CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

No. 569

Torben G. Andersen, Gökhan Cebiroglu, and Nikolaus Hautsch

Volatility, Information Feedback and Market Microstructure Noise: A Tale of Two Regimes

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

The CFS Working Paper Series

presents ongoing research on selected topics in the fields of money, banking and finance. The papers are circulated to encourage discussion and comment. Any opinions expressed in CFS Working Papers are those of the author(s) and not of the CFS.

The Center for Financial Studies, located in Goethe University Frankfurt's House of Finance, conducts independent and internationally oriented research in important areas of Finance. It serves as a forum for dialogue between academia, policy-making institutions and the financial industry. It offers a platform for top-level fundamental research as well as applied research relevant for the financial sector in Europe. CFS is funded by the non-profit-organization Gesellschaft für Kapitalmarktforschung e.V. (GfK). Established in 1967 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia. GfK members comprise major players in Germany's financial industry. The funding institutions do not give prior review to CFS publications, nor do they necessarily share the views expressed therein.

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

Volatility, Information Feedback and Market Microstructure Noise: A Tale of Two Regimes*

Torben G. Andersen† Gökhan Cebiroglu‡ Nikolaus Hautsch§

This Version: February 6, 2017

Abstract

We extend the classical "martingale-plus-noise" model for high-frequency prices by an error correction mechanism originating from prevailing mispricing. The speed of price reversal is a natural measure for informational efficiency. The strength of the price reversal relative to the signal-to-noise ratio determines the signs of the return serial correlation and the bias in standard realized variance estimates. We derive the model's properties and locally estimate it based on mid-quote returns of the NASDAQ 100 constituents. There is evidence of mildly persistent local regimes of positive and negative serial correlation, arising from lagged feedback effects and sluggish price adjustments. The model performance is decidedly superior to existing stylized microstructure models. Finally, we document intraday periodicities in the speed of price reversion and noise-to-signal ratios.

Keywords: volatility estimation; market microstructure noise; price reversal; momentum trading, contrarian trading

JEL classification: C58, C32, G14

*For helpful comments and discussions we thank the participants of the "Jan Mossin Memorial Symposium on Financial Markets", Norwegian School of Economics, Bergen, 2016, the Guangzhou 2016 Symposium on Financial Engineering and Risk Management, the 9th Annual Meeting of the Society for Financial Econometrics, University of Aarhus, 2016, the Conference on "New Developments in Measuring and Forecasting Financial Volatility", Durham, USA, 2016, and of research seminars at Humboldt University, Berlin, and Universitat Pompeu Fabra, Barcelona. Andersen acknowledges support from CREATES, Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. Hautsch acknowledges financial support from the Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF).

†Kellogg School of Management, Northwestern University, NBER and CREATES. Email: t-andersen@kellogg.northwestern.edu.

‡Faculty of Business, Economics and Statistics, University of Vienna. Email: goekhan.cebiroglu@univie.ac.at.

§Faculty of Business, Economics and Statistics, University of Vienna and Center for Financial Studies, Frankfurt. Email: nikolaus.hautsch@univie.ac.at.

1 Introduction

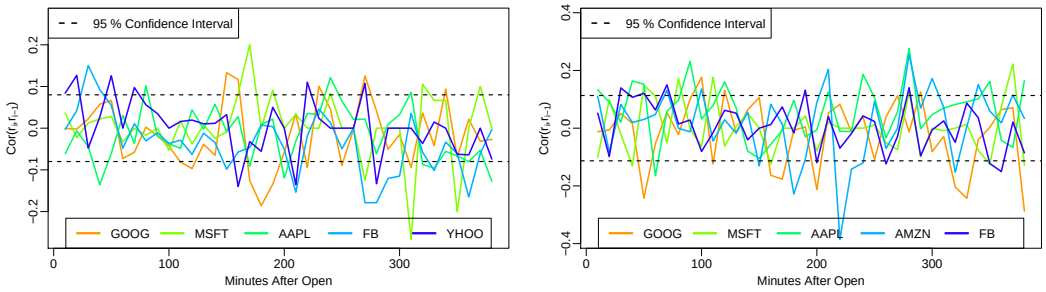
A major theme in the recent financial econometrics literature is the efficient extraction of daily or intra-daily volatility measures from high-frequency asset price data. A common starting point is to assume that micro-level prices, or quotes, are governed by an underlying latent semi-martingale process which, however, is polluted by noise. This noise component is typically attributed to market microstructure frictions, causing observed micro-level prices to deviate from the efficient (semi-)martingale process.

It is well-known that this so-called "market microstructure noise" (henceforth MMN) induces autocorrelation in high-frequency returns and may cause substantial biases in volatility estimators if one utilizes very high frequency returns. A classical result is that discretely sampled high-frequency returns, generated by a (semi-)martingale process polluted by i.i.d. noise, will display negative serial correlation and follow an MA(1) process. This relationship is already described by Roll (1984) in the context of transaction returns subject to bid-ask bounce effects. In a fully parametric setting, this property can be exploited to construct estimators of both the "efficient," or fundamental variance and the MMN variance, see, e.g., Zhang et al. (2005).

Hansen and Lunde (2006) show, however, that the dependence patterns in high-frequency returns typically are not consistent with an MA(1) representation, but tend to follow more general processes. This is particularly true if the high-frequency observations originate from the midpoint of bid and ask quotes (mid-quotes), rather than from transaction prices. Moreover, the former are less impacted by bid-ask bounce and, almost by construction, provide a closer approximation to the underlying "efficient" price. Given that the negative serial correlation is mitigated in mid-quote returns, one may wonder if the standard discrete-time representation of high-frequency asset returns, as following a random walk plus i.i.d. noise process, remains a sensible first-order approximation to the price dynamics. Figure 1 shows the empirical first-order autocorrelation in 1- and 2-second mid-quote returns, computed for local intervals of 10 minutes for constituents of the NASDAQ 100, over randomly selected trading days.

In the absence of serial correlation in returns, we expect less than five violations on either side of the confidence band. However, in Figure 1, we find about four times that number. Furthermore, the violations are scattered over the trading day and their timing differ across stocks. In addition, positive dependence appears roughly as likely as negative dependence. These findings are representative for the full cross-section of stocks and it is among the key empirical facts we scrutinize in this paper. It suggests that, in general, MMN cannot be i.i.d., but must have more general properties that accommodate *both* positive and negative serial correlation. We may also need to allow for the possibility that the noise is endogenous, i.e., correlated with fundamental price innovations. Among others, Hansen and Lunde (2006)

Figure 1: First-order autocorrelations in 1-second (left) and 2-second (right) mid-quote returns for a selection of stocks in the NASDAQ 100 index for two random days (January 2 and 9, 2014, on the left and right, respectively). The autocorrelations are computed over adjacent non-overlapping intervals of 10 minutes from open to close. We depict statistics for Apple, Cisco, Facebook, Google, Microsoft, and Yahoo. The dashed lines indicate the asymptotic 95% confidence band.



conclude that the MMN is (contemporaneously) *negatively* correlated with the efficient price.

The literature on high-frequency-based volatility estimation has responded by constructing modified procedures that allow for serial correlation of unknown form. The identification of the efficient price variation is obtained through the assumption that all the systematic serial correlation stems from a stationary noise process. It follows that consistent estimators may concentrate on the long-range variation, exploiting techniques along the lines of the long-run variance estimators in econometrics, as developed by Newey and West (1987) or Andrews (1991).[1] This nonparametric reduced-form approach is agnostic about the properties of the noise and focuses exclusively on estimating the quadratic variation of the random walk component.

Consequently, the critical notion in the high-frequency volatility literature remains that of a latent semi-martingale process whose value is known at all times by the active informed market participants. This requirement is obviously stringent for returns sampled at very high frequencies. The investor information set is not perfect, and it may not be updated instantaneously, but be subject to information acquisition and processing delay. That is, investors may operate with slightly stale or incomplete information. For example, they may not be able to determine whether an increase in the asset price signals the arrival of positive news or it stems from the random arrival of net buying orders from liquidity traders (noise). In the former case, the price likely continues on an upward trajectory as more investors learn of the news or infer its presence from the trading activity while, in the second case, the price will tend to revert back to the prior level. While an astute trader can avoid to *systematically* misprice the asset, this only implies that the valuation is unbiased on average. That is, during an episode with sustained imbalances in net liquidity demand or an unusual intensity of privately informed trading, even fully rational agents do not possess perfect information and they will err in their instantaneous valuation

[1] For a review of nonparametric noise-robust return variation estimators, see, e.g., the survey of Andersen et al. (2010) or Ait-Sahalia and Jacod (2014).

of the asset. On the other hand, rationality does imply that simple momentum and contrarian strategies are unprofitable. The fact that positive and negative return autocorrelation patterns are about equally common is consistent with the latter prediction.

The primary objective of this paper is to explore whether the high-frequency price dynamics displays features consistent with delays in the acquisition, processing or reaction to innovations in the underlying efficient price. Hence, we extend the classical martingale-plus-noise model by incorporating a notion of sequential learning or lagged response which allows for a more structural treatment of the observed temporal dependence patterns and serves to build a bridge to the related market microstructure literature. Motivated by empirical observations, as illustrated in Figure 1, we develop a model for mid-quote prices, which deviates from a pure martingale-plus-noise assumption by allowing for temporal feedback effects due to local discrepancies between the observed and efficient price. Our starting point is that those deviations, arising from frictions such as market microstructure noise or liquidity effects, in turn generate forces that try to correct these inefficiencies. As a consequence, the price discovery process is subject to a constant interplay between these counteracting forces, and corrective feedback plays a substantial role in explaining the temporal dependencies in mid-quote returns.

In fact, in terms of empirical plausibility, the findings in Figure 1 present a challenge. They suggest that the dependence structure in the mid-quote returns varies substantially over time. To retain a simple, transparent framework that facilitates interpretation, we stipulate that the model only applies locally over a short time interval, and then allow the parameters to vary across these non-overlapping windows. Estimating the model separately across short intervals allows the changing information and trading environment in the market to guide the time-varying parameter estimates.[2] Furthermore, the use of a local approximating model with constant volatility enables us to sidestep the difficult task of modeling the intraday patterns and persistent dynamics of the market activity, which almost inevitably would result in severe misspecification. Instead, we can focus the model on the critical features that distinguish it from existing stylized microstructure models and compare their performance across a large number of stocks and time intervals.

The main implication of our new representation is that the temporal dependence structure of observed returns arises from two separate channels. First, in line with the classical literature, market prices are disturbed by random noise arising due to market microstructure frictions or noise trading unrelated to information. The second channel is due to past "mispricing", i.e., deviations between past observed and efficient prices. The speed by which observed prices react to inherent mispricing is governed by a parameter that reflects the degree or "efficiency" with which investors are able to decipher the market signals and swiftly identify innovations

[2]This mimics the asymptotic devise often used in high-frequency volatility estimation where the inference theory is established under the assumption of constant volatility over local (and asymptotically shrinking) time intervals.

in the efficient price. Mispricing generates trading opportunities for informed traders in the spirit of Kyle (1985). Hence, informed arbitrage activity ensures that deviations between the observed and efficient price remain bounded and that the effects of non-informational shocks get mitigated, as informed agents impound their knowledge into observed prices through trading, and other agents infer the presence of information.

To allow for this type of features, the model incorporates an error correction component, reminiscent of the partial adjustment mechanism proposed by Amihud and Mendelson (1987) for modeling daily returns. The motivation for thinking along the lines of sequential learning in this context is inspired by Kyle (1985) and Vives (1995). Thus, we combine notions from high-frequency statistics and market microstructure theory. As a result, we establish the mechanism of temporal feedback as an essential component in models for high-frequency prices. We study the properties of the model in a simple discrete-time framework of (locally) constant volatility. In such a setting, the model depends on three key parameters, namely the volatility of the random walk process (the so-called "fundamental volatility"), the "noise level" corresponding to the variance of the noise process, and the informational efficiency parameter governing the speed of reversals due to mispricing. We show that such a three-parameter setting is quite flexible and can capture a wide range of stochastic behavior in mid-quote returns. As an important special case, our framework contains the classical "martingale-plus-noise" model, where the noise component, however is negatively correlated with the efficient price. Our temporal feedback mechanism, arising from local mispricing, therefore provides a natural explanation for the occurrence of so-called endogenous noise as analyzed, e.g., by Kalnina and Linton (2008).[3]

Our more structural approach for capturing the dynamics between the observed and efficient prices is related to a few prior studies that seek to link the properties of noise more explicitly to the process of trading and the underlying microstructure, e.g., Diebold and Strasser (2013), Chaker (2013) and Li et al. (2016). However, none of these studies explicitly consider temporal feedback effects which we deem crucial in explaining the return dynamics observed in Figure 1.

We show that our framework allows for a natural interpretation of the underlying model components. In particular, the speed of price reversion interacts with the ratio of the noise versus fundamental innovation volatility – the noise-to-signal ratio – to determine the sign of the return autocorrelation. This relation let us classify the (time-varying) market environment into two fundamental regimes. In one regime, the impact of "mispricing" is overcome by the strength of the price reversal, induced by idiosyncratic noise shocks, so that the returns display negative serial correlation, or "contrarian" price behavior. In the other regime, the feedback from "mispricing" generates "momentum," or positive autocorrelation, in the observed returns. Moreover, the regimes have important consequences for standard volatility measures

[3]See also Sheppard (2013) for a related concept to measure the speed of a market.

based on high frequency returns, e.g., Andersen et al. (2003). In the contrarian regime, the realized volatility will tend to exaggerate the underlying quadratic variation of the efficient price, inducing the well-known upward-shaped (for increasing sampling frequencies) volatility signature plots, e.g., Hansen and Lunde (2006). In the other regime, however, high-frequency sampling causes downward-shaped signature plots which are not readily explained within the traditional "martingale-plus-noise" setting, but arise naturally in our framework.

In its basic form, the model can be conveniently represented in a linear state-space form, allowing for standard pseudo-maximum likelihood estimation via the Kalman filter. Thus, we may readily construct a consistent local volatility estimator. Moreover, we develop a generalized representation in which specific parameter configurations allow us to encompass alternative models of interest, including the Amihud and Mendelson (1987) version of the lagged price adjustment hypothesis and the standard martingale-plus-noise specification. As such, our approach provides the first step towards the development of new high-frequency based volatility estimators that simultaneously retain a link to the existing market microstructure literature.

Utilizing high-frequency data from NASDAQ trading, processed via the data platform LOBSTER[4], we estimate the model over local intraday intervals based on high-frequency mid-quote returns sampled at different frequencies. We confirm that the model parameters vary over time. In line with our model, we identify local regimes of contrarian and momentum behavior which are triggered by the (estimated) speed of price reversion in relation to the noise-to-signal ratio. These regimes are persistent, but generally quite short-lived, with slightly higher persistence for the more liquid assets. The local estimates for the speed of price reversion point towards a relatively high degree of informational efficiency, although the price updating clearly also displays an element of sluggishness. We also document the presence of strong intraday periodicities in the speed of price discovery and the relative proportion of informational variance to total variance. Overall, our evidence strongly favors our newly developed specification with lagged price adjustment relative to other popular stylized microstructure models.

The remainder of the paper is structured as follows. In Section 2, we present the underlying model. Section 3 discusses the model's implications for return autocovariances and the identification of underlying market regimes. Section 4 illustrates implications for realized variances and resulting volatility signature plots. In Section 5, we discuss our estimation procedure. Section 6 provides the empirical results for our basic model, while Section 7 provides a more general representation that enables us to test the relative performance of a number of alternative models. Section 8 concludes. Auxiliary results and proofs are collected in the Appendix.

[4]See `https://lobsterdata.com/`.

2 The High-Frequency Asset Price Dynamics

We consider a model in discrete time, $i \in \{0, 1, 2, \ldots, T\}$, defined over an equidistant time grid with (minimal) interval length $\Delta = 1$. At each discrete point in time, we observe the logarithm of the quote midpoint for a single financial asset, p_i. Hence, for the time period $[0, T]$, we have a total of $T + 1$ separate log-price observations, and T consecutive continuously compounded returns, $p_i - p_{i-1}$, $i = 1, \ldots, T$. For the interval length $\Delta = 1$, the number of return observations corresponds to the length of the time period, $n = T/\Delta = T$, while for coarser sampling, governed by the integer $\Delta > 1$, the number of non-overlapping returns equals $n = \lfloor T/\Delta \rfloor$. While the notation is general, we focus on the salient features of the high-frequency return dynamics, so we typically think of Δ as representing a short intraday trading interval, on the order of seconds.

