) + (2\sigma_v^2 + \Sigma_0)\mathrm{d}t. \tag{36}$$

When the informed investor has a speed advantage, short run changes in prices are more correlated with innovations in the asset value (i.e., $\mathsf{Cov}(\mathrm{d}p_t^k, \mathrm{d}v_t)$ is higher), but less correlated with the market maker's forecast error (i.e., $\mathsf{Cov}(\mathrm{d}p_t^k, v_t - p_t^k)$ is smaller). Overall, Σ_t^k is identical whether or not the informed investor has a speed advantage.

Equation (36) means that price discovery improves when short run changes in prices are more correlated with (a) news, i.e., $\mathsf{Cov}(\mathrm{d}p_t^k, \mathrm{d}v_t)$ increases, and (b) the direction of the market maker's forecast error, i.e., $\mathsf{Cov}(\mathrm{d}p_t^k, v_t - p_t)$ increases. Proposition 11 shows that the speed advantage of the informed investor has opposite effects on the two dimensions of price discovery. 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. When the informed trader has a speed advantage, returns are more correlated with news because he trades aggressively on news. By contrast, returns are less correlated with the level of the asset value in the fast model, because 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 Proposition 1.

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. More formally, as we see in the Proof of Theorem 1, the informed trader's optimization implies that Σ_t^k decreases linearly over time. The transversality condition for optimization requires that no money is left on the table at $t = 1$, i.e., $\Sigma_1^k = 0$. Since the initial value $\Sigma_0^k = \Sigma_0$ is exogenously given, the evolution of Σ_t^k is the same in both models.

We now consider the effect of HFTN on the volatility of short run returns, $\mathsf{Var}(\mathrm{d}p_t^k)$, with $k \in \{B, F\}$. This volatility has two sources in our model: (a) trading, and (b) quote updates. The second source of volatility is reflected into the quote adjustments

[33] This result is not due to the fact that p_t^k is defined differently in the two models. In both models, $\mathrm{d}p_t^k$ is the price change induced by the quote update right after the market maker observes her noisy signal about $\mathrm{d}v_t$, and by trading right after the informed trader observes $\mathrm{d}v_t$.

due to the news received by the market maker. Thus, following Hasbrouck (1991b), we decompose price volatility into the volatility coming from trades and the volatility coming from quotes,

$$\sigma_p^2 \mathrm{d}t \; = \; \mathsf{Var}(\mathrm{d}p_t^k) \; = \; \mathsf{Var}(\mathrm{d}p_{\text{trades},\,t}^k) + \mathsf{Var}(\mathrm{d}p_{\text{quotes},\,t}^k), \tag{37}$$

where for both $k \in \{B, F\}$, $\mathsf{Var}(\mathrm{d}p_{\text{trades},\,t}^k) = \mathsf{Var}(p_{t+\mathrm{d}t} - q_t)$; if $k = B$, $\mathsf{Var}(\mathrm{d}p_{\text{quotes},\,t}^k) = \mathsf{Var}(q_t - p_t)$; and if $k = F$, $\mathsf{Var}(\mathrm{d}p_{\text{quotes},\,t}^k) = \mathsf{Var}(q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t})$.

Proposition 12. *Whether the informed investor has a speed advantage or not, the instantaneous volatility of returns is constant, and equal to*

$$\sigma_p^2 \; = \; \sigma_v^2 + \Sigma_0. \tag{38}$$

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 to 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.

5 Conclusion

Adverse selection occurs in financial markets because certain investors have either more precise information, or superior speed in accessing or exploiting information. To disentangle the effects of precision and speed on market performance, we have derived the optimal trading strategy of an informed investor when he reacts to news either (i) at the same speed, or (ii) faster than the other market participants, holding information precision constant.

Our main result is that the optimal trading strategy of the informed investor is very different when has a speed advantage versus when he does not. In general, the optimal trading strategy has two components: (i) the forecast error component, which is proportional to the difference between the informed investor's and market makers' estimates of the asset payoff; and (ii) the news trading component, which is proportional to the news, i.e., to the innovation in the asset value. We have shown that the news trading component occurs only when the informed investor has a speed advantage.

As a consequence, with a speed advantage for the informed investor, his optimal portfolio is much more volatile. Also, at very high frequencies his order flow is correlated with subsequent returns, because news trading is driven by news arrivals. These features fit well with some stylized facts about high frequency traders documented in the literature. For instance, the trades of HFTs account for a large fraction of the trading volume. Moreover, the marketable orders of HFTs anticipate very short run price changes. In contrast, we show that the model in which the informed investor has more accurate information, but no speed advantage, cannot explain these facts.

We have defined high frequency trading on news (HFTN) as trading by informed investors with a speed advantage, as in our model. We have two types of empirical predictions: (i) the effect of different market characteristics on HFTN activity; and (ii) the effect of HFTN on certain measures of market quality. As an example of prediction of type (i), we show that an increase in the precision of public news increases news trading, i.e. HFTN activity, yet surprisingly improves liquidity. For type (ii), we find that an increase in HFTN activity (a) increases trading volume; (b) does not affect total price volatility, and (c) increases overall adverse selection, and thus decreases market liquidity.

Our paper is related to a fast growing literature on high frequency trading. One caveat about interpreting our work is that HFTN can be identified only with a subcategory of high frequency trading, and not with all HFT. To draw more general conclusions, one should extend the model in several directions. First, investors' information could refer not only to the asset value, but also to the order flow of other traders. Second, inventory constraints can explain the mean reversion of inventories observed in practice. Third, allowing informed traders to also submit limit orders would extend the model to another important category of HFT, the high frequency market makers.

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 explained in the discussion before Theorem 1, we consider only strategies $\mathrm{d}x_\tau$ of the type $\mathrm{d}x_\tau = \beta_\tau^B(v_\tau - p_\tau)\,\mathrm{d}\tau + \gamma_\tau^B\,\mathrm{d}v_\tau$. 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

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

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

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

For convenience, we now omit the superscript B for the coefficients β, γ, μ, λ. Then the expected profit is

$$\begin{aligned}\pi_t \;&=\; \mathsf{E}\left(\int_t^1 (v_\tau +\,\mathrm{d}v_\tau - p_\tau - \mu_\tau\,\mathrm{d}v_\tau - \lambda_\tau\,\mathrm{d}x_\tau)\,\mathrm{d}x_\tau \mid \mathcal{J}_t^p\right)\\ &=\; \int_t^1 \Big(\beta_\tau V_\tau + (1-\mu_\tau - \lambda_\tau\gamma_\tau)\gamma_\tau\sigma_v^2\Big)\,\mathrm{d}\tau.\end{aligned} \tag{A.3}$$

V_τ can be computed recursively:

$$\begin{aligned}V_{\tau+\,\mathrm{d}\tau} \;&=\; \mathsf{E}\big((v_{\tau+\,\mathrm{d}\tau} - p_{\tau+\,\mathrm{d}\tau})^2 \mid \mathcal{J}_t^p\big)\\ &=\; \mathsf{E}\big((v_\tau +\,\mathrm{d}v_\tau - p_\tau - \mu_\tau\,\mathrm{d}v_\tau - \mu_\tau\,\mathrm{d}e_\tau - \lambda_\tau\,\mathrm{d}x_\tau - \lambda_\tau\,\mathrm{d}u_\tau)^2 \mid \mathcal{J}_t^p\big)\\ &=\; V_\tau + (1-\mu_\tau-\lambda_\tau\gamma_\tau)^2\sigma_v^2\,\mathrm{d}\tau + \mu_\tau^2\sigma_e^2\,\mathrm{d}\tau + \lambda_\tau^2\sigma_u^2\,\mathrm{d}\tau - 2\lambda_t\beta_t V_\tau\,\mathrm{d}\tau.\end{aligned} \tag{A.4}$$

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

$$V_\tau' \;=\; -2\lambda_t\beta_t V_\tau + (1-\mu_\tau-\lambda_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + \lambda_\tau^2\sigma_u^2, \tag{A.5}$$

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 (A.1) and integrate by parts. Since $V_t = 0$, we get

$$\begin{aligned}\pi_t \;=\; & -\frac{V_1}{2\lambda_1} + \int_t^1 V_\tau \Big(\frac{1}{2\lambda_\tau}\Big)' \,\mathrm{d}\tau \\ & + \int_t^1 \bigg(\frac{(1-\mu_\tau-\lambda_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + \lambda_\tau^2\sigma_u^2}{2\lambda_\tau} + (1-\mu_\tau-\lambda_\tau\gamma_\tau)\gamma_\tau\sigma_v^2 \bigg) \,\mathrm{d}\tau.\end{aligned} \tag{A.6}$$

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

$$\lambda_\tau \;=\; \text{constant} \;=\; \lambda. \tag{A.7}$$

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

$$-(1-\mu_\tau-\lambda_\tau\gamma_\tau) + (1-\mu_\tau-\lambda_\tau\gamma_\tau) - \lambda_\tau\gamma_\tau \;=\; 0 \quad \Longrightarrow \quad \gamma_\tau \;=\; 0. \tag{A.8}$$

Thus, there is no news trading in the benchmark model. Note also that the second order condition is $\lambda_\tau > 0$.[34]

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

$$\mu_t \;=\; \frac{\mathsf{Cov}(v_1,\, \mathrm{d}z_t \mid \mathcal{I}_t^p)}{\mathsf{Var}(\,\mathrm{d}z_t \mid \mathcal{I}_t^p)} \;=\; \frac{\mathsf{Cov}(v_0 + \int_0^1 \mathrm{d}v_\tau,\, \mathrm{d}v_t + \mathrm{d}e_t \mid \mathcal{I}_t^p)}{\mathsf{Var}(\,\mathrm{d}v_t + \mathrm{d}e_t \mid \mathcal{I}_t^p)} \;=\; \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2} \;=\; \mu. \tag{A.9}$$

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, since $\lambda_t = \lambda$ is constant,

$$\lambda \;=\; \frac{\mathsf{Cov}(v_1,\, \mathrm{d}y_t \mid \mathcal{I}_{t+\mathrm{d}t}^q)}{\mathsf{Var}(\,\mathrm{d}y_t \mid \mathcal{I}_{t+\mathrm{d}t}^q)} \;=\; \frac{\mathsf{Cov}(v_1, \beta_t(v_t - p_t)\,\mathrm{d}t + \mathrm{d}u_t \mid \mathcal{I}_{t+\mathrm{d}t}^q)}{\mathsf{Var}(\beta_t(v_t - p_t)\,\mathrm{d}t + \mathrm{d}u_t \mid \mathcal{I}_{t+\mathrm{d}t}^q)} \;=\; \frac{\beta_t \Sigma_t}{\sigma_u^2}, \tag{A.10}$$

where $\Sigma_t = \mathsf{E}\big((v_t - p_t)^2 | \mathcal{I}_t^p\big)$.[35] 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

[34] 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 (A.5) 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.

[35] 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:

$$\Sigma_t' \;=\; -2\lambda\beta_t\Sigma_t + (1-\mu)^2\sigma_v^2 + \mu^2\sigma_e^2 + \lambda^2\sigma_u^2, \tag{A.11}$$

except that it has a different initial condition. If we solve this first order differential equation explicitly, it follows 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 $\Sigma_1 = 0$. By (A.10), we get $\beta_t \Sigma_t = \lambda\sigma_u^2$ is constant. Equation (A.11) then implies that Σ_t' is constant. From $\Sigma_1 = 0$, we get $\Sigma_t = (1-t)\Sigma_0$, and $\beta_t = \frac{\beta_0}{1-t}$. Then, (A.11) becomes $-\Sigma_0 = -2\lambda^2\sigma_u^2 + (1-\mu)^2\sigma_v^2 + \mu^2\sigma_e^2 + \lambda^2\sigma_u^2$. Since $\mu = \frac{\sigma_v^2}{\sigma_v^2+\sigma_e^2}$, we get $\lambda^2\sigma_u^2 = \Sigma_0 + \frac{\sigma_v^2\sigma_e^2}{\sigma_v^2+\sigma_e^2}$, which implies (15). Then, $\beta_0 = \frac{\lambda\sigma_u^2}{\Sigma_0}$ and $\beta_t = \frac{\beta_0}{1-t}$ imply (13).

Fast model: The informed trader has the same objective function as in (A.1):

$$\pi_t \;=\; \mathsf{E}\left(\int_t^1 (v_1 - p_{\tau+\mathrm{d}\tau})\,\mathrm{d}x_\tau \mid \mathcal{J}_t^p\right). \tag{A.12}$$

but here we use q_t instead of p_t as a state variable. From (18), we obtain

$$\begin{aligned} q_{\tau+\mathrm{d}\tau} \;&=\; \mu_\tau^F\,\mathrm{d}z_\tau + l_\tau^F\,\mathrm{d}y_\tau, \quad \text{with} \\ l_\tau^F \;&=\; \lambda_\tau^F - \mu_\tau^F\rho_\tau^F. \end{aligned} \tag{A.13}$$

As explained in the discussion before Theorem 1, we consider only strategies $\mathrm{d}x_\tau$ of the type (17), $\mathrm{d}x_\tau = \beta_\tau^F(v_\tau - q_\tau)\,\mathrm{d}\tau + \gamma_\tau^F\,\mathrm{d}v_\tau$. For $\tau \geq t$, denote by

$$V_\tau \;=\; \mathsf{E}\big((v_\tau - q_\tau)^2 \mid \mathcal{J}_t^p\big). \tag{A.14}$$

For convenience, we now omit the superscript F for the coefficients β, γ, μ, λ, ρ, l. The expected profit is

$$\begin{aligned} \pi_t \;&=\; \left(\int_t^1 (v_\tau + \,\mathrm{d}v_\tau - q_\tau - \lambda_\tau\,\mathrm{d}x_\tau)\,\mathrm{d}x_\tau \mid \mathcal{J}_t^p\right) \\ &=\; \int_t^1 \Big(\beta_\tau V_\tau + (1-\lambda_\tau\gamma_\tau)\gamma_\tau\sigma_v^2\Big)\,\mathrm{d}\tau. \end{aligned} \tag{A.15}$$

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

$$\begin{aligned} V_{\tau+\,\mathrm{d}\tau} \;&=\; \mathsf{E}\big((v_{\tau+\,\mathrm{d}\tau} - q_{\tau+\,\mathrm{d}\tau})^2 \mid \mathcal{J}^p_t\big) \\ &=\; V_\tau + (1-\mu_\tau - l_\tau\gamma_\tau)^2\sigma_v^2\,\mathrm{d}\tau + \mu_\tau^2\sigma_e^2\,\mathrm{d}\tau + l_\tau^2\sigma_u^2\,\mathrm{d}\tau - 2l_t\beta_t V_\tau\,\mathrm{d}\tau. \end{aligned} \tag{A.16}$$

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

$$V'_\tau \;=\; -2l_t\beta_t V_\tau + (1-\mu_\tau - l_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + l_\tau^2\sigma_u^2, \tag{A.17}$$

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

$$\begin{aligned} \pi_t \;&=\; -\frac{V_1}{2l_1} + \int_t^1 V_\tau\Big(\frac{1}{2l_\tau}\Big)'\,\mathrm{d}\tau \\ &\quad+ \int_t^1 \bigg(\frac{(1-\mu_\tau - l_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + l_\tau^2\sigma_u^2}{2l_\tau} + (1-\lambda_\tau\gamma_\tau)\gamma_\tau\sigma_v^2\bigg)\bigg)\,\mathrm{d}\tau. \end{aligned} \tag{A.18}$$

Since $V_\tau > 0$ can be arbitrarily chosen, in order to get an optimum we must have $\big(\frac{1}{2l_\tau}\big)' = 0$, which is equivalent to $l_\tau =$ 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

$$-(1-\mu_\tau - l_\tau\gamma_\tau) + (1-\lambda_\tau\gamma_\tau) - \lambda_\tau\gamma_\tau \;=\; 0 \quad\Longrightarrow\quad \gamma_\tau \;=\; \frac{\mu_\tau}{2\lambda_\tau - l_\tau} \;=\; \frac{\mu_\tau}{\lambda_\tau + \mu_\tau\rho_\tau}. \tag{A.19}$$

Thus, we obtain a nonzero news 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: $l_\tau > 0$: see Footnote 34.

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

$$\begin{aligned} \lambda_t \;&=\; \frac{\mathsf{Cov}_t(v_1,\,\mathrm{d}y_t)}{\mathsf{Var}_t(\,\mathrm{d}y_t)} \;=\; \frac{\mathsf{Cov}_t(v_1, \beta_t(v_t - p_t)\,\mathrm{d}t + \gamma_t\,\mathrm{d}v_t +\,\mathrm{d}u_t)}{\mathsf{Var}(\beta_t(v_t-p_t)\,\mathrm{d}t + \gamma_t\,\mathrm{d}v_t +\,\mathrm{d}u_t)} \;=\; \frac{\beta_t\Sigma_t + \gamma_t\sigma_v^2}{\gamma_t^2\sigma_v^2 + \sigma_u^2}, \\ \rho_t \;&=\; \frac{\mathsf{Cov}_t(\,\mathrm{d}z_t,\,\mathrm{d}y_t)}{\mathsf{Var}_t(\,\mathrm{d}y_t)} \;=\; \frac{\gamma_t\sigma_v^2}{\gamma_t^2\sigma_v^2 + \sigma_u^2}, \\ \mu_t \;&=\; \frac{\mathsf{Cov}_t(v_1,\,\mathrm{d}z_t - \rho_t\,\mathrm{d}y_t)}{\mathsf{Var}_t(\,\mathrm{d}z_t - \rho_t\,\mathrm{d}y_t)} \;=\; \frac{-\rho_t\beta_t\Sigma_t + (1-\rho_t\gamma_t)\sigma_v^2}{(1-\rho_t\gamma_t)^2\sigma_v^2 + \rho_t^2\sigma_u^2 + \sigma_e^2}. \end{aligned} \tag{A.20}$$

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 (A.17) as V_t, and $\Sigma_t' = -\Sigma_0$, we obtain

$$-\Sigma_0 \;=\; -2l_t\beta_t\Sigma_t + (1-\mu_\tau - l_\tau\gamma_\tau)^2\sigma_v^2 + \mu_\tau^2\sigma_e^2 + l_\tau^2\sigma_u^2. \tag{A.21}$$

We now define the following constants:

$$\begin{aligned} a &= \frac{\sigma_u^2}{\sigma_v^2}, \quad b = \frac{\sigma_e^2}{\sigma_v^2}, \quad c = \frac{\Sigma_0}{\sigma_v^2}, \\ g &= \frac{\gamma^2}{a}, \quad \tilde{\lambda} = \lambda\gamma, \quad \tilde{\rho} = \rho\gamma, \quad \nu = \frac{\beta_0\Sigma_0}{\sigma_u^2}\gamma, \quad \tilde{l} = l\gamma. \end{aligned} \tag{A.22}$$

With these notations, equations (A.19)–(A.21) become

$$\begin{aligned} \tilde{\lambda} &= \mu(1-\tilde{\rho}), \quad \tilde{\lambda} = \frac{\nu+g}{1+g}, \quad \tilde{\rho} = \frac{g}{1+g}, \quad \mu = \frac{1-\nu}{1+b(1+g)} \\ c &= \frac{2\nu}{g} - (1-\mu-\tilde{l})^2 - \mu^2 b - \frac{\tilde{l}^2}{g}. \end{aligned} \tag{A.23}$$

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

$$\nu \;=\; \frac{1-(1+b)g - bg^2}{2+b+bg} \;=\; \frac{1+g}{2+b+bg} - g. \tag{A.24}$$

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

$$\tilde{\lambda} = \frac{1}{2+b+bg}, \quad \tilde{\rho} = \frac{g}{1+g}, \quad \mu = \frac{1+g}{2+b+bg}, \quad \tilde{l} = \frac{1-g}{2+b+bg}, \tag{A.25}$$

$$1+c \;=\; \frac{(1+bg)(1+g)^2}{g(2+b+bg)^2}. \tag{A.26}$$

From (A.22), we get

$$\gamma \;=\; a^{1/2}g^{1/2}, \quad \beta_0 \;=\; \frac{\sigma_u^2}{\Sigma_0\gamma}\nu \;=\; \frac{a}{c\gamma}\nu \;=\; \frac{a^{1/2}}{cg^{1/2}}\nu. \tag{A.27}$$

From (A.24) and (A.26), we get $\nu = \frac{1+g}{2+b+bg} - g = \frac{g(2+b+bg)}{(1+g)(1+bg)}\left(\frac{(1+g)^2(1+bg)}{g(2+b+bg)^2} - \frac{(1+g)(1+bg)}{2+b+bg}\right) = \frac{g(2+b+bg)}{(1+g)(1+bg)}\left(c+1-\frac{(1+g)(1+bg)}{2+b+bg}\right) = \frac{g(2+b+bg)}{(1+g)(1+bg)}\left(c+(1-g)\frac{(1+b+bg)}{2+b+bg}\right)$. We compute $\beta_0 = \frac{a^{1/2}g^{1/2}(2+b+bg)}{(1+g)(1+bg)}\left(1+\frac{1-g}{c}\frac{1+b+bg}{2+b+bg}\right)$. Using again (A.26), we get

$$\beta_0 \;=\; \frac{a^{1/2}}{(1+c)^{1/2}(1+bg)^{1/2}}\left(1+\frac{1-g}{c}\frac{1+b+bg}{2+b+bg}\right). \tag{A.28}$$

Now substitute a, b, c from (A.22) in equations (A.25)–(A.28) to obtain equations (19)–(24). Moreover, the second order conditions $\lambda + \mu\rho > 0$ and $l > 0$ are equivalent to $g \in (-1, 1)$.

Finally, we show that the equation $1 + c = \frac{(1+bg)(1+g)^2}{g(2+b+bg)^2}$ has a unique solution $g \in (-1, 1)$, which in fact lies in $(0, 1)$. This can be shown by noting that

$$F_b(g) \;=\; 1 + c, \quad \text{with} \quad F_b(x) \;=\; \frac{(1 + bx)(1 + x)^2}{x(2 + b + bx)^2}. \tag{A.29}$$

One verifies $F_b'(x) = \frac{(x+1)(x-1)(2+b+3bx)}{x^2(2+b+bx)^3}$, 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 $g \in (0, 1)$ so that $F_b(g) = 1 + c$.[36]

A.2 Useful Comparative Statics

To compare the fast and benchmark models, and to do some comparative statics for the coefficients involved in Theorem 1, we prove the following result.