We assume that the price dynamics evolve according to the following scheme,

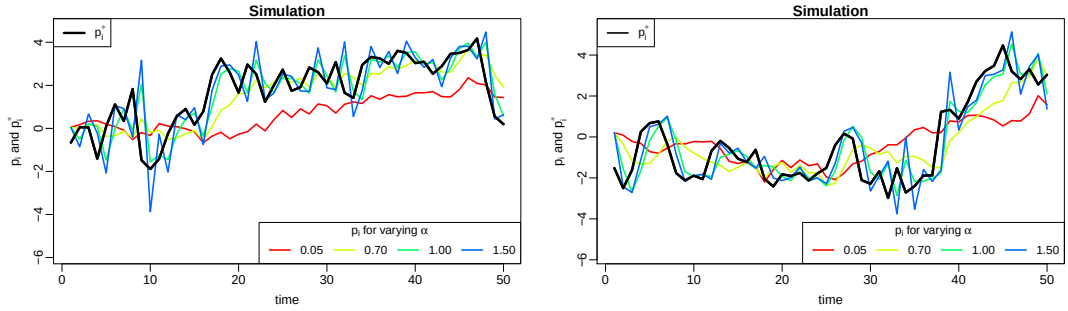
$$p_i = p_{i-1} - \alpha(p_{i-1} - p^*_{i-1}) + \varepsilon_i, \quad 0 < \alpha < 2, \tag{1}$$

$$p^*_i = p^*_{i-1} + \varepsilon^*_i, \tag{2}$$

where p^*_i denotes the "efficient" (full information and frictionless) log price at time i, that follows a random walk process with i.i.d. white noise innovations ε^*_i and $\mathbb{V}[\varepsilon^*_i] = \sigma^2_{\varepsilon^*}$. Accordingly, the observed mid-quote price p_i is driven by two components. It is directly impacted by an i.i.d. noise process ε_i with $\mathbb{V}[\varepsilon_i] = \sigma^2_\varepsilon$. This source of randomness is assumed to arise from non-informational sources, such as random exogenous trading motives or market microstructure frictions. It is, therefore, reasonable to assume that this source of randomness is uncorrelated with informational innovations, i.e., $\mathbb{E}[\varepsilon_i \varepsilon^*_i] = 0$. Second, p_i is influenced by the lagged deviation between p_i and p^*_i, which we denote by $\mu_i = p_i - p^*_i$. The "mispricing" component, μ_i, is akin to the "classical" MMN component, as it represents the discrepancy between the observed and efficient price. As illustrated subsequently in Section 3.2, an immediate consequence is that the observed (mid-quote) returns will display serial dependence.

Econometrically, the component $\alpha\mu_{i-1}$ serves as an error correction term which pushes the mid-quote back towards the "equilibrium" or "efficient" level, $p_i = p^*_i$. For example, if the past observed price exceeds the efficient price, i.e., $\mu_{i-1} > 0$, there is an induced negative mean-reversion component in the price process, whose strength is controlled by the parameter α. This error correction mechanism is motivated by the intuition that risk-averse agents with incomplete information form unbiased, yet not error-free, expectations about the efficient price, and therefore develop a notion of temporary mispricing. This creates opportunities for (risky) speculative trading that pulls the price back towards efficiency, i.e., when prices appear

Figure 2: Two simulated trajectories of p_i^* and p_i from the model defined through equations (1) - (2) for different values of α. The simulations are based on $\sigma^2_{\varepsilon^*} = 1$, $\sigma^2_{\varepsilon} = 1/10$, and thus $\lambda = 1/10$.



overvalued, informed agents short the asset. As such, the model captures temporal feedback effects through a minimalistic reduced-form approach. Earlier contributions, including Kyle (1985) and Vives (1995), discuss various settings and economic factors that render prolonged price reaction patterns consistent with sequential learning models.

Figure 2 illustrates the relationship between p_i and p_i^* as a function of α. Low values of α imply that the information transmission is slow, i.e., lagged pricing errors μ_i induce only weak price adjustments. In the extreme case of $\alpha = 0$, efficient and observed prices are entirely decoupled. At the other extreme, $\alpha > 1$, we have over-shooting, and the induced price changes $(p_i - p_{i-1})$ exceed the pricing errors μ_{i-1}.

To keep the model transparent and tractable, we assume that both the noise variance σ^2_{ε} and the variance of the "efficient" returns, $\mathbb{V}[\varepsilon_i^*] = \mathbb{V}[p_i^* - p_{i-1}^*] = \sigma^2_{\varepsilon^*}$, are constant over time. This is clearly unrealistic if $[0, T]$ covers a substantial time period, e.g., a day or even an hour, but it arguably provides a reasonable approximation for short (intraday) intervals. In fact, this approach mimics the theoretically motivated assumptions invoked for the development of inference techniques with high-frequency data in Mykland and Zhang (2009) and Bibinger et al. (2014), who explicitly rely on a locally fixed approximation of the underlying process.[5]

Alternatively, the model can be couched in a more standard random-walk-plus-i.i.d.-noise type of representation,

$$p_i = p_i^* + \mu_i, \tag{3}$$

$$\mu_i = (1 - \alpha)\mu_{i-1} + \varepsilon_i^{\mu}, \tag{4}$$

where $\varepsilon_i^{\mu} = \varepsilon_i - \varepsilon_i^*$ denote the innovations to the mispricing process μ_i, corresponding to the discrepancy between the shock to the noise and information components. Since these processes

[5]In Section 6, we estimate the model over short local intervals of 5, 10 and 30 minutes, rendering the constant volatility parameter a tenable assumption.

are i.i.d. white noise and independent, $\sigma_\mu^2 = \sigma_{\varepsilon^*}^2 + \sigma_\varepsilon^2$.

The specification (3) highlights the role of μ_i as a "classical" MMN component, causing deviations between p_i and p_i^*. Due to the temporal feedback induced by equation (1), however, μ_i in equation (4) is serially dependent, and follows a mean zero AR(1) process with autoregressive parameter $(1-\alpha)$. It is stationary for $0 < \alpha < 2$ with unconditional variance,

$$\mathbb{V}[\mu_i] = \frac{\sigma_\mu^2}{\alpha(2-\alpha)}. \tag{5}$$

The fact that the innovations to the mispricing component, ε_i^μ, are a linear function of both ε_i and ε_i^* has important implications for the covariance structure of ε_i^μ and ε_i^*. It takes the form

$$\Sigma := \begin{bmatrix} \mathbb{E}[(\varepsilon_i^\mu)^2] & \mathbb{E}[\varepsilon_i^\mu \varepsilon_i^*] \\ \mathbb{E}[\varepsilon_i^\mu \varepsilon_i^*] & \mathbb{E}[(\varepsilon_i^*)^2] \end{bmatrix} = \begin{bmatrix} \mathbb{E}[\mu_i^2] & \mathbb{E}[\mu_i \varepsilon_i^*] \\ \mathbb{E}[\mu_i \varepsilon_i^*] & \mathbb{E}[(\varepsilon_i^*)^2] \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_{\varepsilon^*}^2 & -\sigma_{\varepsilon^*}^2 \\ -\sigma_{\varepsilon^*}^2 & \sigma_{\varepsilon^*}^2 \end{bmatrix}.$$

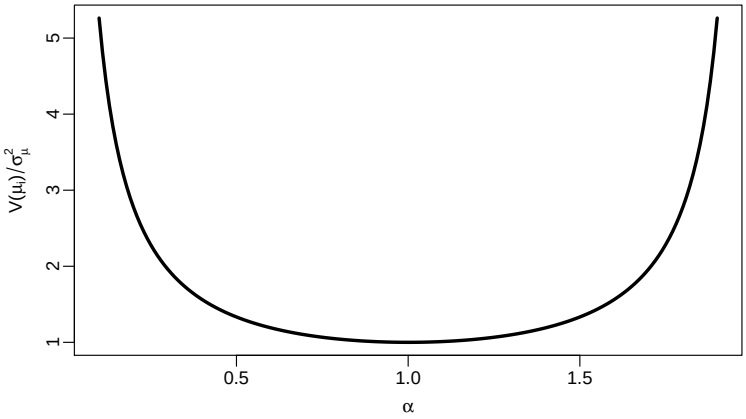
Thus, the model structure implies that μ_i and ε_i^* are negatively correlated, rendering the mispricing or MMN component μ_i "endogenous." It is a direct consequence of the temporal feedback relating p_i to the *lagged* efficient price p_{i-1}^*. This is further in line with the empirical evidence reported by, e.g., Hansen and Lunde (2006).

The error correction mechanism also ensures that the mid-quote price equals the efficient price in expectation, i.e., $\mathbb{E}[\mu_i] = 0$ and $\mathbb{E}[p_i] = \mathbb{E}[p_i^*]$, for $0 < \alpha < 2$. Nevertheless, the observed and efficient prices deviate over time, and the average discrepancy is related to the variance of μ_i, i.e., $\mathbb{V}[\mu_i]$. Specifically, if the observed prices p_i equal the efficient prices p_i^* throughout, $\mathbb{V}[\mu_i] = 0$, but as they deviate more systematically, $\mathbb{V}[\mu_i]$ grows larger. As such, $\mathbb{V}[\mu_i]$ serves as an indicator for the degree of informational inefficiency of the price process.

Intuitively, all else equal, rapid error correction, $\alpha \approx 1$, enhances price efficiency. Likewise, a low degree of noise, i.e., small σ_ε^2, shrinks the gap between the observed and efficient price, as it lowers ε_i^μ and therefore $\mathbb{V}[\mu_i]$. We depict the relationship between the "inefficiency" measure $\mathbb{V}[\mu_i]$ and the strength of the price reversion α in Figure 3 for $0 < \alpha \leq 1$ and $\sigma_\mu^2 = 1$. The relationship is symmetric in $|\alpha - 1|$ for $1 < \alpha < 2$. The figure conveys the pronounced sensitivity of $\mathbb{V}[\mu_i]$ to the strength of the error correction mechanism, with the most informative price scenario, or minimal price inefficiency, attained for complete one-period error correction ($\alpha = 1$).

In the benchmark case of $\alpha = 1$, where feedback perfectly removes prevailing mispricing,

Figure 3: Illustration of $\mathbb{V}[\mu_i]/\sigma_\mu^2$ depending on α for $\alpha \in (0, 2)$.



inherited from the previous period, the model can be written as,

$$p_i = p_{i-1}^* + \varepsilon_i \tag{6}$$

$$= p_i^* + \varepsilon_i^\mu, \tag{7}$$

with ε_i^μ collapsing to $\varepsilon_i^\mu = \varepsilon_i - \varepsilon_i^* = p_i - p_i^* = \mu_i$, and ε_i^μ is i.i.d. with $\mathbb{E}[\varepsilon_i^* \epsilon_i^\mu] = -\sigma_{\varepsilon^*}^2$.

This particular representation of the special case $\alpha = 1$ highlights a critical distinction relative to the classical random-walk-plus-iid-noise model, given by $p_i = p_i^* + \varepsilon_i$, see, e.g., Ait-Sahalia et al. (2005) and Bandi and Russell (2008). According to (6), our approach implies that the mid-quote price p_i is affected by the *lagged* efficient price, p_{i-1}^*, and not the *contemporaneous* efficient price p_i^*, as in the "classical" model. The temporal backshift in p_i^* is an integral feature of the model, reflecting the fact that observed prices are updated in response to recent trades (or quote updates) and newly available information, but with some delay. The reformulation in (7) illustrates that this lagged response renders the noise process endogenous. This type of endogeneity typically is imposed through explicit statistical assumptions regarding the noise terms, see, e.g., Kalnina and Linton (2008), but it arises structurally within our specification.

In summary, our parsimonious model, featuring just three parameters (α, $\sigma_{\varepsilon^*}^2$ and σ_ε^2), endows a simple high-frequency asset pricing process with distinct dynamic properties that are manifestations of underlying economic forces. The subsequent sections explore the theoretical ramifications for the associated high-frequency return process, while the empirical evidence in Section 6 illustrates how this minimalistic framework captures salient features of the data generating process over short intraday trading periods.

3 Second Order Return Moments

This section characterizes the conditions for excess volatility and serial correlation in the observed returns implied by the model. These features turn out to be interrelated and closely associated with the impact of MMN versus smoothing in the return variation process.

3.1 Excess Return Variation

We define $r_i = p_i - p_{i-1}$ and $r_i^* = p_i^* - p_{i-1}^*$ as the mid-quote continuously compounded and "efficient" return, respectively. The observed returns are then given by

$$r_i = -\alpha\mu_{i-1} + \varepsilon_i \,,$$

and the return variance is

$$\begin{aligned} \mathbb{V}[r_i] &= \sigma_\varepsilon^2 + \alpha^2\mathbb{V}[\mu_i], \\ &= \frac{1}{2-\alpha}(2\sigma_\varepsilon^2 + \alpha\sigma_{\varepsilon^*}^2)\,. \end{aligned} \tag{8}$$

Clearly, $\mathbb{V}[r_i] \geq \sigma_\epsilon^2$, which is a simple consequence of the fact that informationally-driven price variation adds to the non-informational noise component, ε_i, to generate overall return variation. The strength of the information component is governed by the size of α, reflecting the speed by which information about the efficient price is impounded in the observed price. For the benchmark case of $\alpha = 1$, the return variation is equivalent to the following result, well-known from the classical i.i.d. MMN model, $\mathbb{V}[r_i] = \sigma_{\varepsilon^*}^2 + 2\sigma_\varepsilon^2$. This scenario involves "excess volatility," as the total return variation exceeds the variation stemming from the information component. However, in our more general case, we do not necessarily encounter excess volatility. This critically depends on the relative amount of noise in the system along with the speed by which informational shifts are transmitted back into market prices.

To formalize the discussion, it is convenient to introduce notation for the ratio of non-informational versus informational variance, which also plays a crucial role in the characterization of other key features of the model. We define the *noise-to-signal ratio* λ as follows,

$$\lambda = \sigma_\varepsilon^2/\sigma_{\varepsilon^*}^2\,. \tag{9}$$

We may then write the unconditional return variance as

$$\mathbb{V}[r_i] \;=\; \sigma_{\varepsilon^*}^2\,\frac{2\lambda + \alpha}{2-\alpha}\,, \tag{10}$$

and it follows that,

$$\mathbb{V}[r_i] \leq \mathbb{V}[r_i^*] \qquad \text{if} \qquad \lambda \leq 1-\alpha, \tag{11}$$

$$\mathbb{V}[r_i] > \mathbb{V}[r_i^*] \qquad \text{otherwise.} \tag{12}$$

Consequently, in terms of excess return variation, we have two separate regimes, determined by the size of the noise-to-signal ratio, λ, and the sluggishness of the information feedback, $1-\alpha$. The impact of λ is expected; noise elevation translates into higher excess price variation. The ambiguity about excess volatility stems from the fact that deviations from the lagged efficient price induce a temporal correction process that smoothes the price path. For now, we focus on the more empirically relevant range of $0 < \alpha \leq 1$. If $\alpha = 1$, the impact of smoothing vanishes, and we are back in the classical MMN result, so the coefficient $1-\alpha$ is an indicator of the strength of the smoothing effect, which serves to reduce the return variability. Hence, the overall effect is dependent on the relative size of the excess return variation generated by the noise component and the volatility smoothing effect of the partial adjustment mechanism.

In the first regime, where $\lambda < 1-\alpha$, we obtain a reversal of the "classical" i.i.d. noise characterization of the market dynamics, where excess volatility invariably is associated with price inefficiency. Instead, we now find observed mid-quotes fluctuating less than the efficient price, yet the market is also less informationally efficient, as the low value of α slows down the transmission of price-relevant information. Formally, one may readily verify that, despite $\mathbb{V}[\mu_i]$ being monotonously decreasing (and price efficiency increasing) in α, $\mathbb{V}[r_i]$ is strictly increasing in α. Thus, lower values of α increase the average degree of mispricing, while simultaneously shrinking overall return volatility. As such, the regime provides a striking contrast to the classical i.i.d. MMN model, linking excess volatility to large pricing errors.

In the second regime, the noise proportion, λ, exceeds the sluggishness in the information transmission, $(1-\alpha)$. In this case, p_i is subject to a relatively higher degree of non-informational disturbances ε_i, or a faster transmission of price-relevant information. This induces return variation that exceeds the efficient price variation. Note that this regime trivially encompasses all scenarios for which there is overshooting (excess extrapolation of the efficient price information), $\alpha > 1$, and where $\lambda > 1$ (the noise variation overwhelms the informational variation), as the condition for $\mathbb{V}[r_i] > \mathbb{V}[r_i^*]$ automatically is satisfied in these circumstances.