Lemma A.1. *With the notations in Theorem 1, the following inequalities are true:*

$$\mu^F \;<\; \mu^B, \quad \lambda^F \;>\; \lambda^B, \quad \beta_0^F \;<\; \beta_0^B. \tag{A.30}$$

Proof. Recall that in the proof of Theorem 1, we have denoted

$$a \;=\; \frac{\sigma_u^2}{\sigma_v^2}, \quad b \;=\; \frac{\sigma_e^2}{\sigma_v^2}, \quad c \;=\; \frac{\Sigma_0}{\sigma_v^2}. \tag{A.31}$$

We start by showing that

$$\mu^F \;=\; \frac{1 + g}{2 + b + bg} \;<\; \mu^B \;=\; \frac{1}{1 + b}. \tag{A.32}$$

By computation, this is equivalent to $g < 1$, which is true since $g \in (0, 1)$.

We show that

$$\lambda^F \;=\; \frac{(1 + c)^{1/2}}{a^{1/2}} \frac{1}{(1 + bg)^{1/2}(1 + g)} \;>\; \lambda^B \;=\; \frac{c^{1/2}}{a^{1/2}} \left(1 + \frac{b}{c(b + 1)}\right)^{1/2}. \tag{A.33}$$

After squaring the two sides, and using $1 + c = \frac{(1+bg)(1+g)^2}{g(2+b+bg)^2}$, we need to prove that

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

$\frac{1}{g(2+b+bg)^2} > c+1-\frac{1}{1+b}$, or equivalently $\frac{1}{1+b} > \frac{(1+bg)(1+g)^2}{g(2+b+bg)^2} - \frac{1}{g(2+b+bg)^2} = \frac{2+b+g+2bg+bg^2}{(2+b+bg)^2}$. This reduces to proving $1 + b + (1 - g)(1 + bg) > 0$, which is true, since $b > 0$ and $g \in (0, 1)$.

In the proof of Theorem 1, we have $\nu = \frac{1+g}{2+b+bg} - g = \frac{g(2+b+bg)}{(1+g)(1+bg)}\left(c+(1-g)\frac{(1+b+bg)}{2+b+bg}\right) > 0$. But $\frac{1+g}{2+b+bg} > g$ implies $bg < \frac{1-g}{1+g}$. We now show that

$$\beta_0^F = \frac{a^{1/2}}{cg^{1/2}}\left(\frac{1+g}{2+b+bg} - g\right) < \beta_0^B = \frac{a^{1/2}}{c}\left(c + \frac{b}{1+b}\right)^{1/2}, \tag{A.34}$$

where we use (A.24) and (A.27) for β_0^F, and (13) for β_0^B. Using (A.26), the desired inequality is equivalent to $\frac{1}{g}\frac{(1-g-bg-bg^2)^2}{(2+b+bg)^2} < c + 1 - \frac{1}{1+b} = \frac{(1+bg)(1+g)^2}{g(2+b+bg)^2} - \frac{1}{1+b}$, or $\frac{1}{1+b} < \frac{4+3b+bg(2-b)-bg^2(1+2b)-b^2g^3}{(2+b+bg)^2}$. After some algebra, this is equivalent to $bg^2(1+g)^2 + bg(1+4g+g^2) < 3+2b$. We use $bg < \frac{1-g}{1+g}$ (proved above) to show that $bg^2(1+g)^2 < g(1-g^2)$ and $bg(1 + 4g + g^2) < (1 - g)\frac{1+4g+g^2}{1+g} < (1 - g)(1 + 3g)$. Then, it is sufficient to prove that $g(1 - g^2) + (1 - g)(1 + 3g) < 3 + 2b$, or $1 + 3g - 3g^2 - g^3 < 3 + 2b$. For this, it is sufficient to prove $1 + 3g - 3g^2 < 3 + 2b$. But $1 + 3g - 3g^2$ attains its maximum value of $1 + \frac{3}{4}$ at $g = \frac{1}{2}$, and $1 + \frac{3}{4} < 3 + 2b$. $\square$

A.3 Proof of Proposition 1

See Lemma A.1.

A.4 Proof of Proposition 2

For the benchmark model, as in Theorem 1 and Lemma A.1, we have

$$\beta_0^B = \frac{\sigma_u}{\Sigma_0^{1/2}}\left(1 + \frac{\sigma_v^2\sigma_e^2}{\sigma_v^2+\sigma_e^2}\right)^{1/2} = \frac{a^{1/2}}{c}\left(c + \frac{b}{1+b}\right)^{1/2}. \tag{A.35}$$

From the first equality, since $\frac{\sigma_v^2\sigma_e^2}{\sigma_v^2+\sigma_e^2}$ is increasing in σ_v, so is β_0^B. From the second equality, β_0^B is increasing in $a = \frac{\sigma_u^2}{\sigma_v^2}$ and $b = \frac{\sigma_e^2}{\sigma_v^2}$, and decreasing in $c = \frac{\Sigma_0}{\sigma_v^2}$, and thus β_0^B is increasing in σ_u and σ_e, and decreasing in Σ_0.

As in the proof of Theorem 1, let $F(b, x) = \frac{(1+bx)(1+x)^2}{x(2+b+bx)^2}$, with $\frac{\partial F}{\partial b} = -\frac{(1+x)^2(2+bx+bx^2)}{x(2+b+bx)^3}$ and $\frac{\partial F}{\partial x} = -\frac{(1-x)(1+x)(2+b+3bx)}{x^2(2+b+bx)^3}$. Since $g \in (0, 1)$ is the solution of $F(b, g(b, c)) = 1 + c$, by differentiating with respect to b and c, respectively, we get $\frac{\partial F}{\partial b} + \frac{\partial F}{\partial x}\frac{\partial g}{\partial b} = 0$, and

$\frac{\partial F}{\partial x}\frac{\partial g}{\partial c} = 1$. We compute

$$\frac{\partial g}{\partial b} \;=\; -\frac{g(1+g)(2+bg+bg^2)}{(1-g)(2+b+3bg)}, \quad \frac{\partial g}{\partial c} \;=\; -\frac{g^2(2+b+bg)^3}{(1-g)(1+g)(2+b+3bg)}. \tag{A.36}$$

This implies that g is decreasing in b and c. Since $\gamma^F = \frac{\sigma_u}{\sigma_v}g^{1/2} = a^{1/2}g^{1/2}$, γ^F is increasing in a (and σ_u); and decreasing in b, c (and σ_e, Σ_0).

From the proof of Theorem 1, we also have

$$\beta_0^F \;=\; \frac{a^{1/2}}{cg^{1/2}}\left(\frac{1+g}{2+b+bg} - g\right). \tag{A.37}$$

Since g does not depend on a, β_0^F is increasing in a (and σ_u). We use (A.36) to compute $\frac{\partial \beta_0^F}{\partial b} = -\frac{a^{1/2}g^{1/2}(1+g)^2(2+3b+3bg+b^2g+b^2g^2)}{2c(1-g)(4+4b+8bg+b^2+4b^2g+3b^2g^2)}$, thus β_0^F is decreasing in b (and σ_e). Similarly, we find that $\frac{\partial \beta_0^F}{\partial c}$ is proportional to $-\frac{1}{c^2}\big(\frac{1+g}{2+b+bg} - g\big) + \frac{g(2+b+bg)(2+b+5bg+b^2g+b^2g^2)}{2c(1-g)(2+b+3bg)}$. Substituting $c = \frac{(1+bg)(1+g)^2}{g(2+b+bg)^2} - 1$, we obtain that $\frac{\partial \beta_0^F}{\partial c}$ is proportional to $2(1-g)(2+b+3bg)(g+bg+bg^2-1) + \big((1+bg)(1+g)^2 - g(2+b+bg)^2\big)(2+b+5bg+b^2g+b^2g^2)$. This is a polynomial in b and g, which can be written as $-2(1-g)^2 - b^3g(1-g^4) - P(b,g)$, where $P(b,g)$ is a polynomial with positive coefficients. Since $b > 0$ and $g \in (0,1)$, we get $\frac{\partial \beta_0^F}{\partial c} < 0$, thus β_0^F is decreasing in c (and Σ_0). Similarly, if we denote $\Sigma_v = \sigma_v^2$, we find that $\frac{\partial \beta_0^F}{\partial \Sigma_v}$ is proportional to $b^3g(1+g)^3 + b^2g^3 + 2b^2g^3(1-g) + 13b^2g^2 + 9b^2g + b^2 + 4bg^2 + 4bg + 6bg(1-g^2) + 4g(1-g)$, which is positive. Thus, β_0^F is increasing in σ_v.

A.5 Proof of Proposition 3

In the benchmark model, equation (11) and $\gamma_t^B = 0$ imply that $\mathsf{Var}(\mathrm{d}x_t) = (\beta_t^B)^2\Sigma_t \mathrm{d}t^2 = 0$, since $\mathrm{d}t^2 = 0$. Also, $\mathsf{Var}(\mathrm{d}u_t) = \sigma_u^2\mathrm{d}t$. Thus, $IPR_t^B = \frac{\mathsf{Var}(\mathrm{d}x_t)}{\mathsf{Var}(\mathrm{d}x_t)+\mathsf{Var}(\mathrm{d}u_t)} = 0$.

In the fast model, equation (17) implies $\mathsf{Var}(\mathrm{d}x_t) = (\gamma_t^F)^2\sigma_v^2\mathrm{d}t$, and equation (20) implies $(\gamma_t^F)^2\sigma_v^2 = \sigma_u^2 g$. Therefore, $IPR_t^F = \frac{\sigma_u^2 g\mathrm{d}t}{\sigma_u^2 g\mathrm{d}t+\sigma_u^2\mathrm{d}t} = \frac{g}{g+1}$. From Theorem 1, we know that $g \in (0,1)$.

A.6 Proof of Proposition 4

For $k = \{B, F\}$, we write the equilibrium equations

$$\begin{aligned} \mathrm{d}x_t \;&=\; \beta_t^k(v_t - p_t)\mathrm{d}t + \gamma^k\mathrm{d}v_t, \\ \mathrm{d}p_t \;&=\; \lambda^k\mathrm{d}y_t + \mu^k(\mathrm{d}z_t - \rho^k\mathrm{d}y_t) \;=\; l^k\mathrm{d}y_t + \mu^k\mathrm{d}z_t. \end{aligned} \tag{A.38}$$

We first prove the following useful result.

Lemma A.2. *In both the benchmark and the fast models, i.e., if $k \in \{B, F\}$, and for all $s < t \in (0,1)$,*

$$\begin{aligned}
\mathsf{Cov}(v_s - p_s, v_t - p_t) &= \Sigma_s \left(\frac{1-t}{1-s}\right)^{l^k \beta_0^k}, \\
\frac{1}{\mathrm{d}s}\mathsf{Cov}(\mathrm{d}v_s, v_t - p_t) &= (1 - l^k\gamma^k - \mu^k)\sigma_v^2 \left(\frac{1-t}{1-s}\right)^{l^k \beta_0^k},
\end{aligned} \tag{A.39}$$

where $l^k = \lambda^k - \mu^k \rho^k$.