3.2 Return Auto-Covariances

The determinants of whether returns display excess volatility are also critical for the sign of the return auto-covariances. This is a consequence of the following lemma and corollary.

Lemma 1 (Return Auto-Covariances)**.** *Assume* $\sigma_\varepsilon^2 > 0$, $0 < \alpha < 2$, *and* $h \geq 1$. *Then,*

$$\mathbb{C}ov[r_i, r_{i-h}] \;=\; \psi(h-1)\;\sigma_{\varepsilon^*}^2\;\frac{(1-\alpha-\lambda)}{2-\alpha}, \tag{13}$$

with $\psi(h-1) \;=\; \alpha\,(1-\alpha)^{h-1}$, *and* $\psi(0) = 1$, *if* $\alpha = 1$.

Proof. See Appendix. □

Hence, the relation between market sluggishness, 1-α, and the noise-to-signal ratio, λ, governs the sign of the auto-covariance of observed returns, as detailed below.

Corollary 1. *Assume* $\sigma_\varepsilon^2 > 0$, $0 < \alpha < 2$, *and* $h \geq 1$.

(i) If $0 < \alpha < 1$, *then* $sgn\{\mathbb{C}ov[r_i, r_{i-h}]\} \;=\; sgn\{(1-\alpha) - \lambda\}$.

(ii) If $\alpha = 1$, *then* $\mathbb{C}ov[r_i, r_{i-1}] = -\sigma_\varepsilon^2 < 0$, *and* $\mathbb{C}ov[r_i, r_{i-h}] = 0$, *for* $h > 1$.

(iii) If $1 < \alpha < 2$, *then* $sgn\{\mathbb{C}ov[r_i, r_{i-h}]\} = sgn\{(-1)^h\}$.

For the benchmark case of $\alpha = 1$, the error correction is complete within one period. This eliminates any temporal dependency in the return component associated with the efficient price, while the temporary noise shock is eliminated in the following trading session, inducing the standard first-order negative serial correlation in returns, also observed in the classical i.i.d. noise model. For the region $0 < \alpha < 1$, the informational trading channel generates only partial updating, inducing an element of price smoothing and, critically, at tendency towards positive return serial correlation, as the price adjustment is spread across multiple periods. When α and λ are sufficiently low, such that $1 - \alpha > \lambda$, this effect dominates and the observed mid-point returns display positive auto-correlation. Intuitively, we expect this scenario to capture situations in which risk-averse and partially informed agents are actively trading on their signals, enabling others to infer the presence of informed trading from the underlying market activity. As long as these trading patterns are sustained over some time, and possibly supported by the contemporaneous public information flow, rational market participants will continue to revise the probability of informed trading upward, adding to the forces generating price momentum. This regime is also characterized by subdued return variation due to the prevalence of price smoothing, as documented in Section 3.1.

In contrast, when the noise-to-signal ratio exceeds the degree of market sluggishness, observed returns are negatively autocorrelated. In this regime, the price changes and pricing errors μ_i are dominated by the noise component, and this manifests itself in mean-reversion for the observed price. The resulting return pattern is qualitatively consistent with "contrarian" trading behavior. As shown in Section 3.1, in this regime, the variance of the observed returns

exceeds the informationally efficient variance, as the noise shocks and their subsequent reversals creates excessive price variation.

Furthermore, we note that, for $0 < \alpha < 1$, the sign of the auto-correlations are identical for all $h \geq 1$, indicating that both the "contrarian" and "momentum" regimes imply a certain persistence in the return dependencies. To the extent that these regimes provide a good approximation to the local return dynamics over short, yet non-trivial, trading intervals, there may be an opportunity to capture these episodes empirically. In contrast, for $\alpha > 1$, the first-order return serial correlation is negative, and the signs alternate for $h > 1$. This dynamic behavior is induced by period-by-period over-correction of the mispricing. While such behavior may be plausible over longer horizons, it is hard to envision this type of high-frequency shifts in the return dependencies for mid-quotes across short intraday trading intervals.

To exemplify the occurrence of persistent return correlation patterns, we illustrate the return dynamics implied by the model for two simple scenarios.

Example 1 (Positive serial correlation)**.** *Consider $\alpha = 0.5$ and $\lambda = 0$, and assume the efficient and observed price initially coincide, i.e., $p_0 = p^* = 0$. At time $i = 1$, new information ($\varepsilon_1^* = 2$) triggers a change in the efficient price to $p_1^* = 2$. After this shock, the efficient price remains constant, i.e., $\varepsilon_i^* = 0$ for $i \geq 2$. Then the observed returns $r_i = -\alpha\mu_{i-1}$ at times $i > 0$ obey,*

$$
\begin{aligned}
r_1 &= -\frac{1}{2}(0-0) \;=\; 0\,, & p_1 &= p_0 + r_1 \;=\; 0\,.\\
r_2 &= -\frac{1}{2}(0-2) \;=\; 1 > 0\,, & p_2 &= p_1 + r_2 \;=\; 1\,.\\
r_3 &= -\frac{1}{2}(1-2) \;=\; 1/2 > 0\,, & p_3 &= p_2 + r_3 \;=\; 3/2\,.
\end{aligned}
$$

Hence, the innovation in the efficient price at $i = 1$ generates a sequence of positively correlated returns for $i > 1$.

Example 2 (Negative serial correlation)**.** *Consider $\alpha = 1/2$ and $\lambda = \infty$, and assume $p_0 = p_0^* = 0$. A non-informational noise shock arrives at $i = 1$, i.e. $\varepsilon_1 = 2$, $\varepsilon_i = 0$ for $i > 1$. In the absence of information arrivals, $\varepsilon_i^* = 0$ for $i \geq 0$, the observed returns evolve as follows,*

$$
\begin{aligned}
r_1 &= -\frac{1}{2}(0-0) + 2 \;=\; 2\,, & p_1 &= p_0 + r_1 \;=\; 2\,.\\
r_2 &= -\frac{1}{2}(2-0) + 0 \;=\; -1 > 0\,, & p_2 &= p_1 + r_2 \;=\; 1\,.\\
r_3 &= -\frac{1}{2}(1-0) + 0 \;=\; -1/2 < 0\,, & p_3 &= p_2 + r_3 \;=\; 1/2\,.
\end{aligned}
$$

Hence, non-informational noise shock at $i = 1$ generates a sequence of returns, for $i > 1$, that are negatively correlated with the original return innovation.

The examples portray extreme scenarios with $1-\alpha > \lambda$ and $1-\alpha < \lambda$, respectively, generating pronounced positive and negative autocorrelation in returns. It is often assumed that mid-quote returns, at least on average, display little serial correlation. In our setting, this is feasible only if $1-\alpha = \lambda$. This will be approximately true, if the noise component is relatively small for the mid-quote returns, while prices impound new information swiftly, so that $1-\alpha$ is small also. More generally, zero return correlation is consistent with some noise in the returns as long as the impact is mitigated by a certain sluggishness in market prices. However, a relatively large noise component, $\lambda \geq 1$, will always induce negative serial correlation and excess volatility in returns due to strong price reversals.

4 Implications for the Realized Variance

A natural application of the developed model is to explore the implications for volatility measurement based on high-frequency return data. It is well-known that significant biases may arise if the underlying martingale assumption for the return innovations are violated to a significant degree. Our model and data exploited in the empirical work, embody features that render such consequences quite plausible. Hence, in this section we explore the basic properties of the quadratic variation of the observed price process implied by equations (1) and (2). We demonstrate that the identified regimes translate into distinct empirical features in the standard return variation measures. This may have important implications for both the interpretation and design of bias-adjustment procedures for such measures, when constructed from returns sampled at high frequencies.

In this context, it is important to recognize the manner in which the findings depend on the sampling frequency. To this end, we introduce additional notation to keep track of prices and returns for different interval lengths. As before, the underlying observations reside in the interval $[0,T]$, with $\Delta = 1$ being the highest possible sampling frequency, and T being measured in terms of the unit of the corresponding (smallest possible) sampling interval. To compute the high-frequency volatility measures at a lower frequency, $\Delta \geq 1$, we exploit the log prices, $p_0, p_\Delta, \ldots, p_{i\Delta}, \ldots, p_T$, at equidistant points in calendar time, $i = 1, \ldots, T/\Delta - 1, T/\Delta$, with (short) grid size Δ, where we assume, for simplicity and without loss of generality, that T/Δ is an integer. Thus, at frequency Δ, we observe a total of $n = T/\Delta$ non-overlapping continuously compounded returns, $r_{i\Delta} = p_{i\Delta} - p_{(i-1)\Delta}$.

The corresponding realized return variance measure at time T is given as

$$RV_T^\Delta = \sum_{i=1}^{T/\Delta} r_{i\Delta}^2. \tag{14}$$

In extension of equation (8), the following theorem provides the expected value of the model-implied return variation, given we are in the empirically relevant region, $0 < \alpha < 1$. For a discrete time model, this quantity is also known as the previsible quadratic variation or the sharp bracket process, and we exploit the common notation, $\langle p \rangle_T^{\Delta} = \mathbb{E}\left[RV_T^{\Delta}\right]$.

Theorem 1. *For $0 < \alpha < 1$, the expected time-T realized variance sampled at calendar time grid size Δ equals,*

$$\langle p \rangle_T^{\Delta} = T \cdot \sigma_{\varepsilon^*}^2 + T \cdot \sigma_{\varepsilon^*}^2 \cdot \phi(\Delta) \frac{\lambda - (1-\alpha)}{(2-\alpha)}, \tag{15}$$

$$\textit{with} \qquad \phi(\Delta) = \frac{2}{\alpha\Delta}\left(1 - (1-\alpha)^{\Delta}\right). \tag{16}$$

The mapping $\Delta \mapsto \phi(\Delta)$, from $\mathbb{R}_+$ into $\left(0, -\frac{2}{\alpha}\ln(1-\alpha)\right)$ is strictly decreasing with,

(i) $\lim_{\Delta \to 0} \phi(\Delta) = -\frac{2}{\alpha}\ln(1-\alpha)$,

(ii) $\lim_{\Delta \to \infty} \phi(\Delta) = 0$.

Proof. See Appendix. □

We obtain the following corollary, mimicking the conditions in Section 3.1, determining whether we reside in a regime inducing upward or downward biases in the observed return variance relative to the efficient or fundamental return variance.

Corollary 2. *For $0 < \alpha < 2$,*

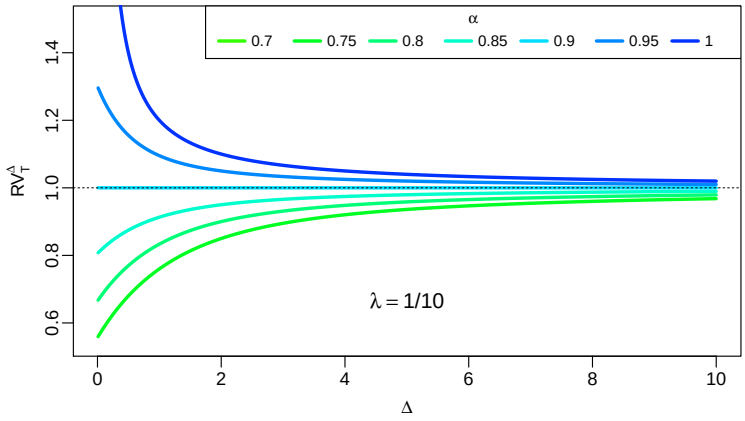
(i) if $\lambda > (1-\alpha)$, then $T \cdot \sigma_{\varepsilon^*}^2 < \langle p \rangle_T^{\Delta}$,

(ii) if $\lambda < (1-\alpha)$, then $T \cdot \sigma_{\varepsilon^*}^2 > \langle p \rangle_T^{\Delta}$.

Equation (15) of Theorem 1 decomposes the expected quadratic variation into two terms. The first signifies the expected variation of the efficient price ($T \cdot \sigma_{\varepsilon^*}^2$), reflecting the arrival of fundamental information over the interval $[0, T]$. The second term captures the impact of the noise component and the feedback between the efficient and observed prices. Its sign depends on two separate factors, namely the noise-to-signal ratio λ and the sluggishness of price updating $(1-\alpha)$. These components mirror those relevant for our prior discussion of the determinants for the unconditional return variance and excess volatility.

Theorem 1 demonstrates that, for our model, more frequent sampling of the returns enlarges the (absolute) bias of the volatility measure, whether it is positive or negative. Moreover, if the bias is positive, then large values of α and λ (quicker correction of pricing errors and a

Figure 4: Theoretical quadratic variation $\langle p \rangle_T^{\Delta}$ as a function of the sampling interval Δ for different realizations of α. Exemplary parameters are $T = 1$, $\sigma^2_{\varepsilon^*} = 1$, $\lambda = 1/10$. The dotted line corresponds to the efficient level, i.e., $\sigma^2_{\varepsilon^*} T$, which is obtained for $\alpha = 0.9$ and $\lambda = 1 - \alpha$.



more prominent role for noise), all else equal, increase the average realized return volatility and inflates the upward bias in this volatility measure, which is a natural consequence of the regime characterization given in the corollary.

The two regimes are graphically illustrated in Figure 4. The underlying forces determining the two regimes are analogous to those governing the return auto-covariance structure, described in Lemma 1. When the noise-to-signal ratio is sufficiently pronounced, observed returns are negatively autocorrelated, with the noise component dominating the smoothing effect of the error correction. As the sampling interval lengthens, the price adjustment becomes more pronounced, and the quadratic variation decreases monotonically with the size of Δ. Thus, as we shrink Δ, and rely on returns at a higher frequency, the volatility measure grows sharply. Such upward sloping signature plots are typically observed for realized volatility computed from transaction prices. However, there is another possibility. If the noise component is relatively small or the speed of the error correction is sufficiently high, then, the returns are positively serially correlated, and returns over longer horizons are more effectively smoothed over the sampling interval. This scenario generates the opposite effect, with the realized volatility dropping off as the returns are sampled more frequently. While this phenomenon has received only limited attention in the literature, our empirical evidence in Section 6 shows that such patterns can be commonplace for mid-quote returns, as also noted in the work by Hansen and Lunde (2006).

5 Estimation

Recall that the model is written as,

$$r_i = -\alpha\mu_{i-1} + \varepsilon_i,$$
$$\mu_i = (1-\alpha)\mu_{i-1} + \varepsilon_i^{\mu},$$

where $\varepsilon_i^{\mu} = \varepsilon_i - \varepsilon_i^*$, with ε_i and ε_i^{μ} being i.i.d. with variances σ_{ε}^2 and $\sigma_{\mu} = \sigma_{\varepsilon^*}^2 + \sigma_{\varepsilon}^2$. Hence, returns r_i are driven by a latent variable μ_i following an AR(1) process with covariance structure given by

$$\mathbb{E}[\varepsilon_i\mu_{i+h}] = (1-\alpha)^h\sigma_{\varepsilon}^2 \quad \forall h \geq 0,$$
$$\mathbb{E}[\varepsilon_i\mu_{i-h}] = 0 \quad \forall h > 0,$$
$$\mathbb{E}[\varepsilon_i\varepsilon_i^{\mu}] = \sigma_{\varepsilon}^2,$$
$$\mathbb{E}[\varepsilon_i\varepsilon_{i-h}^{\mu}] = 0 \quad \forall h \neq 0.$$

To estimate the model, it is convenient to re-write it in terms of a state-space representation. Denote X_i as a state vector at i with $X_i = (\mu_i, \mu_{i-1}, \varepsilon_i)'$. Then, r_i can be written as

$$r_i = FX_i,$$
$$X_i = GX_{i-1} + w_i,$$

with $F = (0, -\alpha, 1)$ and

$$G := \begin{pmatrix} (1-\alpha) & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The error term vector is given by $w_i = (\varepsilon_i^{\mu}, 0, \varepsilon_i)$ with covariance matrix,

$$\Sigma_w = \begin{pmatrix} \sigma_{\varepsilon}^2 + \sigma_{\varepsilon^*}^2 & 0 & \sigma_{\varepsilon}^2 \\ 0 & 0 & 0 \\ \sigma_{\varepsilon}^2 & 0 & \sigma_{\varepsilon}^2 \end{pmatrix}.$$

Three state equations are required to capture the covariance structure between ε_i^{μ} and ε_i. If we impose normality for ε_i and ε_i^{μ}, we can estimate the parameters α, σ_{ε}^2 and $\sigma_{\varepsilon^*}^2$ based on the Kalman filter, utilizing the corresponding error prediction decomposition of the (Gaussian) likelihood, see, e.g., Harvey (1989). Using pseudo-maximum likelihood (PML) arguments in

an exponential family setting, the estimates are further consistent, although not efficient, under distributional misspecification as long as the conditional mean processes are correctly specified, see, e.g., Gourieroux et al. (1984).