Proof. Denote by $X_t = \mathsf{Cov}(v_s - p_s, v_t - p_t)$. For $t \geq s$, $\mathrm{d}X_t = \mathsf{Cov}(v_s - p_s, \mathrm{d}v_t - \mathrm{d}p_t) = -\mathsf{Cov}(v_s - p_s, \mathrm{d}p_t) = -l^k \beta_t^k X_t \mathrm{d}t = -l^k \frac{\beta_0^k}{1-t} X_t \mathrm{d}t$. Then, $\mathrm{d}\ln(X_t) = l^k \beta_0^k \, \mathrm{d}\ln(1-t)$. Also, at $t = s$, we have $X_s = \Sigma_s$. Thus, we have a first order differential equation, with solution given by the first equation in (A.39).

Denote by $Y_t = \frac{1}{\mathrm{d}s}\mathsf{Cov}(\mathrm{d}v_s, v_t - p_t)$. For $t > s$, $\mathrm{d}Y_t = \frac{1}{\mathrm{d}s}\mathsf{Cov}(\mathrm{d}v_s, \mathrm{d}v_t - \mathrm{d}p_t) = -\frac{1}{\mathrm{d}s}\mathsf{Cov}(\mathrm{d}v_s, \mathrm{d}p_t) = -l^k \beta_t^k Y_t \mathrm{d}t = -l^k \frac{\beta_0^k}{1-t} Y_t \mathrm{d}t$. Then, $\mathrm{d}\ln(Y_t) = l^k \beta_0^k \, \mathrm{d}\ln(1-t)$. At $t = s + \mathrm{d}s$, we have $Y_{s+\mathrm{d}s} = \frac{1}{\mathrm{d}s}\mathsf{Cov}(\mathrm{d}v_s, v_s - p_s + \mathrm{d}v_s - \mathrm{d}p_s) = \frac{1}{\mathrm{d}s}\mathsf{Cov}(\mathrm{d}v_s, \mathrm{d}v_s) - \frac{1}{\mathrm{d}s}\mathsf{Cov}(\mathrm{d}v_s, l^k \mathrm{d}y_t + \mu^k \mathrm{d}z_t) = (1 - l^k\gamma^k - \mu^k)\sigma_v^2$. Thus, we have a first order differential equation, with solution given by the second equation in (A.39).

□

We now prove Proposition 4. For the benchmark model, $l^B = \lambda^B$. Then, using the notations from Lemma A.2, we get

$$\mathsf{Corr}(\mathrm{d}x_t^B, \mathrm{d}x_{t+\tau}^B) = \frac{\mathsf{Cov}(v_t - p_t, v_{t+\tau} - p_{t+\tau})}{\mathsf{Cov}(v_t - p_t)^{1/2}\,\mathsf{Cov}(v_{t+\tau} - p_{t+\tau})^{1/2}} = \frac{\Sigma_t \left(\frac{1-t-\tau}{1-t}\right)^{\lambda^B \beta_0^B}}{\Sigma_t^{1/2}\Sigma_{t+\tau}^{1/2}}. \tag{A.40}$$

Since $\Sigma_s = \Sigma_0(1-s)$, we obtain $\mathsf{Corr}(\mathrm{d}x_t^B, \mathrm{d}x_{t+\tau}^B) = \left(\frac{1-t-\tau}{1-t}\right)^{\lambda^B \beta_0^B - \frac{1}{2}}$.[37]

In the fast model, we use both equations in (A.39) to show that the autocovariance of the informed order flow, $\mathsf{Cov}(\mathrm{d}x_t^F, \mathrm{d}x_{t+\tau}^F)$, is of order $\mathrm{d}t^2$. But the informed order flow variance is of order $\mathrm{d}t$, therefore the autocorrelation is of order $\mathrm{d}t$, which is zero in continuous time.

[37]Note that $\lambda^B \beta_0^B = 1 + \frac{\sigma_v^2 \sigma_e^2}{\Sigma_0(\sigma_v^2 + \sigma_e^2)} > 1$.

A.7 Proof of Proposition 5

For $k \in \{B, F\}$, recall that $AT_t^k = \mathsf{Corr}(\mathrm{d}x_t, q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t})$, and $\mathrm{d}x_t = \beta_t^k(v_t - p_t)\mathrm{d}t + \gamma^k \mathrm{d}v_t$. In the benchmark, equation (8) implies $q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t} = \mu^B \mathrm{d}z_{t+\mathrm{d}t}$. Therefore, $AT_t^B = 0$.

In the fast model, equation (9) implies $q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t} = \mu^F(\mathrm{d}z_t - \rho^F \mathrm{d}y_t)$, thus

$$q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t} \;=\; \mu^F(1 - \rho^F\gamma^F)\mathrm{d}v_t + \mu^F \mathrm{d}e_t - \mu^F\rho^F \mathrm{d}u_t. \tag{A.41}$$

Then, $\frac{1}{\mathrm{d}t}\mathsf{Cov}(\mathrm{d}x_t, q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t}) = \gamma^F\mu^F(1-\rho^F\gamma^F)\sigma_v^2$. Moreover, $\frac{1}{\mathrm{d}t}\mathsf{Var}(\mathrm{d}x_t) = (\gamma^F)^2\sigma_v^2$ and $\frac{1}{\mathrm{d}t}\mathsf{Var}(q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t}) = (\mu^F)^2(1-\rho^F\gamma^F)^2\sigma_v^2 + (\mu^F)^2\sigma_e^2 + (\mu^F)^2(\rho^F)^2\sigma_u^2$. Together, these imply $AT^F = \frac{(1-\rho^F\gamma^F)\sigma_v}{\sqrt{(1-\rho^F\gamma^F)^2\sigma_v^2 + \sigma_e^2 + (\rho^F)^2\sigma_u^2}}$. With $a = \frac{\sigma_u^2}{\sigma_v^2}$ and $b = \frac{\sigma_e^2}{\sigma_v^2}$ as in (A.22), we use the formulas $\gamma^F = a^{1/2}g^{1/2}$ and $\rho^F = \frac{g^{1/2}}{a^{1/2}(1+g)}$ in Theorem 1 to compute $AT^F = \frac{1}{\sqrt{(1+g)(1+b)}}$. This proves (29) for AT_t^F.

A.8 Proof of Proposition 6

In the fast model, denote by $TV^F = \mathsf{Var}(\mathrm{d}y_t)$ the trading volume, and $IPR^F = \frac{\mathsf{Var}(\mathrm{d}x_t)}{\mathsf{Var}(\mathrm{d}y_t)}$ the informed participation rate. Then, by Proposition 3, $TV^F = \sigma_u^2(1+g)$, and $IPR^F = \frac{g}{1+g}$, hence TV^F and IPR^F have the same dependence on σ_e as g. From (A.36), g is decreasing in b, hence also in σ_e. Thus, both TV^F and IPR^F are decreasing in σ_e.

As in equation (A.33), we have $(\lambda^F)^2 = \frac{1+c}{a}\frac{1}{(1+bg)(1+g)^2}$. Using the formula for $\frac{\partial g}{\partial b}$ in (A.36), we compute $\frac{\partial\big((1+bg)(1+g)^2\big)}{\partial b} = -\frac{g(1+g)^3(1+bg)}{1-g} < 0$. Therefore, λ^F is increasing in b, hence in σ_e. Thus, higher precision of the public signal (lower σ_e) implies higher liquidity (lower price impact coefficient λ^F).

A.9 Proof of Proposition 7

As in proof of Proposition 6, the trading volume and the informed participation rate in the fast model are given, respectively, by $TV^F = \sigma_u^2(1+g)$ and $IPR^F = \frac{g}{1+g}$, hence TV^F and IPR^F have the same dependence on σ_v (holding $b = \frac{\sigma_e^2}{\sigma_v^2}$ constant) as g. In the proof of Proposition 2, we analyze $g = g(b, c)$ as a function of $b = \frac{\sigma_e^2}{\sigma_v^2}$ and $c = \frac{\Sigma_0}{\sigma_v^2}$. If we hold b constant, then g depends on σ_v only via c. From equation (A.36) we have $\frac{\partial g}{\partial c} = -\frac{g^2(2+b+bg)^3}{(1-g)(1+g)(2+b+3bg)} < 0$, thus g is decreasing in c. Therefore, g is increasing in σ_v, hence both TV^F and IPR^F are increasing in σ_v.

Denote by $\Sigma_v = \sigma_v^2$. As in proof of Proposition 6, we use $(\lambda^F)^2 = \frac{1+c}{a}\frac{1}{(1+bg)(1+g)^2} = \frac{\Sigma_v+\Sigma_0}{\sigma_u^2}\frac{1}{(1+bg)(1+g)^2}$. We now hold b constant, and differentiate g only with respect to $c = \frac{\Sigma_0}{\Sigma_v}$. After substituting also $1+c = \frac{(1+bg)(1+g)^2}{g(2+b+bg)^2}$, we obtain $\left(\frac{\partial(\lambda^F)^2}{\partial\Sigma_v}\right)_{b=\text{const}} = \frac{1-g+b+bg}{\sigma_u^2(1+g)(1-g)(2+b+bg)} > 0$. Thus, if σ_v increases and $\frac{\sigma_e}{\sigma_v}$ is held constant, λ^F increases.

A.10 Proof of Proposition 8

From (9), we have $r_t = p_{t+1} - p_t = \lambda\Delta y_t + \mu(\Delta z_{t-1} - \rho\Delta y_{t-1})$. Also, $\Delta z_t = \Delta v_t + \Delta e_t$, and $\Delta x_t = \gamma\Delta v_t + O(\Delta t)$. Thus, up to terms of order of Δt, we have

$$r_t \; = \; \lambda\Delta x_t + \lambda\Delta u_t + \mu\Delta x_{t-1} + \mu\Delta e_{t-1} - \mu\rho\Delta x_{t-1} - \mu\rho\Delta u_{t-1}, \tag{A.42}$$

from which we get the first equation in (31). The other two equations in the VAR representation are straightforward.

Next, start with a shock $\Delta x_t = 1$. This has an immediate effect, $r_t = \lambda$, but also a lagged effect, $r_{t+1} = \frac{\mu}{\gamma}(1-\gamma\rho)$. After that, $r_{t+\tau} = 0$ for $\tau \geq 2$, hence

$$\sum_{\tau\geq t} r_\tau \; = \; r_t + r_{t+1} \; = \; \lambda + \frac{\mu}{\gamma}(1-\gamma\rho), \tag{A.43}$$

which implies (32).

A.11 Proof of Proposition 9

We discretize the continuous time fast model, and write $t+1$ instead of $t+\Delta t$. We also remove the superscript F since we focus on the fast model. Since in this section, we work under the assumption that β is constant, equations (17) and (18) imply

$$\Delta x_t \; = \; -\beta s_t \Delta t + \gamma\Delta v_t, \qquad \Delta q_t \; = \; l\Delta y_t + \mu\Delta z_t, \tag{A.44}$$

where $s_t = q_t - v_t$ is the pricing error, and $l = \lambda - \mu\rho$. Then, using (A.44), we rewrite $s_{t+1} = s_t + \Delta q_t - \Delta v_t$ as follows:

$$s_{t+1} \; = \; (1 - l\beta\Delta t)s_t - (1-\mu-l\gamma)\Delta v_t + \eta_t. \tag{A.45}$$

with $\eta_t = l\Delta u_t + \mu\Delta u_t$. Using $\Delta v_t = \frac{1}{\gamma}\Delta x_t + \frac{\beta\Delta t}{\gamma}\, s_t$, we obtain the desired formulas for ϕ and ψ.

Using equations (A.25), we also write the following formulas for the coefficients:

$$\begin{aligned} \gamma &= a^{1/2}g^{1/2}, \quad \lambda = \frac{1}{\gamma}\frac{1}{2+b+bg}, \quad \rho = \frac{1}{\gamma}\frac{g}{1+g}, \quad \mu = \frac{1+g}{2+b+bg}, \\ l &= \frac{1}{\gamma}\frac{1-g}{2+b+bg}, \quad 1-\mu-l\gamma = \frac{b+bg}{2+b+bg}. \end{aligned} \tag{A.46}$$

A.12 Proof of Proposition 10

See Lemma A.1.

A.13 Proof of Proposition 11

We compute $\mathrm{d}\Sigma_t^k = 2\,\mathsf{Cov}(\mathrm{d}v_t - \mathrm{d}p_t^k, v_t - p_t^k) + \mathsf{Cov}(\mathrm{d}v_t - \mathrm{d}p_t^k, \mathrm{d}v_t - \mathrm{d}p_t^k)$. Since the news $\mathrm{d}v_t$ is orthogonal to $v_t - p_t^k$ in both models, $\mathrm{d}\Sigma_t^k = -2\,\mathsf{Cov}(\mathrm{d}p_t^k, v_t - p_t^k) - 2\,\mathsf{Cov}(\mathrm{d}p_t^k, \mathrm{d}v_t) + \mathsf{Var}(\mathrm{d}v_t) + \mathsf{Var}(\mathrm{d}p_t^k)$. But $\frac{1}{\mathrm{d}t}\mathsf{Var}(\mathrm{d}v_t) = \sigma_v^2$; and by Proposition 12, $\sigma_p^2 = \frac{1}{\mathrm{d}t}\mathsf{Var}(\mathrm{d}p_t^k) = \sigma_v^2 + \Sigma_0$. We have just proved (36).

Equation (8) implies that in the benchmark model $\mathrm{d}p_t = \mu^B \mathrm{d}z_t + \lambda^B \mathrm{d}y_t$. Since $\mathrm{d}y_t = \mathrm{d}x_t + \mathrm{d}u_t$, with $\mathrm{d}x_t = O(\mathrm{d}t)$,

$$\mathsf{Cov}(\mathrm{d}p_t, \mathrm{d}v_t) = \mu^B \sigma_v^2 \mathrm{d}t. \tag{A.47}$$

Equation (9) implies that in the fast model, $\mathrm{d}q_t = \lambda^F \Delta y_t + \mu^F(\Delta z_t - \rho^F \Delta y_t)$. Since $\mathrm{d}x_t = \gamma^F\,\mathrm{d}v_t + O(\mathrm{d}t)$, we obtain

$$\mathsf{Cov}(\mathrm{d}p_t, \mathrm{d}v_t) = \Big(\gamma^F(\lambda^F - \mu^F\rho^F) + \mu^F\Big)\sigma_v^2 \mathrm{d}t. \tag{A.48}$$

Then, we prove that $\gamma^F(\lambda^F - \mu^F\rho^F) + \mu^F > \mu^B$. Using (A.46) and (A.32), we need to show that $\frac{2}{2+b+bg} > \frac{1}{1+b}$, which is equivalent to $1 > g$. But this is true, since $g \in (0,1)$.

Finally, from the proof of Theorem 1, in both models $\Sigma_t^k = \Sigma_0(1-t)$.

A.14 Proof of Proposition 12

Denote $\mathsf{Var}(\mathrm{d}p_t^k) = \sigma_p^2 \mathrm{d}t$ the variance of the instantaneous price changes, and we use Theorem 1 to compute the two components of this variance. In the benchmark model, $\mathsf{Var}(p_{t+\mathrm{d}t} - q_t) = (\lambda^B)^2\sigma_u^2 \mathrm{d}t = \left(\Sigma_0 + \frac{\sigma_v^2\sigma_e^2}{\sigma_v^2+\sigma_e^2}\right)\mathrm{d}t$. Also, $\mathsf{Var}(q_t - p_t) = (\mu^B)^2(\sigma_v^2 + \sigma_e^2)\mathrm{d}t =$

$\frac{\sigma_v^4}{\sigma_v^2+\sigma_e^2}\mathrm{d}t$. We obtain the volatility decomposition in the benchmark model,

$$\sigma_p^2 \; = \; \mathsf{Var}(\mathrm{d}p_t) \; = \; \left(\Sigma_0 + \frac{\sigma_v^2\sigma_e^2}{\sigma_v^2+\sigma_e^2}\right) \; + \; \frac{\sigma_v^4}{\sigma_v^2+\sigma_e^2} \; = \; \Sigma_0 + \sigma_v^2. \tag{A.49}$$

Similarly, in the fast model, $\mathsf{Var}(p_{t+\mathrm{d}t} - q_t) = (\lambda^F)^2\big((\gamma^F)^2\sigma_v^2 + \sigma_u^2\big)\mathrm{d}t$, and using equation (A.46) we compute $\mathsf{Var}(p_{t+\mathrm{d}t} - q_t) = \frac{1+g}{g(2+b+bg)^2}\sigma_v^2\,\mathrm{d}t$. Also, $\mathsf{Var}(q_{t+\mathrm{d}t} - p_{t+\mathrm{d}t}) = (\mu^F)^2\big((1-\rho^F\gamma^F)^2\sigma_v^2 + \sigma_e^2 + (\rho^F)^2\sigma_u^2\big)\mathrm{d}t = \frac{(1+g)(1+b+bg)}{(2+b+bg)^2}\sigma_v^2\,\mathrm{d}t$. According to (A.26), $\Sigma_0 + \sigma_v^2 = \sigma_v^2(1+c) = \sigma_v^2\frac{(1+g)^2(1+bg)}{g(2+b+bg)^2}$, thus we have

$$\sigma_p^2 \; = \; \mathsf{Var}(\mathrm{d}q_t) \; = \; \frac{1+g}{g(2+b+bg)^2}\sigma_v^2 \; + \; \frac{(1+g)(1+b+bg)}{(2+b+bg)^2}\sigma_v^2 \; = \; \Sigma_0 + \sigma_v^2. \tag{A.50}$$

We now show that the volatility component coming from quote updates is larger in the benchmark, i.e., $\frac{\sigma_v^2}{\sigma_v^2+\sigma_e^2} = \frac{1}{1+b} > \frac{(1+g)(1+b+bg)}{(2+b+bg)^2}$. The difference is proportional to $3 - g + 2b + bg - bg^2 = 2(1+b) + (1-g)(1+bg) > 0$. Since the total volatility is the same, it also implies that the volatility component coming from the trades is larger in the fast model.

B Sampling at Lower Frequency than Trading Frequency

In this section, we show that the results of Section 3.2 generalize when trades are aggregated over intervals of an arbitrary length Δs. Suppose trading takes place in continuous time, but trades are aggregated over $S > 0$ time intervals of equal length $\frac{1}{S} = \Delta s$. Then, data are indexed by $s = 0, 1, 2, \ldots, S-1$ and the informed order flow at s is $\Delta x_s = x_{s+1} - x_s$. The empirical counterpart of the Informed Participation Rate and the autocorrelation of the informed investor's order flow when data are sampled every Δs periods of time are, respectively,

$$\begin{aligned} &IPR \; = \; \frac{\mathsf{Var}(\Delta x_s)}{\mathsf{Var}(\Delta x_s) + \mathsf{Var}(\Delta u_s)}, \\ &\mathsf{Corr}(\Delta x_s, \Delta x_{s+k}). \end{aligned} \tag{B.1}$$

Proposition B.1. When the sampling interval Δs is small, the empirical informed participation rate in the benchmark increases with Δs and is always below its level in the fast model:

$$
\begin{aligned}
IPR^B &= \frac{(\beta_s^B)^2\Sigma_s}{\sigma_u^2}\Delta s + o(\Delta s), \\
IPR^F &= \frac{(\gamma^F)^2\sigma_v^2}{(\gamma^F)^2\sigma_v^2 + \sigma_u^2} + \frac{o(\Delta s)}{\Delta s}.
\end{aligned} \tag{B.2}
$$

The informed order flow autocorrelation in the fast model increases with the sampling interval Δs and is always below its level in the benchmark:

$$
\begin{aligned}
\mathsf{Corr}(\Delta x_s^B, \Delta x_{s+k}^B) &= \left(\frac{1-s-k\Delta s}{1-s}\right)^{\lambda^B\beta_0^B - \frac{1}{2}} + \frac{o(\Delta s)}{\Delta s}, \\
\mathsf{Corr}(\Delta x_s^F, \Delta x_{s+k}^F) &= \frac{\beta_{s+k}^F\Big(\beta_s^F\Sigma_s + \gamma^F\big(1 - l^F\gamma^F - \mu^F\big)\sigma_v^2\Big)}{(\gamma^F)^2\sigma_v^2} \\
&\quad \times \left(\frac{1-s-k\Delta s}{1-s}\right)^{l^F\beta_0^F} \Delta s + o(\Delta s).
\end{aligned} \tag{B.3}
$$

To define the empirical counterpart of our measure of anticipatory trading, we now consider that trading takes place in discrete time rather than in continuous time. As in (the future Internet) Appendix C, trading takes place over $T > 0$ intervals of equal length $\frac{1}{T} = \Delta t$. Time is indexed by $t = 0, 1, 2, \ldots, T-1$. We assume that Δt is small, and we approximate the equilibrium variables (β_t, γ_t, λ_t, μ_t, ρ_t) in this discrete time model by their continuous time counterpart. The informed trade at time $t+1$ is equal to $\Delta x_t = \beta_t(v_t - q_t)\Delta t + \gamma\Delta v_t$, where q_t is the quote just before the order flow arrives, and p_{t+1} is the execution price.

We consider that the econometrician has data sampled every n trading rounds, i.e., the sampling interval is $\Delta s = n\Delta t$. Therefore the data are a time series indexed by $s = 0, 1, 2, \ldots, \frac{T}{n} - 1$ (assume that T is a multiple of n). The sth observation corresponds to the aggregate order flow during the n trading rounds from $sn+1$ to $sn+n$,

$$
\Delta x_s(n) \;=\; \Delta x_{sn} + \ldots + \Delta x_{sn+n-1}, \tag{B.4}
$$

and the return over this period is

$$
r_s(n) \;=\; p_{sn+n} - p_{sn}. \tag{B.5}
$$

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

$$AT_s(n) \; = \; \mathsf{Corr}\big(\Delta x_s(n), r_{s+1}(n)\big). \tag{B.6}$$

The next result shows that that sampling data at a sufficiently high frequency (i.e., low Δs, or low n) is important for detecting anticipatory trading.

Proposition B.2. In the fast model, when Δt is small, the empirical measure of anticipatory trading is decreasing in $n = \frac{\Delta s}{\Delta t}$, and converges to zero when $n \to +\infty$,

$$AT_s^F(n) \; = \; \frac{\mu^F(1-\rho^F\gamma^F)\sigma_v}{\sqrt{\sigma_v^2+\Sigma_0}} \, \frac{1}{n}. \tag{B.7}$$

The aggregated order flow spans $n = \frac{\Delta s}{\Delta t}$ 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 $\Delta x_s(n)$ is correlated with the next aggregated return $r_{s+1}(n)$. As a result, when n increases, the correlation between $\Delta x_s(n)$ and $r_{s+1}(n)$ decreases. When n becomes too large, the correlation becomes almost zero.