An alternative approach relies on unconditional moment restrictions implied by the model. One such condition is stems from the model-implied quadratic variation and the realized variance,

$$\phi_1(r_i;\alpha;\sigma_\varepsilon^2,\sigma_{\varepsilon^*}^2) = \sigma_{\varepsilon^*}^2 n - \sigma_\mu^2 n\phi(\Delta)\frac{1-\alpha-\lambda}{(2-\alpha)(\lambda+1)} - \sum_{i=1}^{n} r_i^2$$

Further moment conditions are implied by high-order moments of returns r_i,

$$\begin{aligned}\phi_2(r_i;\alpha;\sigma_\varepsilon^2,\sigma_{\varepsilon^*}^2) &= r_i^2 - \frac{1}{2-\alpha}(2\sigma_\varepsilon^2 + \alpha\sigma_{\varepsilon^*}^2),\\ \phi_{2,h}(r_i;\alpha;\sigma_\varepsilon^2,\sigma_{\varepsilon^*}^2) &= r_i r_{i-h} - \psi(h-1)\sigma_{\varepsilon^*}^2\,\frac{1-\alpha-\lambda}{2-\alpha},\end{aligned}$$

with $\psi(h) = \alpha(1-\alpha)^h \geq 0$ defined as in Lemma 1, and $h = 1, 2, \dots$.

A GMM estimator according to Hansen (1982) can then be formulated as,

$$\hat{\theta}(\mathcal{W}) = \arg\min_\theta \left[\frac{1}{n}\sum_{i=1}^{n}\tilde{m}(r_i;\theta)\right]' \mathcal{W}_n \left[\frac{1}{n}\sum_{i=1}^{n}\tilde{m}(r_i;\theta)\right],$$

where $\theta = (\alpha, \sigma_\varepsilon^2, \sigma_{\varepsilon^*}^2)'$, while $\tilde{m}(r_i;\theta) = (\phi_1(r_i;\theta), \phi_2(r_i;\theta), \phi_{2,1}(r_i;\theta), \phi_{2,2}(r_i;\theta), \dots)$ represents a set of model implied moment conditions, and $\mathcal{W}_n$ is a conforming positive definite weighting matrix.

6 Empirical Evidence

We employ high-frequency mid-quote prices of the constituents of the NASDAQ 100 index, covering the full 6 1/2 hour trading day. The underlying data is provided by the LOBSTER database[6] building on NASDAQ's historical TotalView-ITCH data. Our sample covers the first 40 trading days of 2014. The model is estimated by (pseudo-) maximum likelihood using the Kalman filter as described in Section 5.[7]

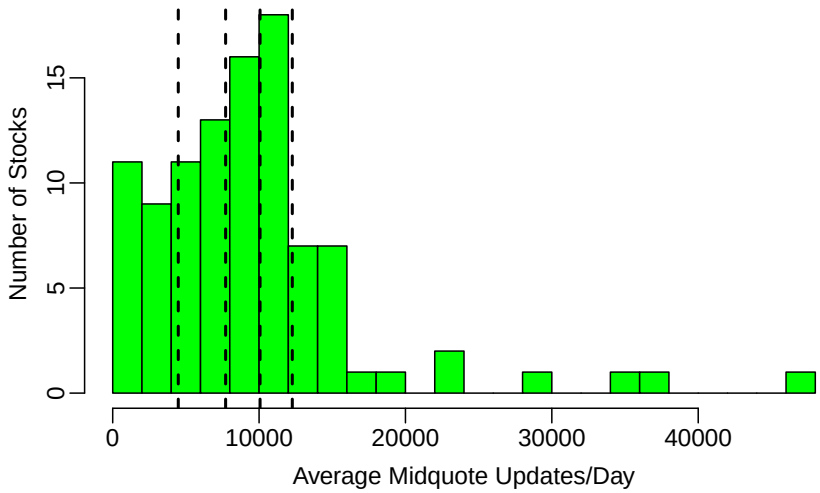
[6] See `https://lobsterdata.com/`.

[7] We also validated these estimates through the method of moments approach using the first three moment conditions provided in Section 5.

6.1 Basic Features of the Sample

A major objective of the paper is to explore the time-varying characteristics of the mid-quote return series across short calendar intervals, spanning 5-30 minutes. Thus, our study focuses on liquid stocks with frequent quote revisions. We divide our sample into five groups based on the quintiles of the average daily number of mid-quote revisions, with Group 1 representing those with the least quote revisions and Group 5 those with the most. The histogram in Figure 5 displays the distribution of daily (mid-) quote revisions, with vertical lines representing the 20'th, 40'th, 60'th and 80'th percentiles. The corresponding quintile means are 2,210, 6,130, 8,710, 11,180, and 19,400 quote revisions. If we sample the mid-quotes every 5 seconds (over 6 1/2 hours of trading), we have 4,680 return observations per day, and sampling each 10 seconds lead to 2,340 return observations. Consequently, for the typical stocks in Group 1, we have on average fewer quote revisions than intra-daily returns at these frequencies, while this ratio is much more favorable for Groups 2 through 5. For this reason, we focus our quantitative analysis primarily on the stocks in Group 2-5. Nonetheless, for illustrative purposes, we also include Group 1 stocks in some aspects of the analysis. Additional details about our five groupings of NASDAQ 100 stocks are provided in Table 4 of the Appendix. It reveals that Group 1 has relatively few quote revisions, in part because it contains low-priced stocks with large (percentage) tick sizes and high depth on the first level of the order book.

Figure 5: Empirical distribution of per-stock averages of daily mid-quote revisions for 100 stocks, based on the first 40 trading days in 2014. The vertical dashed lines indicate the corresponding quintiles, i.e., the 20th, 40th, 60th and 80th percentiles.



One key issue is whether there is evidence of local regimes with persistent return autocorrelations of either sign. This is an empirical question which we partially address in this section by studying the sample return autocorrelation over consecutive small intervals. In so doing,

we maintain that the expected return over a short interval is negligible. Hence, in computing the sample autocorrelations, the mean is treated as known and equal to zero. This provides an improved trade-off between bias and efficiency, as the return innovations are large and render the realized return an extremely imprecise estimator for the underlying mean return.

Table 1: Table reports the percentage of estimated significant first-order return autocorrelations, $\widehat{\mathbb{C}or}(r_i, r_{i-1})$, at the 1%, 5% and 10% level for various specifications of the estimation window T and sampling frequency Δ. The first three columns report the percentage of observations exhibiting significant autocorrelation using a two-sided test; the next six columns report the percentage of observations that exhibit significant positive autocorrelation (Columns 4 to 6) and significant negative autocorrelation (Columns 7 to 9) using corresponding one-sided tests. The last column provides the number of observations in each estimation window, with $n = T \cdot 60/\Delta$. The underlying confidence intervals are based on the asymptotic standard deviation of the empirical first-order autocorrelation, $sd = 1/\sqrt{n}$. Estimates are based on the first 40 trading days of 2014, starting on January 2nd. Each trading day has 390 trading minutes. Thus, for an intraday window of $T = 10min$, the total number of estimates for the 80 stocks in Groups 2-5 is $N = 39 \cdot 40 \cdot 80 = 124,800$. For $T = 5min$, we have $N = 249,600$.

	Two-Sided			One-Sided $\widehat{\mathbb{C}or}(r_i, r_{i-1}) < 0$			One-Sided $\widehat{\mathbb{C}or}(r_i, r_{i-1}) > 0$			
$T = 30min \quad (N = 41,600)$										
Δ	1%	5%	10%	1%	5%	10%	1%	5%	10%	n
$1sec$	25.8	37.2	44.7	18.8	27.0	32.6	11.3	17.7	22.3	1800
$2sec$	15.7	27.3	35.7	9.5	17.3	22.8	10.3	18.4	23.8	900
$5sec$	9.4	19.7	27.1	4.5	9.8	14.4	8.4	17.3	24.4	360
$10sec$	6.4	15.4	22.5	2.3	6.6	10.9	6.6	15.9	23.1	180
$T = 10min \quad (N = 124,800)$										
Δ	1%	5%	10%	1%	5%	10%	1%	5%	10%	n
$1sec$	17.9	28.7	36.2	13.0	20.8	26.2	8.9	15.4	20.3	600
$2sec$	10.3	19.9	27.1	6.6	13.2	18.5	6.8	14.0	19.8	300
$5sec$	5.0	13.0	19.9	2.7	7.6	12.3	4.8	12.2	18.6	120
$10sec$	3.0	9.3	15.6	1.5	5.3	9.2	3.3	10.3	17.1	60
$T = 5min \quad (N = 249,600)$										
Δ	1%	5%	10%	1%	5%	10%	1%	5%	10%	n
$1sec$	14.1	23.9	30.7	10.1	17.2	22.2	7.5	13.6	18.2	300
$2sec$	7.9	15.9	22.3	5.1	10.9	15.7	5.4	11.4	16.7	150
$5sec$	3.5	9.6	15.6	1.9	6.0	10.4	3.3	9.6	15.5	60
$10sec$	1.5	6.0	11.2	0.8	3.8	7.5	1.8	7.3	13.6	30

Table 1 summarizes evidence regarding the statistical significance of return autocorrelations

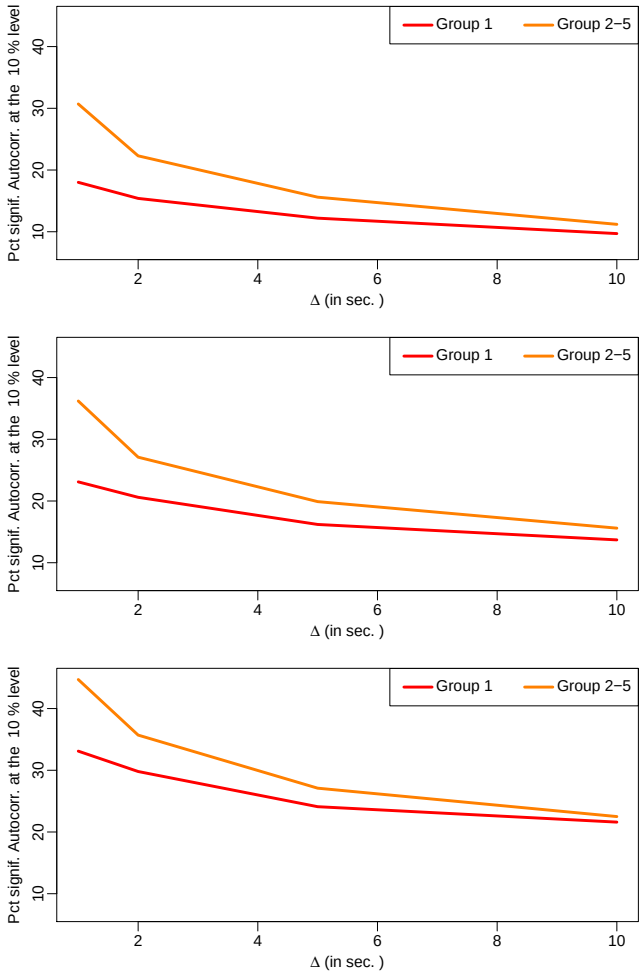
over intervals of 5, 10, and 30 minutes, computed from an intra-interval sampling frequency of 1, 2, 5, and 10 seconds. The results are based on the full set of 80 individual stocks in Groups 2-5, using the first 40 trading days of 2014. The exclusion of stocks from Group 1 (with relatively few quote revisions) is not critical for the conclusions below. The first three columns report the number of times the computed autocorrelation deviates significantly from zero, while the following two sets of three columns report the number of significant *negative* and *positive* autocorrelations, respectively. For the highest sampling frequencies, we have a large number of return observations at our disposal, but one may be concerned about the impact of microstructure noise. At the lower sampling frequencies, the impact of noise is less problematic, but we have fewer observations, so the estimation precision drops and the power to reject the null hypothesis declines. Hence, there is a fundamental trade-off in the designs, and we report results across a broad set of configurations to convey a general sense of the findings. We also caution that we expect the most suitable return frequency and window length to vary over time for a given stock and to differ across stocks. Given the unobserved variation in the market environment and the heterogeneity across stocks, we merely apply the assumption of locally constant return autocorrelations to search for robust features of the data that speak to critical aspects of the high-frequency return dynamics.

Table 1 reveals that we have many more statistically significant autocorrelations than expected under the null hypothesis of uncorrelated high-frequency returns. Moreover, we seem to acquire additional power to reject the null hypothesis if we expand the number of observations, either through a higher sampling frequency or a longer time interval. Another important feature of Table 1 is the asymmetry in the autocorrelations. For all sampling frequencies, interval lengths, and significance levels, we observe a larger number of positive than negative autocorrelations. Since the test for significant non-zero autocorrelations is symmetric, the hypothesis that positive or negative serial-correlation occurs with equal probability is also readily tested. Given the large number of observations, we find this test to reject the null hypothesis with an extreme degree of significance. In other words, the distribution of the return autocorrelations is definitely positively skewed. This confirms our conjecture that the dynamics of the high-frequency mid-quote returns are very different from those derived from transaction prices, for which the bid-ask bounce induces strong *negative* serial correlation. At the same time, it is also evident that the dynamics varies over time, consistent with the idea that the return autocorrelation changes sign across the states of the market. In particular, we find evidence for a predominance of "momentum" relative to "contrarian" regimes, although both are quite common.

Figure 6 depicts the proportion of estimated return autocorrelations that are significant at the 10% level as a function of the sampling frequency for both Group 1 and Group 2-5 stocks. The qualitative pattern is identical for both set of stocks but – as expected – the large proportion

of zero returns in Group 1 is associated with a noticeable drop in the relative frequency of significant autocorrelations.

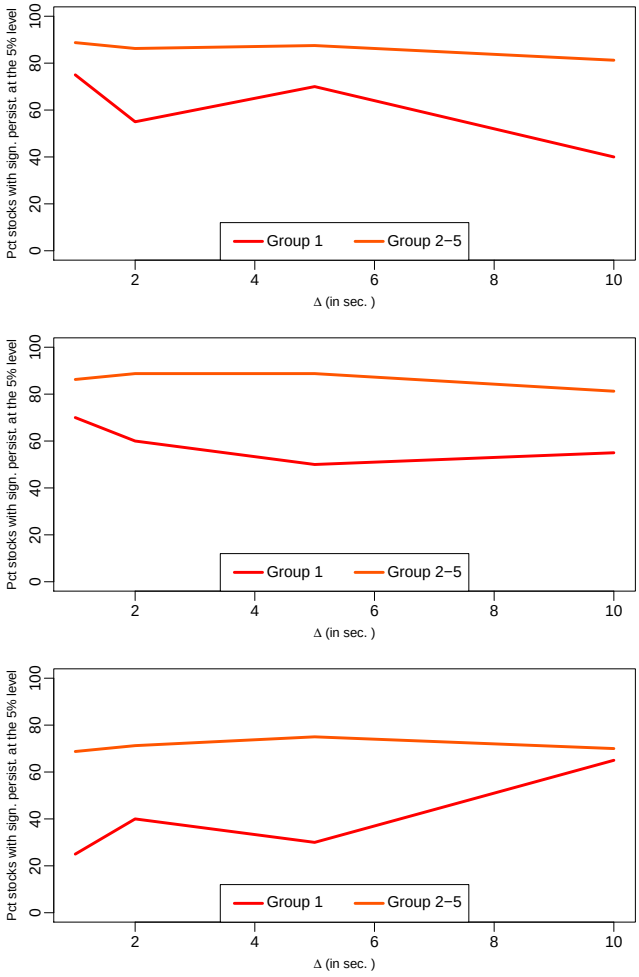
Figure 6: The figures show the proportion of significant first order return autocorrelations at the 10% significance level for different settings of the intraday estimation period T and sampling frequency Δ. The two curves refer to the 20 stocks in Group 1 and the 80 stocks in Group 2-5, respectively. The first graph shows the result for $T = 5min$, the second for $T = 10min$, and the last for $T = 30min$.