B.1 Proof of Proposition B.1

The sth trade in the data is $\Delta x_s = \int_s^{s+\Delta s} \beta_t(v_t - p_t)\,\mathrm{d}t + \gamma\,\mathrm{d}v_t$. When Δs is small, β_t is approximately constant over the interval $[s, s+\Delta s]$. Thus, in the benchmark model we have $\Delta x_s^B = \beta_s^B(v_s - p_s)\Delta s + o(\Delta s)$, since $\gamma^B = 0$. This implies $\mathsf{Var}(\Delta x_s^B) = (\beta_s^B)^2\Sigma_s(\Delta s)^2 + o\big((\Delta s)^2\big)$. Also, $\mathsf{Var}(\Delta u_s) = \sigma_u^2\Delta s$, which yields the informed participation rate in (B.2). Using Lemma A.2 in Appendix A, we obtain $\mathsf{Cov}(\Delta x_s^B, \Delta x_{s+k}^B) = \beta_{s+k}^B\beta_s^B\Sigma_s\left(\frac{1-s-k\Delta s}{1-s}\right)^{\lambda^B\beta_0^B}(\Delta s)^2 + o\big((\Delta s)^2\big)$, which proves the first equation in (B.3).

In the fast model, $\Delta x_s^F = \beta_s^F(v_s - p_s)\Delta s + \gamma^F\Delta v_s + o(\Delta s)$. Then, $\mathsf{Var}(\Delta x_s^F) = (\gamma^F)^2\sigma_v^2\Delta s + o(\Delta s)$, which implies the informed participation rate in (B.2). Using Lemma A.2, $\mathsf{Cov}(\Delta x_s^F, \Delta x_{s+k}^F) = \beta_{s+k}^F\Big(\beta_s^F\Sigma_s + \gamma^F\big(1 - l^F\gamma^F - \mu^F\big)\sigma_v^2\Big)\left(\frac{1-s-k\Delta s}{1-s}\right)^{l^F\beta_0^F}(\Delta s)^2 + o((\Delta s)^2)$, where $l^F = \lambda^F - \mu^F\rho^F$, which yields the second equation in (B.3).

B.2 Proof of Proposition B.2

We have $\mathsf{Cov}(\Delta x_s(n), r_{s+1}(n)) = \gamma^F\mu^F(1-\rho^F\gamma^F)\sigma_v^2\Delta t + o(\Delta t)$. Also, $\mathsf{Var}(\Delta x_s(n)) = (\gamma^F)^2\sigma_v^2\Delta s + o(\Delta s)$, and $\mathsf{Var}(r_{s+1}) = (\sigma_v^2 + \Sigma_0)\Delta s + o(\Delta s)$, where we use Proposition 12 for the last equation. These prove (B.7).

C Internet Appendix

In this Appendix, we present the discrete time version of the fast and benchmark models in subsections C.1 and C.2.

C.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:

$$
\begin{aligned}
p_t &= \mathsf{E}(v_t \mid \mathcal{I}_t^p), \qquad & \mathcal{I}_t^p &= \{\Delta y_1, \ldots, \Delta y_t, \Delta z_1, \ldots, \Delta z_{t-1}\}, \\
q_{t+1} &= \mathsf{E}(v_t \mid \mathcal{I}_t^q), \qquad & \mathcal{I}_t^q &= \{\Delta y_1, \ldots, \Delta y_t, \Delta z_1, \ldots, \Delta z_t\}.
\end{aligned}
\tag{C.1}
$$

We also denote

$$
\begin{aligned}
\Omega_t &= \mathsf{Var}(v_t \mid \mathcal{I}_t^p), \\
\Sigma_t &= \mathsf{Var}(v_t \mid \mathcal{I}_t^q).
\end{aligned}
\tag{C.2}
$$

Definition C.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 $t = 1, \ldots, T-1$.[38]

The next result shows that if the pricing rule is linear, then the informed trader's strategy is also linear, and furthermore it can be decomposed into a forecast error component, $\beta_t(v_t - q_t)\Delta t$, and a news trading component, $\tilde{\gamma}_t \Delta v_t$, where $\tilde{\gamma}_t \equiv \gamma_t - \beta_t \Delta t = \frac{\alpha_t l_t \mu_t}{\lambda_t - \alpha_t l_t^2}$ (see (C.6)).

[38] 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.

Theorem C.1. Any equilibrium with a linear pricing rule must be of the form

$$\begin{aligned} \Delta x_t &= \beta_t(v_{t-1}-q_t)\Delta t+\gamma_t\Delta v_t, \\ p_t &= q_t+\lambda_t\Delta y_t, \\ q_{t+1} &= p_t+\mu_t(\Delta z_t-\rho_t\Delta y_t), \end{aligned} \tag{C.3}$$

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

$$\begin{aligned} \lambda_t &= \frac{\beta_t\Sigma_{t-1}+\gamma_t\sigma_v^2}{\beta_t^2\Sigma_{t-1}\Delta t+\gamma_t^2\sigma_v^2+\sigma_u^2}, \\ \mu_t &= \frac{\left(\sigma_u^2+\beta_t^2\Sigma_{t-1}\Delta t-\beta_t\gamma_t\Sigma_{t-1}\right)\sigma_v^2}{\left(\beta_t^2\Sigma_{t-1}\Delta t+\gamma_t^2\sigma_v^2+\sigma_u^2\right)\sigma_e^2+\left(\beta_t^2\Sigma_{t-1}\Delta t+\sigma_u^2\right)\sigma_v^2}, \\ l_t &= \lambda_t-\rho_t\mu_t \;=\; \frac{\beta_t\Sigma_{t-1}(\sigma_v^2+\sigma_e^2)+\gamma_t\sigma_v^2\sigma_e^2}{\left(\beta_t^2\Sigma_{t-1}\Delta t+\gamma_t^2\sigma_v^2+\sigma_u^2\right)\sigma_e^2+\left(\beta_t^2\Sigma_{t-1}\Delta t+\sigma_u^2\right)\sigma_v^2}, \\ \rho_t &= \frac{\gamma_t\sigma_v^2}{\beta_t^2\Sigma_{t-1}\Delta t+\gamma_t^2\sigma_v^2+\sigma_u^2}, \\ \Omega_t &= \Sigma_{t-1}+\sigma_v^2\Delta t-\frac{\beta_t^2\Sigma_{t-1}^2+2\beta_t\gamma_t\Sigma_{t-1}\sigma_v^2+\gamma_t^2\sigma_v^4}{\beta_t^2\Sigma_{t-1}\Delta t+\gamma_t^2\sigma_v^2+\sigma_u^2}\,\Delta t, \\ \Sigma_t &= \Sigma_{t-1}+\sigma_v^2\Delta t \\ &\quad -\frac{\beta_t^2\Sigma_{t-1}^2(\sigma_v^2+\sigma_e^2)+\beta_t^2\Sigma_{t-1}\Delta t\sigma_v^4+\sigma_v^4\sigma_u^2+\gamma_t^2\sigma_v^4\sigma_e^2+2\beta_t\gamma_t\Sigma_{t-1}\sigma_v^2\sigma_e^2}{\left(\beta_t^2\Sigma_{t-1}\Delta t+\gamma_t^2\sigma_v^2+\sigma_u^2\right)\sigma_e^2+\left(\beta_t^2\Sigma_{t-1}\Delta t+\sigma_u^2\right)\sigma_v^2}\,\Delta t. \end{aligned} \tag{C.4}$$

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

$$\pi_t \;=\; \alpha_{t-1}(v_{t-1}-q_t)^2+\alpha'_{t-1}(\Delta v_t)^2+\alpha''_{t-1}(v_{t-1}-q_t)\Delta v_t+\delta_{t-1}. \tag{C.5}$$

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

$$\begin{aligned} \beta_t\Delta t &= \frac{1-2\alpha_t l_t}{2(\lambda_t-\alpha_t l_t^2)}, \\ \gamma_t &= \frac{1-2\alpha_t l_t(1-\mu_t)}{2(\lambda_t-\alpha_t l_t^2)} \;=\; \beta_t\Delta t+\frac{\alpha_t l_t\mu_t}{\lambda_t-\alpha_t l_t^2}, \\ \alpha_{t-1} &= \beta_t\Delta t(1-\lambda_t\beta_t\Delta t)+\alpha_t(1-l_t\beta_t\Delta t)^2, \\ \alpha'_{t-1} &= \alpha_t(1-\mu_t-l_t\gamma_t)^2+\gamma_t(1-\lambda_t\gamma_t), \\ \alpha''_{t-1} &= \beta_t\Delta t+\gamma_t(1-2\lambda_t\beta_t\Delta t)+2\alpha_t(1-l_t\beta_t\Delta t)(1-\mu_t-l_t\gamma_t), \\ \delta_{t-1} &= \alpha_t\left(l_t^2\sigma_u^2+\mu_t^2\sigma_e^2\right)\Delta t+\alpha'_t\sigma_v^2\Delta t+\delta_t. \end{aligned} \tag{C.6}$$

The terminal conditions are

$$\alpha_T \ = \ \alpha'_T \ = \ \alpha''_T \ = \ \delta_T \ = \ 0. \tag{C.7}$$

The second order condition is

$$\lambda_t - \alpha_t l_t^2 \ > \ 0. \tag{C.8}$$

Given Σ_0, conditions (C.4)–(C.8) are necessary and sufficient for the existence of a linear equilibrium.

Proof. First, we show that equations (C.4) are equivalent to the zero profit conditions of the market maker. Second, we show that equations (C.6)–(C.8) are equivalent to the informed trader's strategy in (C.3) being optimal. These two steps prove that equations (C.3)–(C.8) describe an equilibrium. Conversely, all equilibria with a linear pricing rule must satisfy these equations since the trading strategy in (C.3) is the best-response to the linear pricing rule.

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:

$$\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \mid \mathcal{I}^q_{t-1} \ \sim \ \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{t-1} & 0 \\ 0 & \sigma_v^2 \end{bmatrix} \right). \tag{C.9}$$

The aggregate order flow at t is of the form

$$\Delta y_t \ = \ \beta_t (v_{t-1} - q_t)\Delta t + \gamma_t \Delta v_t + \Delta u_t. \tag{C.10}$$

Denote by

$$\Phi_t \ = \ \mathsf{Cov}\left(\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix}, \Delta y_t \right) \ = \ \begin{bmatrix} \beta_t \Sigma_{t-1} \\ \gamma_t \sigma_v^2 \end{bmatrix} \Delta t. \tag{C.11}$$

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:

$$\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \mid \mathcal{I}_t \ \sim \ \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right), \tag{C.12}$$

where

$$\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \Phi_t \, \mathsf{Var}(\Delta y_t)^{-1} \Delta y_t = \begin{bmatrix} \beta_t \Sigma_{t-1} \\ \gamma_t \sigma_v^2 \end{bmatrix} \frac{1}{\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2} \, \Delta y_t, \tag{C.13}$$

and

$$\begin{aligned} \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} &= \mathsf{Var}\left(\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \right) - \Phi_t \, \mathsf{Var}(\Delta y_t)^{-1} \Phi_t' \\ = \begin{bmatrix} \Sigma_{t-1} & 0 \\ 0 & \sigma_v^2 \Delta t \end{bmatrix} &- \frac{1}{\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2} \begin{bmatrix} \beta_t^2 \Sigma_{t-1}^2 & \beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 \\ \beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 & \gamma_t^2 \sigma_v^4 \end{bmatrix} \Delta t. \end{aligned} \tag{C.14}$$

We compute

$$p_t - q_t = \mathsf{E}(v_t - q_t \mid \mathcal{I}_t) = \mu_1 + \mu_2 = \frac{\beta_t \Sigma_{t-1} + \gamma_t \sigma_v^2}{\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2} \, \Delta y_t, \tag{C.15}$$

which proves equation (C.4) for λ_t. Also,

$$\begin{aligned} \Omega_t &= \mathsf{Var}(v_t - q_t \mid \mathcal{I}_t) = \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2 \\ &= \Sigma_{t-1} + \sigma_v^2 \Delta t - \frac{\beta_t^2 \Sigma_{t-1}^2 + 2\beta_t \gamma_t \Sigma_{t-1} \sigma_v^2 + \gamma_t^2 \sigma_v^4}{\beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2} \, \Delta t, \end{aligned} \tag{C.16}$$

which proves the formula for Ω_t.