The interpretation of the significant return autocorrelations in Table 1 as representing short-lived regimes, reflecting the state of information processing and risk-return assessment in the market, is not directly testable. However, we can gauge whether the significant return autocorrelations occur randomly or reflect a more permanent feature of the market environment. If the non-zero autocorrelations are indicative of market conditions, we would expect some permanency in the state, even if they tend to be fairly short-lived. On the other hand, if they arise from purely idiosyncratic noise or random shocks, we would not expect to observe any significant persistence in the interval-specific return autocorrelation estimates.

Figure 7 reveals that the interval-to-interval correlation of the estimated first order auto-

Figure 7: The figures show the proportion of stocks exhibiting significant window-to-window first order autocorrelation of the first order return autocorrelation at the 10% level across different settings of the intraday estimation period T and sampling frequency Δ. The 20 stocks in Group 1 and the 80 stocks in Groups 2-5 are displayed separately. The first graph shows the result for $T = 5min$, the second for $T = 10min$ and the last for $T = 30min$.



correlation coefficients is significant (and positive) for a large majority of the stocks in Group 2-5. Moreover, the number of significant coefficients is higher for the shorter time intervals, consistent with regimes that tend to persist across consecutive intervals, but are of relatively short duration. As expected, we find less persistence among Group 1 stocks, which feature fewer mid-quote revisions. Nonetheless, the qualitative findings are similar across all groups. Overall, these results are consistent with our thesis that the dynamics for the mid-quote returns undergo distinct periods of positive and negative serial correlation, linked to variation in the market environment. These effects are at the heart of the model introduced in Section 2.

6.2 Parameter Estimates

We now explore the estimation results for our high-frequency return model defined through equations (1)-(2), based on the Kalman filter approach detailed in Section 5.

Figure 8 depicts the inferred empirical distribution for key parameter estimates and summary statistics, obtained from localized 10 minute windows with a sampling frequency of 5 seconds, across the first 40 trading days in 2014 for our 100 stocks. The figure displays the distribution for the full sample and for the five quintiles of stocks sorted on the intensity of the quote-midpoint revisions. The findings are qualitatively identical for alternative window sizes and sampling frequencies, as evident from Figures 15 to 17 in the Appendix.

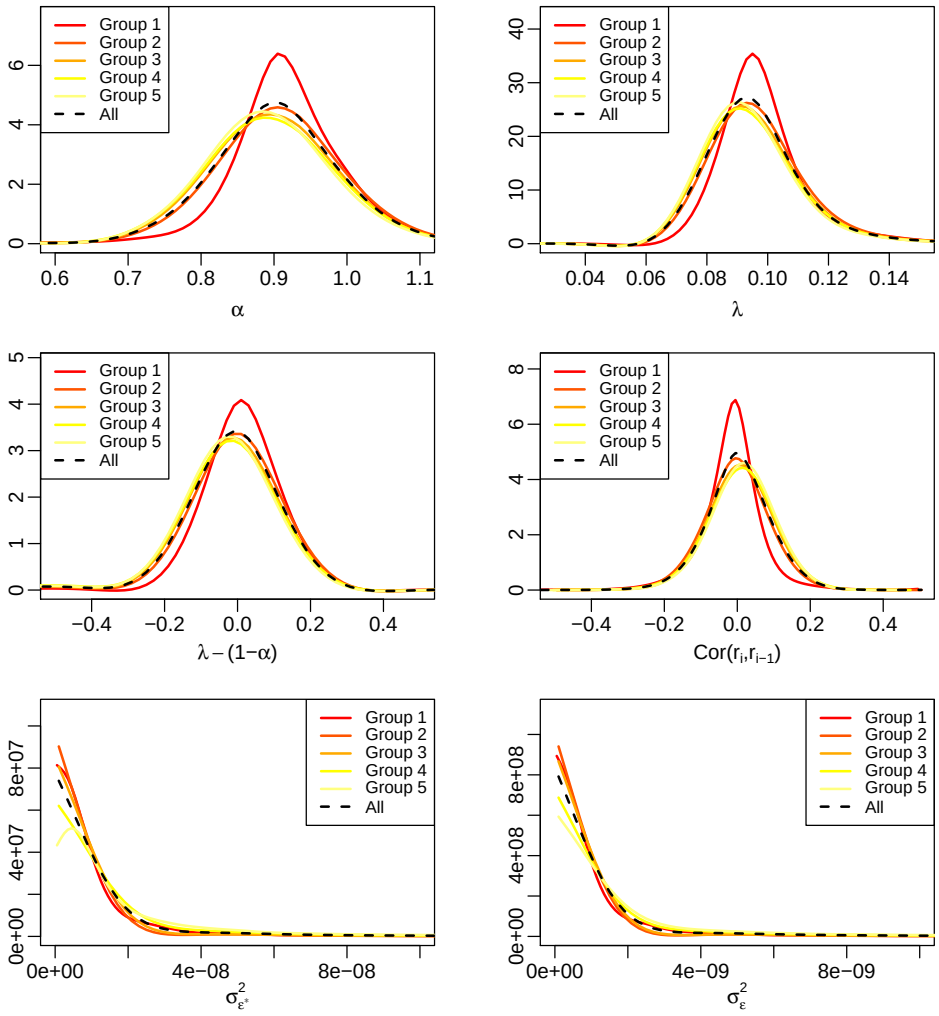
Figure 8 confirms that the stocks in Groups 2-5 generate nearly identical estimates, while those for Group 1 tend to deviate quite significantly from the others. In fact, the distribution for the four largest quintiles in the first two rows coincide to the extent that the curves effectively portray the identical distribution. For the variance of the efficient price and the noise component, minor deviations are visible, as the most liquid stocks (Groups 4-5) generate a slightly flatter distribution, with less mass at very small values and slightly thicker tails on the upside.

Specifically, for stocks in Groups 2-5, the estimates of α are approximately symmetric around a mode of slightly below 0.90, corroborating the hypothesis of a feedback effect and a mildly sluggish price adjustment. Likewise, the estimates of λ are unimodal, but centered on a value just below 0.10. This implies that the "serial correlation condition," $\hat{\lambda} - (1 - \hat{\alpha})$, has a mode near zero, implying very limited return autocorrelation on average. Nonetheless, this distribution is centered on a slightly negative value and displays some left-skewness. This induces a moderate tendency towards positive return serial correlation. This is consistent with the distribution of the first-order return autocorrelation in the middle right panel, which has a small positive mode and a right skew.

In contrast, the stocks in Group 1 generate $\hat{\alpha}$ and $\hat{\lambda}$ estimates that are somewhat larger and distinctly right skewed. That is, stocks with larger tick sizes and fewer quote revisions tend to generate returns that are relatively noisy and react more strongly to past deviations from the efficient price. Consequently, the associated distribution for the "serial correlation condition" has a positive mean and is slightly right-skewed. This is consistent with the absence of a right-skewed distribution for the realized return autocorrelation of Group 1 stocks.

Finally, we note that the estimates for the efficient variance are about tenfold the size of those for the noise variance, but both distributions feature a strong tilt towards low values and a much smaller right tail. It is evident that these two parameter estimates are highly correlated, generating a stable distribution for λ, featuring values predominantly within the $0.05 - 0.15$ range. Clearly, the noise component in the mid-quote returns is quite small. This property is in

Figure 8: The figure depicts the distribution of estimates for $\alpha, \lambda, \lambda - (1-\alpha)$, the return autocorrelation $\mathbb{C}or(r_i, r_{i-1})$, and the variances $\sigma^2_{\varepsilon^*}$, and σ^2_{ε}. The estimates are obtained from all $T = 10$ minute intervals, with intra-interval sampling at the frequency of $\Delta = 2$ seconds, across all 100 stocks in our sample during the first 40 trading days in 2014. The six curves reflect averages across stocks in each of the five groups built on the average daily number of mid-quote revisions, and all combined.



striking contrast to high-frequency transaction returns where the bid-ask bounce is dominantand generates a strong negative return serial correlation.

Overall, we find strong indications of time-variation in the local return dynamics, both from the distribution of the (model-free) return moments and from the distribution of the estimated model parameters. At the same time, the empirical distributions generally range across economically sensible and interpretable values. In Table 2 we provide summary statistics of these distributions for the stocks in Groups 2-5. We exclude Group 1 here to avoid confounding the evidence with the somewhat different features identified from these stocks. Moreover, to avoid excessive distortion from a few outliers, we eliminate all parameter estimates for intervals

where $\sigma^2_{\varepsilon^*}$ exceeds its 99 percentile, and for λ beyond three. The latter is a truly extreme realization given the mean of 0.115 and standard deviation of 0.17 for the 5-second sampling frequency in the truncated sample. The extreme estimates are closely related to unusual values for the realized return variation over the given short interval, so they serve as an effective filter for "abnormal" market conditions that may bias the general inference.[8]

As in Figure 8, Table 2 focuses on 10-minute observation windows, but we report results for 1-second and 5-second sampling along with the 2 seconds used in the figure. We note that the parameter distributions widen as the sampling gets courser and the number of intra-interval observations shrink. This suggests that sampling at the higher frequencies is beneficial. However, price discreteness and other microstructure features tend to have a stronger distortionary effect at high frequencies, which may inflate certain parameters, such as the noise variance estimates, and generate offsetting biases in other parameters. Hence, the use of slightly longer sampling intervals should speak to the robustness of the findings.

Table 2: The table reports percentiles, the mean and standard deviation for the parameter estimates and $\widehat{\mathbb{C}or}(r_i, r_{i-1})$. The statistics are based on 10-minute intervals over the first 40 trading days of 2014 and all stocks in Groups 2-5. Parameter estimates for intervals where $\hat{\sigma}^2_{\varepsilon^*}$ exceeds its 99 percentile or $\hat{\lambda} \leq 3$ are excluded.

$T = 10min, \Delta = 1sec$

	$q5$	$q25$	$Median$	$Mean$	$q75$	$q95$	SD
$\hat{\alpha}$	0.778	0.864	0.909	0.909	0.963	1.052	0.105
$\hat{\lambda}$	0.077	0.088	0.096	0.102	0.106	0.129	0.073
$\hat{\lambda} - (1 - \hat{\alpha})$	−0.150	−0.048	0.005	0.011	0.069	0.181	0.124
$\widehat{\mathbb{C}or}(r_i, r_{i-1})$	−0.144	−0.056	−0.000	−0.008	0.041	0.121	0.081
$\hat{\sigma}^2_{\varepsilon^*}$ $(\cdot 10^{-8})$	0.060	0.140	0.254	0.563	0.517	1.850	1.219
$\hat{\sigma}^2_{\varepsilon}$ $(\cdot 10^{-8})$	0.005	0.013	0.025	0.055	0.050	0.180	0.125

$T = 10min, \Delta = 2sec$

	$q5$	$q25$	$Median$	$Mean$	$q75$	$q95$	SD
$\hat{\alpha}$	0.726	0.846	0.905	0.891	0.961	1.056	0.138
$\hat{\lambda}$	0.072	0.086	0.095	0.104	0.106	0.138	0.101
$\hat{\lambda} - (1 - \hat{\alpha})$	−0.210	−0.070	−0.001	−0.004	0.066	0.188	0.158
$\widehat{\mathbb{C}or}(r_i, r_{i-1})$	−0.149	−0.053	0.000	0.002	0.058	0.151	0.091
$\hat{\sigma}^2_{\varepsilon^*}$ $(\cdot 10^{-8})$	0.116	0.279	0.516	1.177	1.058	3.875	2.670
$\hat{\sigma}^2_{\varepsilon}$ $(\cdot 10^{-8})$	0.009	0.026	0.048	0.118	0.101	0.394	0.287

$T = 10min, \Delta = 5sec$

	$q5$	$q25$	$Median$	$Mean$	$q75$	$q95$	SD

[8]Unusual σ^2_{ε} estimates typically arise from exceptionally tranquil market environments, where the efficient price variance $\sigma^2_{\varepsilon^*}$ is small and λ correspondingly large, or from intervals in which there is a large price change (jump) in the efficient price. The former scenario renders the effective sample small and parameter estimates unreliable over the interval, while the latter scenario violates the parameter constancy assumption, generating an inherent misspecification that renders the interpretation of estimation results problematic.

$\hat{\alpha}$	0.344	0.812	0.890	0.856	0.964	1.086	0.201
$\hat{\lambda}$	0.001	0.083	0.094	0.115	0.111	0.209	0.170
$\hat{\lambda} - (1 - \hat{\alpha})$	−0.433	−0.117	−0.017	−0.029	0.073	0.237	0.242
$\widehat{\mathbb{C}or}(r_i, r_{i-1})$	−0.177	−0.059	0.013	0.016	0.091	0.209	0.118
$\hat{\sigma}^2_{\varepsilon^*}$ $(\cdot 10^{-8})$	0.274	0.710	1.363	3.236	2.918	11.105	7.249
$\hat{\sigma}^2_{\varepsilon}$ $(\cdot 10^{-8})$	0.003	0.059	0.121	0.327	0.277	1.228	0.772

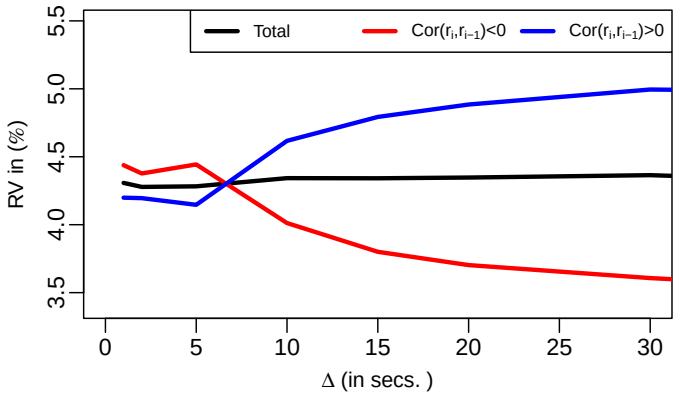
In fact, we do find a higher mean and median value for $\hat{\alpha}$ estimated at one second than for the lower frequencies. For the former, half of the estimates fall within $(0.86, 0.96)$ and 90% within $(0.77, 1.06)$. At the other extreme, for 5-second sampling, $(0.81, 0.97)$ covers half the estimates, while the $5\% - 95\%$ percentile span is dramatically expanded to encompass $(0.34, 1.09)$. In comparison, the median value for $\hat{\lambda}$ is remarkably stable at 0.094 to 0.096 across the sampling frequencies, although the identical inflation of the standard errors and associated variation in the point estimates is observed again, with the 90% range spanning $(0.07, 0.13)$ at one second and $(0.00, 0.21)$ at five seconds. This suggests that the noise variance for the quote midpoint typically is about only one-tenth of the innovation variance for the efficient price. Combined, these results imply that the distribution of the parametric serial correlation condition $\hat{\lambda} - (1 - \hat{\alpha})$ is centered on a slightly negative value and is mildly left-skewed. Consistent with this finding, as stipulated in Corollary 1, the empirical first-order return serial correlation is slightly positive on average and mildly skewed towards the right. Finally, for $\sigma^2_{\varepsilon^*}$ and σ^2_{ε} we confirm that the efficient price variance generally is about ten times larger than the noise variance. Moreover, the strong right skew noted in Figure 8 is apparent for both variables, even after our truncation of the more extreme outliers.

Another implication of our detection of local regimes with predominantly positive or negative autocorrelation, depending on the relative strength of the noise and price smoothing effects, is that the annualized realized volatility measures, locally, will be biased downward or upward relative to the efficient price variation, as outlined in Corollary 2. To assess whether our parametric model and estimation procedure captures such effects, Figure 9 depicts the realized volatility signature plot, obtained by averaging over all Group 2-5 stocks and 10-minute intervals. The traditional plot (black) is virtually flat from frequencies spanning one to thirty seconds, documenting that we avoid systematic microstructure biases in the volatility estimates, even at the highest frequencies we employ. In addition, the figure depicts the signature plots conditioned on the sign of the parametric serial correlation condition, $\hat{\lambda} - (1 - \hat{\alpha})$, with parameter estimates obtained at the five second frequency. If the sign is positive, the condition implies a negative return autocorrelation and associated downward sloping signature plot. This prediction is strongly supported by the plot. Likewise, when the sign is negative, we obtain the upward sloping volatility signature plot. Nonetheless, we caution that these realized variance measures

are strongly downward biased relative to the expected values in the sample, because we have eliminated some extreme return outliers through our filtering. As a consequence, we should not place much emphasis on the actual levels or relative height of the plots in the figure, but only the shapes of the individual curves.[9]

We conclude that our localized parametric model captures the state of the return moments quite robustly across a wide range of return frequencies, with direct implications for the accuracy of realized volatility measures as indicators of the underlying efficient price variation. In general, when the market enters the positive return autocorrelation state, associated with a substantial degree of smoothing of the fluctuations in the fundamental price, there is a systematic tendency to underestimate the underlying volatility through the standard realized volatility measure. The opposite scenario occurs when negative return autocorrelation, typically reflecting reversals of idiosyncratic noise shocks, are dominant, and the realized volatility measures tend to overstate the fundamental return volatility.