Next, to compute $q_{t+1} = \mathsf{E}(v_t \mid \mathcal{I}_t^q)$, we start from the same prior as in (C.9), but we consider the impact of both the order flow at t and the market maker's signal at $t+1$:

$$\begin{aligned} \Delta y_t &= \beta_t (v_{t-1} - q_t) \Delta t + \gamma_t \Delta v_t + \Delta u_t, \\ \Delta z_t &= \Delta v_t + \Delta e_t. \end{aligned} \tag{C.17}$$

Denote by

$$\begin{aligned} \Psi_t &= \mathsf{Cov}\left(\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix}, \begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix} \right) = \begin{bmatrix} \beta_t \Sigma_{t-1} & 0 \\ \gamma_t \sigma_v^2 & \sigma_v^2 \end{bmatrix} \Delta t, \\ V_t^{yz} &= \mathsf{Var}\left(\begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix} \right) = \begin{bmatrix} \beta_t^2 \Sigma_{t-1} \Delta t + \gamma_t^2 \sigma_v^2 + \sigma_u^2 & \gamma_t \sigma_v^2 \\ \gamma_t \sigma_v^2 & \sigma_v^2 + \sigma_e^2 \end{bmatrix} \Delta t. \end{aligned} \tag{C.18}$$

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:

$$\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \mid \mathcal{I}_t^q \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right), \tag{C.19}$$

where

$$\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \Psi_t \, (V_t^{yz})^{-1} \begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix} = \frac{\begin{bmatrix} \beta_t \Sigma_{t-1}(\sigma_v^2 + \sigma_e^2)\Delta y_t \; - \; \beta_t\gamma_t\Sigma_{t-1}\sigma_v^2\Delta z_t \\ \gamma_t\sigma_v^2\sigma_e^2\Delta y_t \; + \; (\beta_t^2\Sigma_{t-1}\Delta t + \sigma_u^2)\sigma_v^2\Delta z_t \end{bmatrix}}{(\beta_t^2\Sigma_{t-1}\Delta t + \gamma_t^2\sigma_v^2 + \sigma_u^2)\sigma_e^2 + (\beta_t^2\Sigma_{t-1}\Delta t + \sigma_u^2)\sigma_v^2}, \tag{C.20}$$

and

$$\begin{aligned} \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} &= \mathsf{Var}\left(\begin{bmatrix} v_{t-1} - q_t \\ \Delta v_t \end{bmatrix} \right) - \Psi_t \, (V_t^{yz})^{-1} \, \Psi_t' \\ &= \begin{bmatrix} \Sigma_{t-1} & 0 \\ 0 & \sigma_v^2 \Delta t \end{bmatrix} - \frac{\begin{bmatrix} \beta_t^2\Sigma_{t-1}^2(\sigma_v^2 + \sigma_e^2) & \beta_t\gamma_t\Sigma_{t-1}\sigma_v^2\sigma_e^2 \\ \beta_t\gamma_t\Sigma_{t-1}\sigma_v^2\sigma_e^2 & (\beta_t^2\Sigma_{t-1}\Delta t + \gamma_t^2\sigma_e^2 + \sigma_u^2)\sigma_v^4 \end{bmatrix}}{(\beta_t^2\Sigma_{t-1}\Delta t + \gamma_t^2\sigma_v^2 + \sigma_u^2)\sigma_e^2 + (\beta_t^2\Sigma_{t-1}\Delta t + \sigma_u^2)\sigma_v^2} \, \Delta t. \end{aligned} \tag{C.21}$$

Therefore,

$$\begin{aligned} q_{t+1} - q_t &= \mu_1 + \mu_2 \\ &= \frac{\big(\beta_t\Sigma_{t-1}(\sigma_v^2 + \sigma_e^2) + \gamma_t\sigma_v^2\sigma_e^2\big)\Delta y_t \; + \; \big(\sigma_u^2 + \beta_t^2\Sigma_{t-1}\Delta t - \beta_t\gamma_t\Sigma_{t-1}\big)\sigma_v^2\Delta z_t}{\big(\beta_t^2\Sigma_{t-1}\Delta t + \gamma_t^2\sigma_v^2 + \sigma_u^2\big)\,\sigma_e^2 + \big(\beta_t^2\Sigma_{t-1}\Delta t + \sigma_u^2\big)\,\sigma_v^2} \end{aligned} \tag{C.22}$$

$$= l_t\Delta y_t \; + \; \mu_t\Delta z_t \; = \; (\lambda_t - \rho_t\mu_t)\Delta y_t \; + \; \mu_t\Delta z_t, \tag{C.23}$$

which proves equation (C.4) for μ_t, l_t, and ρ_t. Also,

$$\begin{aligned} \Sigma_t &= \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2 \\ &= \Sigma_{t-1} + \sigma_v^2\Delta t - \frac{\beta_t^2\Sigma_{t-1}^2(\sigma_v^2 + \sigma_e^2) + \beta_t^2\Sigma_{t-1}\Delta t\sigma_v^4 + \sigma_v^4\sigma_u^2 + \gamma_t^2\sigma_v^4\sigma_e^2 + 2\beta_t\gamma_t\Sigma_{t-1}\sigma_v^2\sigma_e^2}{\big(\beta_t^2\Sigma_{t-1} + (\beta_t + \gamma_t)^2\sigma_v^2 + \sigma_u^2\big)\,\sigma_e^2 + \big(\beta_t^2\Sigma_{t-1} + \sigma_u^2\big)\,\sigma_v^2} \, \Delta t, \end{aligned} \tag{C.24}$$

which proves the formula for Σ_t.

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 (C.5): $\pi_t = \alpha_{t-1}(v_{t-1} - q_t)^2 + \alpha'_{t-1}(\Delta v_t)^2 + \alpha''_{t-1}(v_{t-1} - q_t)\Delta v_t + \delta_{t-1}$. 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 (C.7). 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

$$\pi_t \;=\; \max_{\Delta x}\; \mathsf{E}\Big((v_t - p_t)\Delta x + \pi_{t+1} \mid \mathcal{I}^q_t, v_t, \Delta v_t\Big). \tag{C.25}$$

The last two equations in (C.3) imply that the quote q_t evolves by $q_{t+1} = q_t + l_t\Delta y_t + \mu_t\Delta z_t$, where $l_t = \lambda_t - \rho_t\mu_t$. This implies that the informed trader's choice of Δx affects the trading price and the next quote by

$$\begin{aligned} p_t &= q_t + \lambda_t(\Delta x + \Delta u_t), \\ q_{t+1} &= q_t + l_t(\Delta x + \Delta u_t) + \mu_t\Delta z_t. \end{aligned} \tag{C.26}$$

Substituting these into the Bellman equation, we get

$$\begin{aligned} \pi_t \;=\; & \max_{\Delta x}\; \mathsf{E}\Big(\Delta x(v_{t-1} + \Delta v_t - q_t - \lambda_t\Delta x - \lambda_t\Delta u_t) \\ & + \alpha_t(v_{t-1} + \Delta v_t - q_t - l_t\Delta x - l_t\Delta u_t - \mu_t\Delta v_t - \mu_t\Delta e_t)^2 \;+\; \alpha'_t\Delta v_{t+1}^2 \\ & + \alpha''_t(v_{t-1} + \Delta v_t - q_t - l_t\Delta x - l_t\Delta u_t - \mu_t\Delta v_t - \mu_t\Delta e_t)\Delta v_{t+1} \;+\; \delta_t\Big) \\ =\; & \max_{\Delta x}\; \Delta x(v_{t-1} - q_t + \Delta v_t - \lambda_t\Delta x) \\ & + \alpha_t\Big((v_{t-1} - q_t - l_t\Delta x + (1-\mu_t)\Delta v_t)^2 + (l_t^2\sigma_u^2 + \mu_t^2\sigma_e^2)\Delta t\Big) \;+\; \alpha'_t\sigma_v^2\Delta t \\ & +\; 0 \;+\; \delta_t. \end{aligned} \tag{C.27}$$

The first order condition with respect to Δx is

$$\Delta x \;=\; \frac{1 - 2\alpha_t l_t}{2(\lambda_t - \alpha_t l_t^2)}(v_{t-1} - q_t) \;+\; \frac{1 - 2\alpha_t l_t(1-\mu_t)}{2(\lambda_t - \alpha_t l_t^2)}\Delta v_t, \tag{C.28}$$

and the second order condition for a maximum is $\lambda_t - \alpha_t l_t^2 > 0$, which is (C.8). 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 (C.6). We substitute Δx_t in the formula for π_t to obtain

$$\begin{aligned} \pi_t \;=\; & \Big(\beta_t\Delta t(1-\lambda_t\beta_t\Delta t) + \alpha_t(1-l_t\beta_t\Delta t)^2\Big)(v_{t-1}-q_t)^2 \\ & + \Big(\alpha_t(1-\mu_t-l_t\gamma_t)^2 + \gamma_t(1-\lambda_t\gamma_t)\Big)\Delta v_t^2 \\ & + \Big(\beta_t\Delta t + \gamma_t(1-2\lambda_t\beta_t\Delta t) + 2\alpha_t(1-l_t\beta_t\Delta t)(1-\mu_t-l_t\gamma_t)\Big)(v_{t-1}-q_t)\Delta v_t \\ & + \alpha_t\big(l_t^2\sigma_u^2 + \mu_t^2\sigma_e^2\big)\Delta t \;+\; \alpha_t'\sigma_v^2\Delta t \;+\; \delta_t. \end{aligned} \tag{C.29}$$

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 (C.6). □

We now briefly discuss the existence of a solution for the recursive system given in Theorem C.1. The system of equations (C.4)–(C.6) can be numerically solved backwards, starting from the boundary conditions (C.7). We also start with an arbitrary value of $\Sigma_T > 0$.[39] By backward induction, suppose α_t and Σ_t are given. One verifies that equation (C.4) for Σ_t implies

$$\Sigma_{t-1} \;=\; \frac{\Sigma_t\big(\sigma_v^2\sigma_u^2 + \sigma_v^2(\sigma_u^2+\gamma_t^2\sigma_e^2)\big) - \sigma_v^2\sigma_u^2\sigma_e^2\Delta t}{\big(\sigma_u^2\sigma_e^2 + \sigma_v^2(\sigma_u^2+\gamma_t^2\sigma_e^2) + \beta_t^2\Delta t^2\sigma_v^2\sigma_e^2 - 2\gamma_t\beta_t\Delta t\sigma_v^2\sigma_e^2\big) - \Sigma_t\beta_t^2\Delta t\big(\sigma_v^2+\sigma_e^2\big)}. \tag{C.30}$$

Then, in equation (C.4) we can rewrite λ_t, μ_t, l_t as functions of $(\Sigma_t, \beta_t, \gamma_t)$ instead of $(\Sigma_{t-1}, \beta_t, \gamma_t)$. Next, we use the formulas for β_t and γ_t to express λ_t, μ_t, l_t as functions of $(\lambda_t, \mu_t, l_t, \alpha_t, \Sigma_t)$. This gives a system of polynomial equations, whose solution λ_t, μ_t, l_t depends only on (α_t, Σ_t). Numerical simulations show that the solution is unique under the second order condition (C.8), 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.