Figure 9: The figures shows the average realized variance against the sampling frequency Δ, over periods where $\hat{\lambda} > (1 - \hat{\alpha})$ (red curve), $\hat{\lambda} \leq (1 - \hat{\alpha})$ (blue curve), and all periods (black curve). The parameter estimates used to evaluate these conditions are based on five-second returns. The underlying realized returns are filtered to eliminate excessive outliers, as discussed in the main text and in Table 2.



We conclude this section by illustrating the correspondence between the parameter estimates and the return autocorrelations for individual 10-minute intervals. Figure 10 shows the time series plots of the estimated model parameters and the model-implied versus empirical first-order return autocorrelations based on 2-second sampling over one random trading day for Yahoo and Microsoft. The figure is quite representative for other stocks and trading days. First, we note that α fluctuates around the relationship $1 - \alpha = \lambda$ rendering the returns close to uncorrelated. In line with the summary statistics above, however, there are some significant deviations from

[9]We have confirmed that the slope of the two conditional plots are robust to the inclusion of all observations in the sample. In this case, all curves shift upward by a considerable amount, but retain their general shape.

this benchmark. In fact, we observe several distinct scenarios. The most dramatic is for Yahoo, in the interval from 20-30 minutes prior to the market close. The λ parameter is missing as the value exceeds 3, compared to an estimated value around or below 0.13 for the other 38 intervals. This further implies that we dismiss this interval from all subsequent analysis. Hence, the large drop in α to 0.4 for this particular interval – compared with estimates above 0.8 for all the other intervals – is not used in the sequel. This confirms that extreme estimates for the variance related quantities tend to generate (unreasonable) outliers for the remainder of the model parameters as well.

For examples of contrasting scenarios, consider the Microsoft intervals ending respectively in minute 210 and 270. In the former case, we have an increase in the smoothing effect (α drops) and less noise (λ declines), so we have two separate reinforcing factors contributing towards the observed positive spike in the return serial correlation coefficient. In the latter case, we observe a smaller positive spike in the return autocorrelation, as a sharp increase in the smoothing effects (drop in α) is partially offset by a marked increase in the noise variance (λ increases). A noise shock can arise from the arrival of large idiosyncratic and uninformed transactions, while the increased sluggishness may stem from a shift in the market environment that fosters increased uncertainty about the fundamental value of the asset.

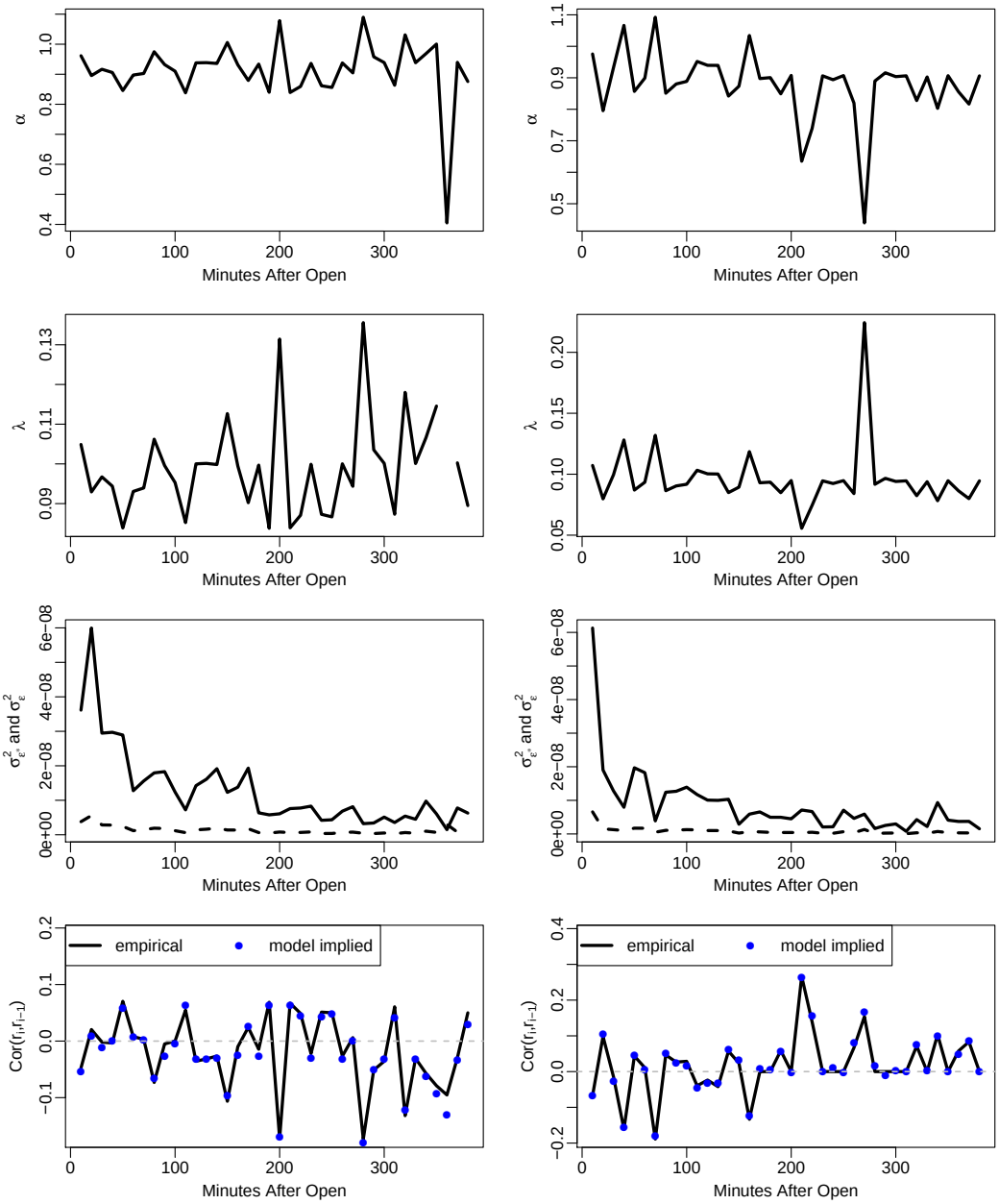
We further note the striking correspondence between the model-implied and empirical return autocorrelations in the bottom row of the figure. It shows that our model parameters can explain the vast majority of the variation in the observed return autocorrelation. In some sense, the model parameters translate the realized high-frequency return dynamics into economically interpretable quantities that we can relate to information and frictional sources. In Section 7, we explore the implications of a more general model specification that allows us to assess the performance of the current model relative to alternative parsimonious representations that are popular in the wider literature. Finally, for both stocks we note that the "fundamental" variance $\sigma^2_{\varepsilon^*}$ and the noise variance σ^2_{ε} follow a distinct intraday pattern with elevated variances during morning trading. This feature is representative of a typical intraday pattern and will be analyzed below, along with the intraday periodicities of α and λ.

6.3 Parameter Variation and Time-of-Day Effects

As discussed in Section 6.2, we observe considerable time variation in the model parameters α and λ. In this section, we explore whether time-of-day effects play a significant role in generating this feature.

Figure 11 reports averages (across stocks and days) of $\hat{\alpha}$, $\hat{\lambda}$, $\hat{\sigma}^2_{\varepsilon^*}$, $\hat{\sigma}^2_{\varepsilon}$, and $\widehat{\mathbb{C}or}[r_i, r_{i-1}]$ per time-of-day, obtained across different sampling frequencies Δ for a window length of $T = 10$ minutes. The horizontal axis depicts the time in minutes after the market open at 9:30.

Figure 10: Time series plots of estimated model coefficients $\hat{\alpha}$, $\hat{\lambda}$, $\hat{\sigma}^2_{\varepsilon^*}$, $\hat{\sigma}^2_{\varepsilon}$, and model-implied and empirical first-order return autocorrelations. Yahoo (left column) and Apple (right column), 6th trading day 2014. $\Delta = 2$ secs and $T = 10$ min.



The intraday plot for $\hat{\sigma}^2_{\varepsilon^*}$ reveals that an L-shape pattern of high volatility in the morning followed by a decay throughout the day is a systematic feature across stocks and time. Thus, the elevated volatility components observed for a couple of individual trading days in Section 6.2 are typical. Our results imply that the fundamental volatility is extraordinarily high during the first few trading minutes, which likely stems from price discovery associated with information

processing of overnight news and orders. Interestingly, the noise variance $\hat{\sigma}_{\varepsilon}^2$ displays the identical qualitative pattern, suggesting that both the informational and noise variance is at a peak in the morning. In fact, the L-shape is even more pronounced for $\hat{\sigma}_{\varepsilon}^2$, yielding generally higher noise-to-signal ratios $\hat{\lambda}$ in early trading. The relative elevation of noise in the morning likely reflects caution among market intermediaries and traders as they anticipate the high level of fundamental price uncertainty after the open, leading to widening spreads and lower effective market depth.[10]

The speed of price reversals, α, also tends to be slightly elevated after the open, when information is processed during the first trading minutes. This suggests a fairly rapid incorporation of fundamental news after the open. Subsequently, α declines and then shows a mild increase towards the market close, generating a distorted u-pattern. The increase late in the day may be due to an increase in liquidity associated with a rise in (uninformed) trading as market participants seek to attain their desired positions prior to the market close. This is also consistent with a minor elevation in λ towards the market close, as an increase in non-informational trading generates additional price pressure, but no permanent innovation in the fundamental price.

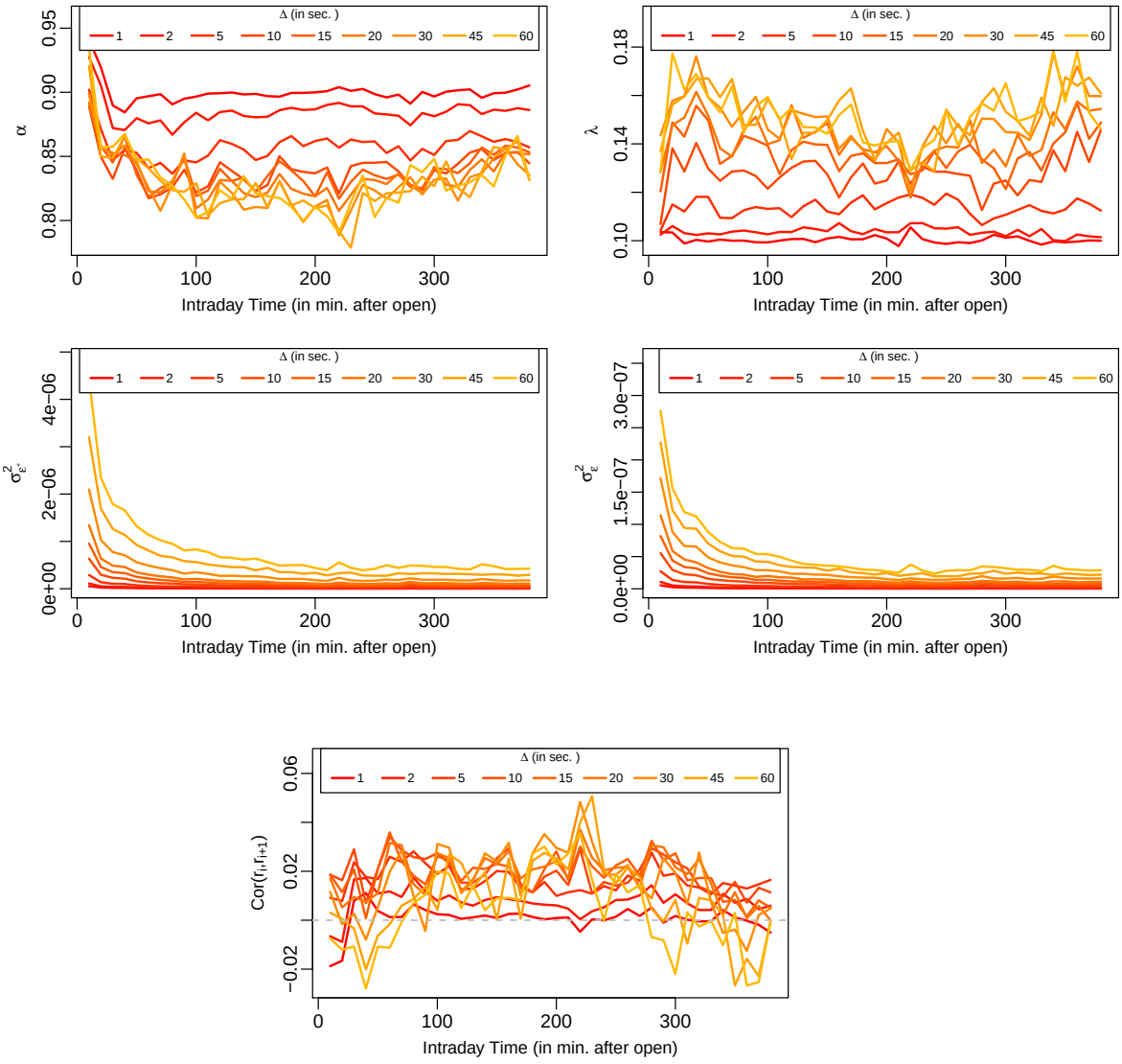
In theory, the systematic intraday variation in α and λ governs the sign of the return autocorrelation, see, e.g., Corollary 1. In fact, we do detect weak intraday patterns in the actual first-order return autocorrelations. They form an inverted U-shape, associated with lower (more negative) serial correlation following the market opening and before the close, and higher (positive) autocorrelations during mid-day. This pattern is in accordance with rapid price discovery, and diminished sluggishness in prices, during morning trading, which renders the elevated noise component dominant compared to the price sluggishness. Later on, prices turn more sluggish (α drops) while the noise shocks are mitigated, leading to a reversal where sluggish adjustment to fundamental news becomes dominant and induces a degree of momentum in prices. Finally, the end of the trading day sees lots of trading, and prices likely react swiftly to fundamental news, leading the noise term to become the main source of serial correlation in mid-quotes, inducing a tendency towards "contrarian" price dynamics.

Overall, we find evidence corroborating the idea that liquid market conditions early and late in the trading day foster efficient discovery and processing of fundamental news. At such times, the main component behind temporary misalignments in prices is MMN, consistent with negative return serial correlation. In contrast, when the information and order flow is slower, the resources devoted to price discovery shrink and it may be more difficult to identify the less prevalent news items among the ever-present noise. This tends to enhance price sluggishness,

[10]Even if it is not transparent from the figure, the intraday shape of $\hat{\sigma}_{\varepsilon^*}^2$ and $\hat{\sigma}_{\varepsilon}^2$ is broadly similar across all sampling frequencies. The cascading effect, arising from the higher variance levels for the larger values of Δ, is a natural byproduct of aggregation, as price innovations and noise shocks cumulate over longer horizons.

generating conditions that favor an element of momentum in the price formation.

Figure 11: The figure depicts the intraday averages of the parameter estimates for Δ ranging from 1 to 60 seconds and $T = 10$ minutes. The underlying realized returns are filtered to eliminate excessive outliers, as discussed in the main text and in Table 2.



Beyond the intraday pattern, the shift in the level of the parameter estimates across the sampling frequencies is also noteworthy. As observed in Table 2, we find α declining and λ increasing in Δ. To assess this feature, we provide the temporal aggregation properties for both parameters in the following theorem, under the strong assumption that our underlying model is correctly specified at a high reference frequency for all stocks and at all times. Moreover, for tractability we impose explicit distributional assumptions.

Theorem 2. *Assume the price dynamics are given by equations* (1) *and* (2)*. Moreover, let the*

observed price process p_i, sampled at step size $\Delta \geq 1$, with observations for $i = 0, \Delta, 2 \cdot \Delta, \ldots T - \Delta$, governed by the following relations,

$$p_{i+\Delta} = p_i - \alpha_\Delta (p_i - p_i^*) + \varepsilon_{i+\Delta,\Delta}, \qquad p_{i+\Delta}^* = p_i^* + \hat{\varepsilon}_{i+\Delta,\Delta}^* \tag{17}$$

where $\varepsilon_{i+\Delta,\Delta} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_{\varepsilon,\Delta})$ and $\hat{\varepsilon}^*_{i+\Delta,\Delta} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_{\varepsilon^*,\Delta})$ for all $i \in \{k \cdot \Delta, k = 0, 1, 2, 3, \ldots\}$.

Then, for $\Delta \geq 1$ and $0 < \alpha < 1$, we have,

$$\alpha_\Delta = 1 - (1 - \alpha)^\Delta \tag{18}$$

$$\sigma^2_{\varepsilon,\Delta} = g_\varepsilon \sigma^2_\varepsilon + g_{\varepsilon^*} \sigma^2_{\varepsilon^*} \tag{19}$$

$$\sigma^2_{\varepsilon^*,\Delta} = \Delta \cdot \sigma^2_{\varepsilon^*}, \tag{20}$$

with

$$g_\varepsilon = \frac{1}{\alpha(2-\alpha)} \left\{ 1 - (1-\alpha)^{2\Delta} \right\}, \tag{21}$$

$$g_{\varepsilon^*} = \frac{1}{\alpha(2-\alpha)} \left\{ 4(1-\alpha)^\Delta - (1-\alpha)^{2\Delta} + \alpha \left(2 \left(1 - (1-\alpha)^\Delta \right) + (2-\alpha)\Delta \right) - 3 \right\}. \tag{22}$$

Proof. See Appendix. □

Equation 18 implies that α_Δ is strictly increasing in Δ, with $\lim_{\Delta \to \infty} \alpha_\Delta = 1$. Thus, any given pricing error should be mitigated over time and, for sufficiently long horizons, it vanishes altogether. In addition, Theorem 2 allows us to derive the noise-to-signal ratio λ_Δ for models estimated at lower frequencies ($\Delta > 1$) as a function of the noise-to-signal ratio of the baseline model ($\Delta = 1$),

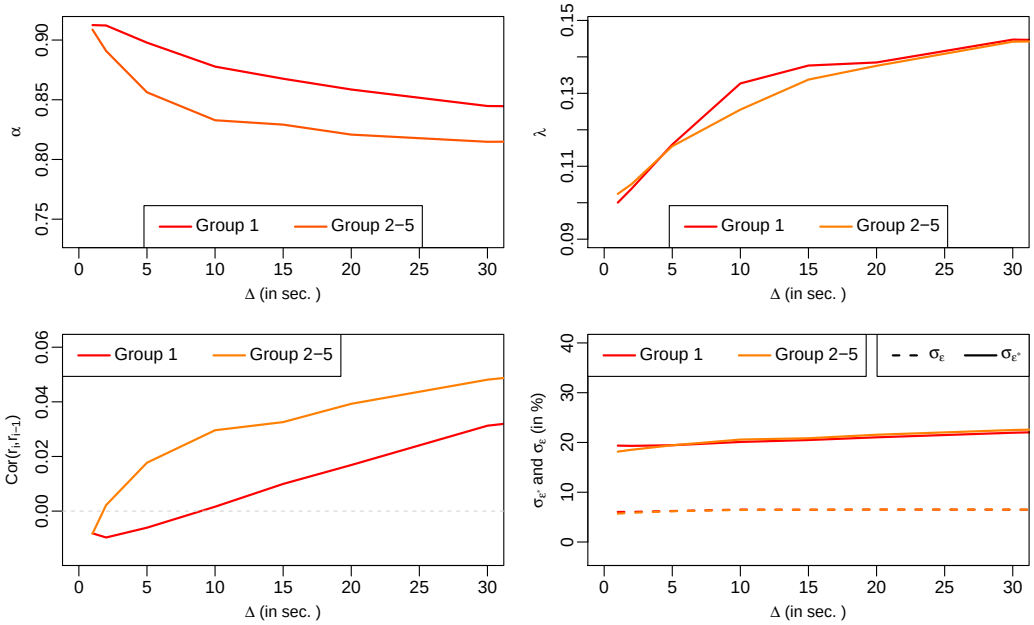
$$\lambda_\Delta = \frac{\sigma^2_{\varepsilon,\Delta}}{\sigma^2_{\varepsilon^*,\Delta}} = \frac{1}{\Delta} \left(g_\varepsilon \lambda + g_{\varepsilon^*} \right). \tag{23}$$

As $\lim_{\Delta \to \infty} g_\varepsilon = \frac{1}{\alpha(2-\alpha)}$ and $\lim_{\Delta \to \infty} g_{\varepsilon^*}/\Delta = \frac{\alpha(2-\alpha)}{\alpha(2-\alpha)} = 1$, we have $\lim_{\Delta \to \infty} \lambda_\Delta = 1$. Hence, the "noise variance" in the aggregated model, $\sigma^2_{\varepsilon,\Delta}$, actually constitutes a mixture of the noise and informational variance of the baseline model, with the unit root component associated with efficient price shocks being the sole determinant of the longer term price variation. Therefore, at a more aggregate level, $\varepsilon_{i+\Delta,\Delta}$ does not represent a pure noise innovation, but also incorporates the cumulative impact of micro-granular feedback stemming from underlying shifts in the fundamental price.

As observed, these aggregation results provide a theoretical benchmark under the assumption that there is a unique "baseline frequency" for which the stipulated model, defined through

equations (1) – (2), applies uniformly across all stocks and at all times. In turn, this simple dynamic at very high frequencies then triggers the low-frequency dynamics outlined in the theorem. The predicted positive association between λ and Δ is readily observed in Figure 12, suggesting that the basic mechanism of the model is at work for the highest frequencies and thus imbues the noise variance σ_ε^2 with the cumulative effect of innovations to the efficient price.

Figure 12: Average estimates for stocks in Group 1 and Group 2-5 of α, λ, $\sigma_{\varepsilon^*}^2$, σ_ε^2 and $\mathbb{C}or(r_i, r_{i-1})$ for different sampling frequencies Δ and $T = 10$ minutes. The underlying realized returns are filtered to eliminate excessive outliers, as discussed in the main text and in Table 2.



For estimates of α, however, the theoretical aggregation results are violated. Although the α estimates do start increasing for large values of Δ (not reported), we observe the opposite tendency for low values of Δ. This is not truly surprising. It is evident that the model parameters vary over time, for example, due to much higher return variance components in the early hours of trading. We would expect variation across stocks as well. In fact, we already noted that stocks in Group 1 behave systematically different from the remainder. We also anticipate the flow of fundamental information to fluctuate across days, rendering factors such as trading activity, information acquisition, order book depth and bid-ask spreads variable as well. Hence, it is impossible for any single baseline frequency to provide valid inference for all stock-time pairings. We further note from Figure 11 that the strength of the feedback is especially strong for low Δ's, suggesting that high-frequency trading algorithms might dominate at these frequencies. This points towards heterogeneous groups of traders, strategies, or algorithms, operating at different frequencies and inducing feedback effects over intervals of different lengths. That is, the observed relationship between α and Δ may be driven by an overlay of distinct pricing error

corrections, each playing out at a different pace. In addition, Table 2 reveals the presence of a substantial number of outliers towards the left tail of the distribution of estimated α's at the five second frequency, implying a lack of robustness in the mean estimates. Thus, it is evident that we should not expect our simple local approximation to deliver clean aggregation results.

These features also manifest themselves in the bottom row of Figure 12. The increasing return autocorrelation across frequencies is consistent with the drop in α for the longer sampling intervals, as it implies enhanced temporal smoothing of the underlying return fluctuations. The (annualized) variance estimates are almost flat, but mildly upward sloping. For the noise variance, this is expected as the efficient price innovations slowly build into this component. For the fundamental variance, this is likely due to heterogeneity in the baseline frequency of the model, as the extent of some of these fundamental shocks only become identifiable at lower frequencies. Yet, the relative stability in these (annualized) variance parameters suggests that qualitative interpretation of the results remain meaningful despite the evidence of some misspecification.

7 A Model Generalization

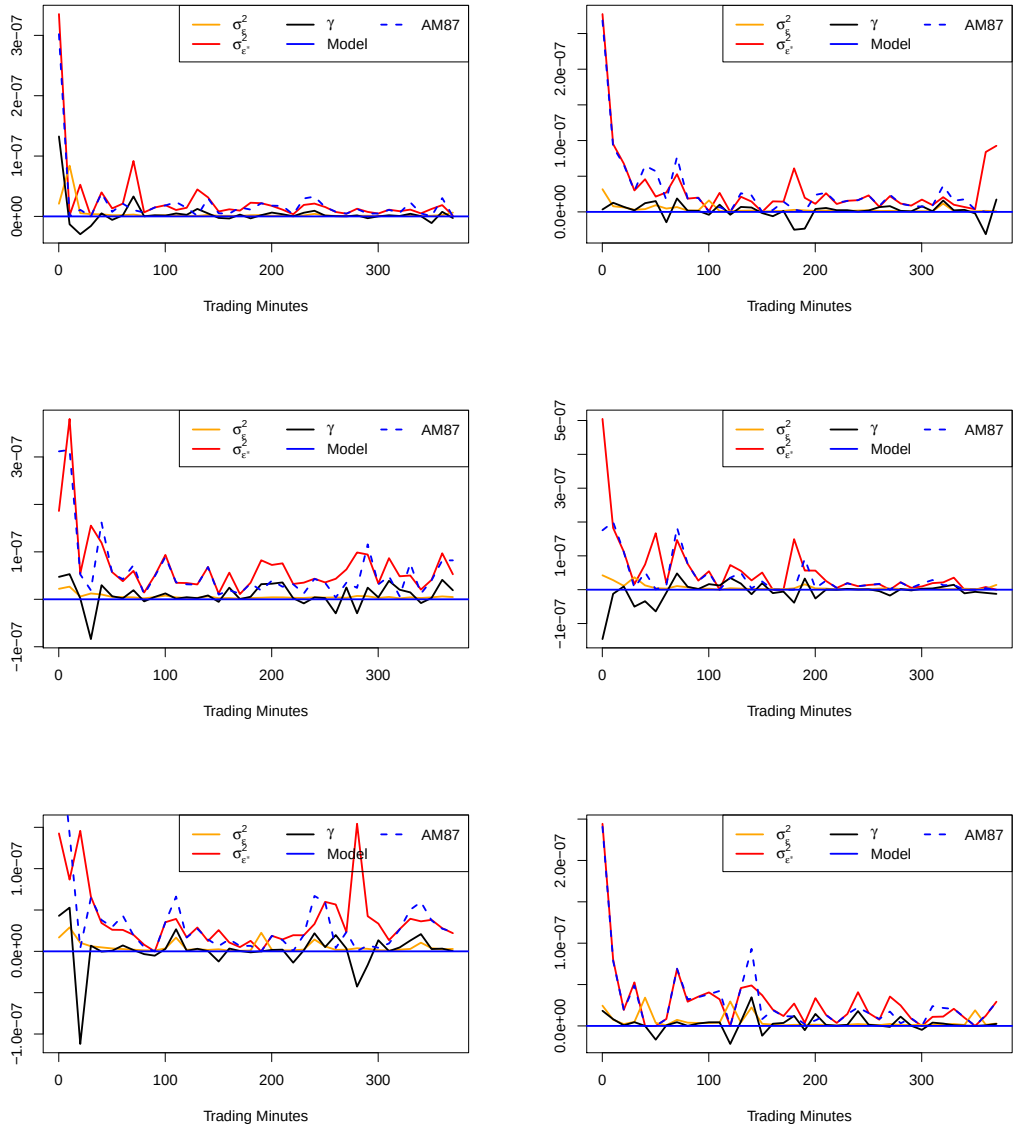
A useful generalization of the model (1)-(2) is obtained by relaxing the uncorrelatedness of ε_i and ε_i^*. In fact, by assuming $\mathbb{E}[\varepsilon_i \varepsilon_i^*] = \gamma$, we enable the model to nest important special cases. In particular, if $\gamma = \alpha \sigma_{\varepsilon^*}^2$, the model can be re-written as

$$p_i = p_{i-1} - \alpha(p_{i-1} - p_i^*) + \tilde{\varepsilon}_i, \tag{24}$$

with the zero mean innovation term $\tilde{\varepsilon}_i := \varepsilon_i - \alpha \varepsilon_i^*$ and $\mathbb{V}[\tilde{\varepsilon}_i] = \sigma_\varepsilon^2 - \alpha^2 \sigma_{\varepsilon^*}^2$. Moreover, one may readily check that the innovation in the efficient price, ε_i^*, is uncorrelated with the innovation term in the observed mid-quote price, i.e., $\mathbb{E}[\tilde{\varepsilon}_i \varepsilon_i^*] = 0$.

Consequently, the price dynamics in equation (24) mimics the one in equation (1), except that the partial price adjustment is based on the discrepancy between the lagged observed price and *concurrent* efficient price. Economically, this updating mechanism is motivated by the idea that traders update their notion of the efficient price instantaneously and modify their quotes (or transaction prices) accordingly. This representation of the information feedback mechanism has been proposed by Amihud and Mendelson (1987), where it was applied for *daily* returns. In that context, updating based on the *current* efficient price is plausible compared to one-day old information. However, in the high-frequency environment, featuring incessant revisions to the order book, near continuous news feeds, and multiple trades per second, it is less plausible that traders have the capability to gauge the efficient price every instant. Instead, we assume

Figure 13: Figure depicts some of the estimation results of the generalized model for the first two trading days of 2014 for AAPL (first panel), AMZN (second panel), and MSFT (third panel) The left figure corresponds to the first trading day; the second figure corresponds to the second trading day. Results are shown for $T = 10$ minutes and $\Delta = 10$ seconds. We report the estimates of the noise variance and the efficient noise variance (σ^2_ε and $\sigma^2_{\varepsilon^*}$) and the estimate of $\lambda = \mathbb{E}(\varepsilon_i \varepsilon^*_i)$. $\gamma = 0$ corresponds to our initial model, while $\gamma = \alpha\sigma^2_{\varepsilon^*}$ corresponds to partial adjustment according to Amihud and Mendelson (1987).



that updates regarding the fundamental price occur with a slight, and potentially time-varying, delay. In other words, we hypothesize that our base model ($\gamma = 0$) is more suitable at high frequencies (small Δ), while model (24) is more plausible for large Δ. A major advantage of our generalized model is that both updating mechanisms can be tested straightforwardly via simple restrictions on γ.

Another special case of great interest arises for $\gamma = \sigma^2_{\varepsilon^*}$, corresponding to $\alpha = 1$ in the parameterization above. This leads to the "classical" random-walk-plus-iid-noise model, see, e.g., Ait-Sahalia et al. (2005) and Bandi and Russell (2008)),

$$p_i = p_i^* + \varepsilon_i$$

with $\mathbb{E}[\varepsilon_i \varepsilon_i^*] = 0$. As discussed in Section 2, this formulation does not capture the endogeneity of the noise which, in our specification, arises naturally from temporal back-shifts stemming from the presence of the lagged pricing error in the adjustment mechanism.