C.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

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

is as follows. At $t = 0$, the informed trader observes v_0. At each $t = 1, \dots, T-1$, 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

$$\begin{aligned} p_t &= \mathsf{E}(v_t \mid \mathcal{I}_t^p), \qquad \mathcal{I}_t^p = \{\Delta y_1, \dots, \Delta y_t, \Delta z_1, \dots, \Delta z_t\}, \\ q_t &= \mathsf{E}(v_t \mid \mathcal{I}_t^q), \qquad \mathcal{I}_t^q = \{\Delta y_1, \dots, \Delta y_{t-1}, \Delta z_1, \dots, \Delta z_t\}. \end{aligned} \tag{C.31}$$

We also denote

$$\begin{aligned} \Sigma_t &= \mathsf{Var}(v_t \mid \mathcal{I}_t^p), \\ \Omega_t &= \mathsf{Var}(v_t \mid \mathcal{I}_t^q). \end{aligned} \tag{C.32}$$

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 forecast error component, $\beta_t(v_t - q_t)$.

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

$$\begin{aligned} \Delta x_t &= \beta_t (v_t - q_t) \Delta t, \\ p_t &= q_t + \lambda_t \Delta y_t, \\ q_t &= p_{t-1} + \mu_{t-1} \Delta z_t, \end{aligned} \tag{C.33}$$

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

$$\begin{aligned} \lambda_t &= \frac{\beta_t \Sigma_t}{\sigma_u^2}, \\ \mu_t &= \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2}, \\ \Omega_t &= \frac{\Sigma_t \sigma_u^2}{\sigma_u^2 - \beta_t^2 \Sigma_t \Delta t}, \\ \Sigma_{t-1} &= \Sigma_t + \frac{\beta_t^2 \Sigma_t^2}{\sigma_u^2 - \beta_t^2 \Sigma_t \Delta t} \Delta t - \frac{\sigma_v^2 \sigma_e^2}{\sigma_v^2 + \sigma_e^2} \Delta t. \end{aligned} \tag{C.34}$$

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

$$\pi_t \;=\; \alpha_{t-1}(v_t - q_t)^2 + \delta_{t-1}. \tag{C.35}$$

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

$$\begin{aligned}
\beta_t \Delta t \;&=\; \frac{1 - 2\alpha_t\lambda_t}{2\lambda_t(1 - \alpha_t\lambda_t)}, \\
\alpha_{t-1} \;&=\; \beta_t\Delta t(1 - \lambda_t\beta_t\Delta t) + \alpha_t(1 - \lambda_t\beta_t\Delta t)^2, \\
\delta_{t-1} \;&=\; \alpha_t\big(\lambda_t^2\sigma_u^2 + \mu_t^2(\sigma_v^2 + \sigma_e^2)\big)\Delta t + \delta_t.
\end{aligned} \tag{C.36}$$

The terminal conditions are

$$\alpha_T \;=\; \delta_T \;=\; 0. \tag{C.37}$$

The second order condition is

$$\lambda_t(1 - \alpha_t\lambda_t) \;>\; 0. \tag{C.38}$$

Given Σ_0, conditions (C.34)–(C.38) are necessary and sufficient for the existence of a linear equilibrium.

Proof. First, we show that equations (C.34) are equivalent to the zero profit conditions of the market maker. Second, we show that equations (C.36)–(C.38) are equivalent to the informed trader's strategy 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

$$\Phi_t \;=\; \mathsf{Cov}(v_t - q_t, \Delta y_t) \;=\; \beta_t\Omega_t\Delta t. \tag{C.39}$$

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

$$\begin{aligned}
p_t \;&=\; q_t + \lambda_t\Delta y_t, \\
\lambda_t \;&=\; \Phi_t\,\mathsf{Var}(\Delta y_t)^{-1} \;=\; \frac{\beta_t\Omega_t}{\beta_t^2\Omega_t\Delta t + \sigma_u^2}, \\
\Sigma_t \;&=\; \mathsf{Var}(v_t - q_t) - \Phi_t\,\mathsf{Var}(\Delta y_t)^{-1}\Phi_t' \;=\; \Omega_t - \frac{\beta_t^2\Omega_t^2}{\beta_t^2\Omega_t\Delta t + \sigma_u^2}\,\Delta t \\
&=\; \frac{\Omega_t\sigma_u^2}{\beta_t^2\Omega_t\Delta t + \sigma_u^2}.
\end{aligned} \tag{C.40}$$

To obtain the equation for λ_t, note that the above equations for λ_t and Σ_t imply $\frac{\lambda_t}{\Sigma_t} = \frac{\beta_t}{\sigma_u^2}$. The equation for Ω_t is obtained by solving for Σ_t in the last equation of (C.40).

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

$$\Psi_t \;=\; \mathsf{Cov}(v_t - p_{t-1}, \Delta z_t) \;=\; \sigma_v^2 \Delta t. \tag{C.41}$$

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

$$\begin{aligned}
q_t &= p_{t-1} + \mu_t \Delta z_t, \\
\mu_t &= \Psi_t \,\mathsf{Var}(\Delta z_t)^{-1} \;=\; \frac{\sigma_v^2}{\sigma_v^2 + \sigma_e^2}, \\
\Omega_t &= \mathsf{Var}(v_t - p_{t-1}) - \Psi_t \,\mathsf{Var}(\Delta z_t)^{-1} \Psi_t' \;=\; \Sigma_{t-1} + \sigma_v^2 \Delta t - \frac{\sigma_v^4}{\sigma_v^2 + \sigma_e^2}\,\Delta t \\
&= \Sigma_{t-1} + \frac{\sigma_v^2 \sigma_e^2}{\sigma_v^2 + \sigma_e^2}\,\Delta t.
\end{aligned} \tag{C.42}$$

Thus, we prove the equation for μ_t. Note that equation (C.42) gives a formula for Σ_{t-1} as a function of Ω_t, and we already proved the formula for Ω_t as a function of Σ_t in (C.34). We therefore get Σ_{t-1} as a function of Σ_t, which is the last equation in (C.34).

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 (C.35): $\pi_t = \alpha_{t-1}(v_t - q_t)^2 + \delta_{t-1}$. 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 (C.37). 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

$$\pi_t \;=\; \max_{\Delta x} \;\mathsf{E}\Big((v_t - p_t)\Delta x + \pi_{t+1} \mid \mathcal{I}_t^q, v_t, \Delta v_t\Big). \tag{C.43}$$

The last two equations in (C.33) show that the quote q_t evolves by $q_{t+1} = q_t + l_t \Delta y_t + \mu_t \Delta z_{t+1}$. This implies that the informed trader's choice of Δx affects the trading price and the next quote by

$$\begin{aligned} p_t &= q_t + \lambda_t(\Delta x + \Delta u_t), \\ q_{t+1} &= q_t + \lambda_t(\Delta x + \Delta u_t) + \mu_t \Delta z_{t+1}. \end{aligned} \tag{C.44}$$

Substituting these into the Bellman equation, we get

$$\begin{aligned} \pi_t &= \max_{\Delta x} \; \mathsf{E}\Big(\Delta x(v_t - q_t - \lambda_t \Delta x - \lambda_t \Delta u_t) \\ &\quad + \alpha_t(v_t + \Delta v_{t+1} - q_t - \lambda_t \Delta x - \lambda_t \Delta u_t - \mu_t \Delta z_{t+1})^2 \;+\; \delta_t\Big) \\ &= \max_{\Delta x} \; \Delta x(v_t - q_t - \lambda_t \Delta x) \\ &\quad + \alpha_t\Big((v_t - q_t - \lambda_t \Delta x)^2 + (\lambda_t^2 \sigma_u^2 + \mu_t^2(\sigma_v^2 + \sigma_e^2))\Delta t\Big) \;+\; \delta_t. \end{aligned} \tag{C.45}$$

The first order condition with respect to Δx is

$$\Delta x \;=\; \frac{1 - 2\alpha_t \lambda_t}{2\lambda_t(1 - \alpha_t \lambda_t)}(v_t - q_t), \tag{C.46}$$

and the second order condition for a maximum is $\lambda_t(1 - \alpha_t \lambda_t) > 0$, which is (C.38). 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 (C.36). We substitute Δx_t in the formula for π_t to obtain

$$\pi_t \;=\; \Big(\beta_t \Delta t(1 - \lambda_t \beta_t \Delta t) + \alpha_t(1 - \lambda_t \beta_t \Delta t)^2\Big)(v_t - q_t)^2 \;+\; \alpha_t\big(\lambda_t^2 \sigma_u^2 + \mu_t^2(\sigma_v^2 + \sigma_e^2)\big)\Delta t \;+\; \delta_t. \tag{C.47}$$

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 (C.36). □

Equations (C.34) and (C.36) form a system of equations. As before, it is solved backwards, starting from the boundary conditions (C.37), and so that $\Sigma_t = \Sigma_0$ at $t = 0$.

References

[1] Admati, Admati, and Paul Pfleiderer (1988): "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies*, 1, 3–40.

[2] Back, Kerry (1992): "Insider Trading in Continuous Time," *Review of Financial Studies*, 5, 387–409.

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

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

[5] Baron, Matthew, Jonathan Brogaard, and Andrei Kirilenko (2012): "The Trading Profits of High Frequency Traders," *Working paper*, University of Washington.

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

[7] Boulatov, Alex, Terrence Hendershott, and Dmitry Livdan (2012): "Informed Trading and Portfolio Returns," forthcoming *Review of Economic Studies*.

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

[9] Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (2012): "High Frequency Trading and Price Discovery," Working Paper.

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

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

[12] 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.

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

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

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

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

[17] Gross-Klussmann, 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.

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

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

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

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

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

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

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

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

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

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

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

[29] 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.

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

[31] Martinez, Victor H., and Ioanid Roşu (2012): "High Frequency Traders, News and Volatillity," Working Paper.

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

[33] Menkveld, Albert, Siem Koopman, and André Lucas (2007): "Modeling Around-the-Clock Price Discovery for Cross-Listed Stocks Using State Space Methods," *Journal of Business and Economics Statistics*, 25, 213–225.

[34] Patell, James, and Mark Wolfson (1984): "The intraday speed of adjustment of stock prices to earnings and dividend announcements," *Journal of Financial Economics*, 13, 223–252.

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

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

[37] Scholtus, Martin, van Dijk Dick, and Bart Frijns (2012): "Speed, Algorithmic Trading, and Market Quality around Macroeconomic News Announcements," Working paper, Erasmus University Rotterdam.

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

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

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