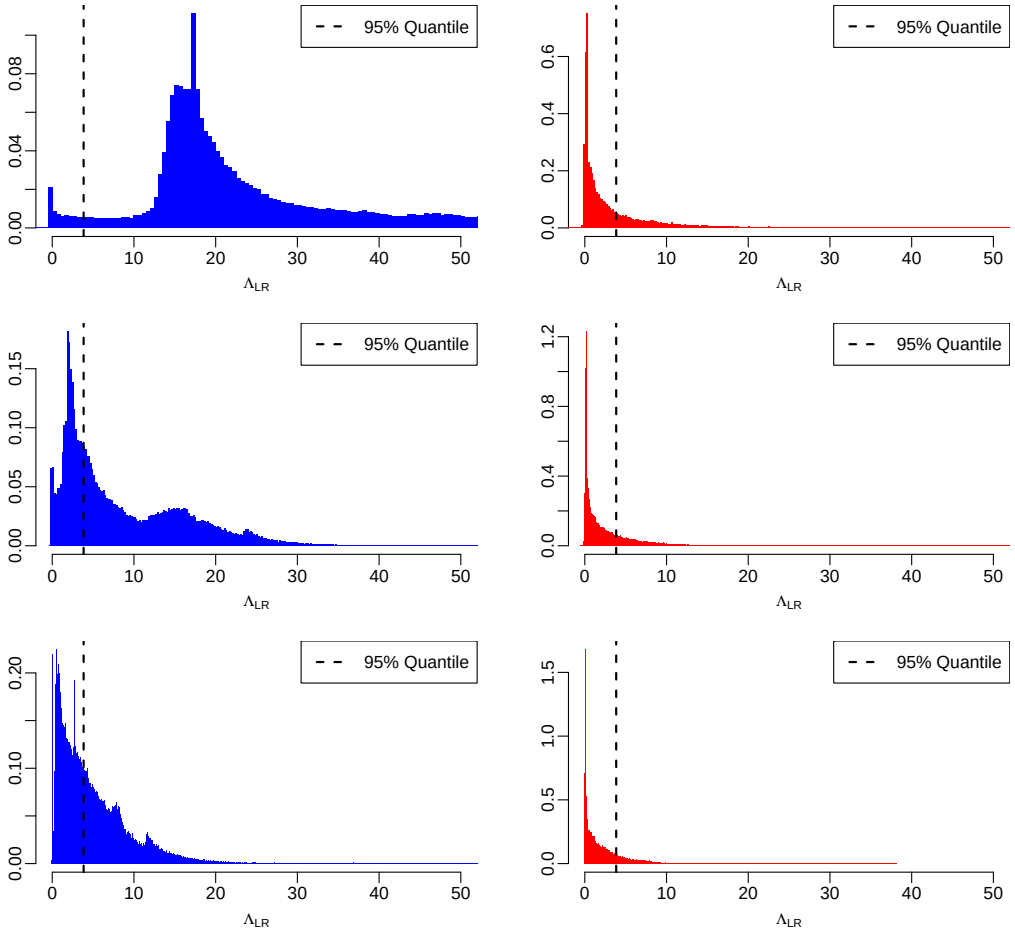
The generalized model is readily estimated from the approach outlined in Section 5. However, it is necessary to modify the covariance matrix of $w_i = (\varepsilon_i^\mu, 0, \varepsilon_i)$ as follows,

$$\Sigma_w = \begin{pmatrix} \sigma^2_\varepsilon + \sigma^2_{\varepsilon^*} - 2\alpha\sigma^2_{\varepsilon^*} & 0 & \sigma^2_\varepsilon - \alpha\sigma^2_{\varepsilon^*} \\ 0 & 0 & 0 \\ \sigma^2_\varepsilon - \alpha\sigma^2_{\varepsilon^*} & 0 & \sigma^2_\varepsilon \end{pmatrix}.$$

Figure 13 displays the time series of estimates for γ, the quantity $\alpha\sigma^2_{\varepsilon^*}$ as well as σ^2_ε and $\sigma^2_{\varepsilon^*}$ for the first two trading days of 2014 for Apple, Amazon, and Microsoft using ten-minute estimation windows and sampling every ten seconds. These illustrations are generally representative of the features observed over other trading days and stocks. It is evident that the estimates of γ vary around zero. In contrast, the restriction imposed by the Amihud and Mendelson (1987) model, namely the equality of γ and $\alpha\sigma^2_{\varepsilon^*}$ is strongly violated due to erratic fluctuations in the latter series. This suggest that our base model, assuming $\gamma = 0$, and thus an error correction mechanism working through the *lagged* pricing errors, is more in line with the high-frequency evidence than the specification adopted in Amihud and Mendelson (1987). Clearly, the identical verdict carries over to the random-walk-plus-iid-noise model, which further requires $\gamma = \sigma^2_{\varepsilon^*}$, or $\alpha = 1$ within the Amihud and Mendelson (1987) representation. Again, the $\sigma^2_{\varepsilon^*}$ estimate fluctuates much too widely compared to the estimate for γ.

To generate a more formal statistical test of these restrictions, we estimate the generalized model with an unconstrained value for $\mathbb{E}[\varepsilon_i \varepsilon_i^*] = \gamma$, and then again with respectively the Amihud and Mendelson (1987) constraint, $\gamma = \alpha\sigma^2_{\varepsilon^*}$, and the constraint of our base model, $\gamma = 0$, imposed. Next, we perform likelihood ratio tests on all available local intervals. Figure 14 depicts the cross-sectional (across local windows and stocks) distribution of the resulting likelihood ratio statistics, $\Lambda_{LR} = 2(\ln \mathcal{L}_U - \ln \mathcal{L}_R)$, where $\ln \mathcal{L}_U$ denotes the log-likelihood value of the generalized (unrestricted) model and $\ln \mathcal{L}_R$ denotes the log likelihood value under the restriction $\gamma = \alpha\sigma^2_{\varepsilon^*}$ in the upper panel and the restriction $\gamma = 0$ in the lower panel. Consistent with the suggestive evidence from Figure 13, we find that the proportion of rejections

Figure 14: Figure shows the likelihood ratio statistics for the hypothesis $H_0 : \gamma = \alpha\sigma^2_{\varepsilon^*}$ (left panel) and $H_0 : \gamma = 0$ (right panel) based on estimates for all stocks and all local windows for $T = 10min$ and $\Delta = 2$ seconds (top), $\Delta = 5$ seconds (middle) and $\Delta = 10$ seconds (bottom).



for the Amihud and Mendelson (1987) model is much higher than our base model. Moreover, the rejection frequencies for our base line model is largely independent of Δ, while the likelihood ratio statistics decline markedly for larger values of Δ in the case of the Amihud and Mendelson (1987) model. These findings are consistent with our conjecture that an updating mechanism based on $(p_{i-1} - p^*_i)$ is inappropriate for high sampling frequencies, but becomes more tenable as Δ increases. Nonetheless, we also acknowledge that the drop in the rejection rate likely also, in part, is associated with a general loss in power arising from less precise inference at the lower sampling frequencies.

Table 3 complements the evidence by reporting the average and median values of Λ_{LR} and corresponding rejection frequencies at the 5% significance level. For $\Delta = 2$ seconds, the Amihud and Mendelson (1987) model is rejected in over 96% of the available intervals, while the base model with $\gamma = 0$ is rejected in only 27% of all cases. These numbers decline

Table 3: The table reports summarizes the results for the likelihood-ratio test statistics $\Lambda_{LR} = 2(\ln \mathcal{L}_U - \ln \mathcal{L}_R)$ for the hypothesis $H_0 : \gamma = \alpha\sigma^2_{\varepsilon^*}$ and $H_0 : \gamma = 0$ based on estimates for all 100 NASDAQ index constituents over the first 40 trading days in 2014, using 39 10-minute estimation windows per day (sample size $N = 152,000$). The table reports the mean and median of Λ_{LR} and the percentage of rejections at the 5-% level for $\Delta = 2sec$, $\Delta = 5sec$ and $\Delta = 10sec$.

	$H_0 : \gamma = \alpha\sigma^2_{\varepsilon^*}$			$H_0 : \gamma = 0$		
Δ	Mean Λ_{LR}	Median Λ_{LR}	reject (in %)	Mean Λ_{LR}	Median Λ_{LR}	reject (in %)
2	19.412	25.127	96.5	1.249	3.494	27.0
5	5.656	8.686	62.3	1.122	2.563	23.0
10	3.744	4.991	49.0	1.162	2.108	18.2

to 49% and 18%, respectively, for $\Delta = 10$ seconds. Finally, we reiterate that our decisive rejection of the Amihud and Mendelson (1987) restriction, $\gamma = \alpha\sigma^2_{\varepsilon^*}$, also implies rejection of the random-walk-plus-iid-noise model, as it constitutes a special case of the former. The fact that the hypothesis $\gamma = 0$ provides a superior approximation to the data suggests that the endogeneity of the market microstructure noise primarily stems from temporal feedback effects.

8 Conclusions

This paper provides a novel perspective on "classical" martingale-plus-noise models for high-frequency mid-quote prices, as commonly used in the literature. The proposed modification of this model class by just one additional parameter is simple but powerful, as it yields a more structural perspective and opens up canonical links to existing approaches in market microstructure theory. We show that well-known properties of market microstructure noise, such as serial correlation as well as correlation with the efficient price ('endogeneity'), arise naturally from temporal feedback due to noise-induced mispricing.

We identify error correction due to prevailing mispricing as a key source of high-frequency dynamics in financial returns. The interplay of the speed of price adjustment and the noise-to-signal ratio governs the autocovariance structure of intraday returns and is shown to have a profound effect on realized volatilities sampled on high frequencies.

Serial correlation in high-frequency returns may therefore be rationalized by the extent of sluggishness in prices and the relative noisiness (or informativeness) of the price innovations. Our empirical analysis shows that the interplay between these two key parameters varies over time and triggers local intervals of 'momentum' behavior (positive autocorrelations) and 'contrarian' trading (negative autocorrelations). An important empirical finding is that particularly regimes of statistically significant positive serial correlations in high-frequency

mid-quote returns (inducing downward sloped realized volatility signature plots) occur more often than commonly expected.

We moreover show that feedback effects based on *lagged* efficient prices are clearly more in line with the high-frequency return dynamics than updating based on *current* efficient prices, as proposed by Amihud and Mendelson (1987) for daily returns. We show that the latter updating mechanism is more appropriate for higher aggregation levels but is soundly rejected at high frequencies. An immediate consequence is that the "classical" martingale-plus-iid-noise model is clearly rejected as well. We conclude that temporal feedback, as captured by the model developed in this paper, may be a critical determinant for the dynamics of high-frequency returns.

The proposed framework allows us to establish closer linkages between high-frequency based volatility estimation and market microstructure analysis and it may be expanded in various directions. From a statistical viewpoint, interesting extensions include a relaxation of the locally constant parameterization or an introduction of more general updating mechanisms, opening up possibilities for the construction of more efficient volatility estimators in a generalized framework.

Finally, our analysis identifies the speed of pricing error correction, or the sluggishness of the price adjustment, as a key parameter driving high-frequency return dynamics and as a natural measure for informational efficiency in the market. Interesting avenues for future research therefore include the cross-sectional variation in this sluggishness parameter and its links to liquidity characteristics of the assets.

References

AIT-SAHALIA, Y. AND J. JACOD (2014): *High-Frequency Financial Econometrics*, Princeton University Press.

AIT-SAHALIA, Y., P. A. MYKLAND, AND L. ZHANG (2005): "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise," *Review of Financial Studies*, 17, 351–416.

AMIHUD, Y. AND H. MENDELSON (1987): "Trading Mechanisms and Stock Returns: An Empirical Investigation," *Journal of Finance*, 42 (3), 533–553.

ANDERSEN, T. G., T. BOLLERSLEV, F. DIEBOLD, AND P. LABYS (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

ANDERSEN, T. G., T. BOLLERSLEV, AND F. X. DIEBOLD (2010): "Parametric and Non-

Parametric Volatility Measurement," in *Handbook of Financial Econometrics*, ed. by Y. Ait-Sahalia and L. P. Hansen, Elsevier, 67–137.

ANDREWS, D. (1991): "Heteroscedasticity and Autocorrelation Consistent Covaraince Matrix Estimation," *Econometrica*, 59, 817–858.

BANDI, F. AND J. RUSSELL (2008): "Microstructure Noise, Realized Variance, and Optimal Sampling," *Review of Economic Studies*, 75, 339–369.

BIBINGER, M., N. HAUTSCH, P. MALEC, AND M. REISS (2014): "Estimating the quadratic covariation matrix from noisy observations: local method of moments and efficiency," *Annals of Statistics*, 42, 1312–1364.

CHAKER, S. (2013): "Volatility and Liquidity Costs," Tech. Rep. 2013-29, Bank of Canada.

DIEBOLD, F. AND G. STRASSER (2013): "On the Correlation Structure of Microstructure Noise: A Financial Economic Approach," *Review of Economic Studies*, 80, 13041337.

GOURIEROUX, C., A. MONFORT, AND A. TROGNON (1984): "Pseudo maximum likelihood methods: theory," *Econometrica*, 52, 681–700.

HANSEN, L. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

HANSEN, P. R. AND A. LUNDE (2006): "Realized Variance and Market Microstructure Noise," *Journal of Business and Economic Statistics*, 24, 127–161.

HARVEY, A. C. (1989): *Forecasting. Structural Time Series Models and the Kalman Filter.*, Cambridge University Press.

KALNINA, I. AND O. LINTON (2008): "Estimating Quadratic Variation Consistently in the Presence of Endogenous and Diurnal Measurement Error," *Journal of Econometrics*, 147, 47–59.

KYLE, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53 (6), 1315–1335.

LI, Y., S. XIE, AND X. ZHENG (2016): "Efficient Estimation of Integrated Volatility Incorporating Trading Information," Tech. rep.

MYKLAND, P. AND L. ZHANG (2009): "Inference for Continuous Semimartingales Observed at High Frequency," *Econometrica*, 77, 1403–1445.

NEWEY, W. AND K. WEST (1987): "A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covaraince Matrix," *Econometrica*, 55, 703–708.

ROLL, R. (1984): "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *Journal of Finance*, 1127–1140.

SHEPPARD, K. (2013): "Measuring Market Speed," Tech. rep., University of Oxford.

VIVES, X. (1995): "The Speed of Information Revelation in a Financial Market Mechanism," *Journal of Economic Theory*, 67, 187–204.

ZHANG, L., P. MYKLAND, AND Y. AIT-SAHALIA (2005): "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 13941411.

Appendix

Proofs

For general reading, the next lemma can be skipped and only serves as a technical support for the results in the next sections.

Lemma A. The cross-moments of ε_i and μ_i obey

(i) $\mathbb{E}[\mu_{i-h}\varepsilon_i] = 0 \qquad \forall h > 0,$

(ii) $\mathbb{E}[\mu_{i+h}\varepsilon_i] = (1-\alpha)^h\sigma_\varepsilon^2 \qquad \forall h \geq 0.$

Proof. Observe that because of equation (3), we have $\mu_{i+1} = -\alpha\mu_i + \varepsilon_{i+1}^\mu$. Hence, through backward recursion, we obtain the moving average representation for μ_{i+h} and μ_i, i.e.,

$$\mu_{i+h} = (1-\alpha)^{h+1}\mu_{i-1} + \sum_{j=0}^{h}(1-\alpha)^j\varepsilon_{i+h-j}^\mu, \tag{25}$$

$$\mu_i = (1-\alpha)^i\mu_0 + \sum_{j=0}^{i}(1-\alpha)^j\varepsilon_{i-j}^\mu. \tag{26}$$

Further, note that the innovation to the mispricing term is defined as $\varepsilon_i^\mu = \varepsilon_i - \varepsilon_i^*$ and that ε_i and ε_i^* are mutually independent, i.i.d. random noise processes. Therefore,

$$\mathbb{E}[\varepsilon_i\varepsilon_k^\mu] = \delta(i-k)\sigma_\varepsilon^2, \tag{27}$$

where δ denotes the Dirac-Delta function. The lemma now follows from straightforward calculations.

(i) Using equation (26), we have for $i > 0$ and $h > 0$

$$\mathbb{E}[\mu_{i-h}\varepsilon_i] = (1-\alpha)^{i-h}\underbrace{\mathbb{E}[\varepsilon_0^{\mu}\varepsilon_i]}_{=0} + \sum_{j=0}^{i-h}(1-\alpha)^j\mathbb{E}[\varepsilon_{i-h-j}^{\mu}\varepsilon_i]. \tag{28}$$

The last terms vanishes because of equation (27), as $i-h-j \leq i-h < i$.

(ii) From equation (25), we obtain

$$\mathbb{E}[\mu_{i+h}\varepsilon_i] = (1-\alpha)^{h+1}\underbrace{\mathbb{E}[\mu_{i-1}\varepsilon_i]}_{=0} + \sum_{j=0}^{h}(1-\alpha)^j\mathbb{E}[\varepsilon_i\varepsilon_{i+h-j}^{\mu}]. \tag{29}$$

The first term is zero because of property (i). For the second term, equation (27) implies,

$$\sum_{j=0}^{h}(1-\alpha)^j\mathbb{E}[\varepsilon_i\varepsilon_{i+h-j}^{\mu}] = (1-\alpha)^h\sigma_{\varepsilon}^2. \tag{30}$$

□

Proof of Lemma 1. We have, by definition,

$$\mathbb{C}ov(r_{i+h}, r_i) = \mathbb{E}[(r_{i+h} - \underbrace{\mathbb{E}[r_{i+h}]}_{=0})(r_i - \underbrace{\mathbb{E}[r_i]}_{=0})] = \mathbb{E}[r_{i+h}r_i] = \mathbb{E}[(p_{i+h} - p_{i+h-1})(p_i - p_{i-1})]. \tag{31}$$

Now, using, $p_i = \mu_i + p_i^*$, we obtain,

$$\mathbb{C}ov(r_{i+h}, r_i) = \mathbb{E}[(\mu_{i+h} - \mu_{i+h-1} + p_{i+h}^* - p_{i+h-1}^*)(\mu_i - \mu_{i-1} + p_i^* - p_{i-1}^*)]. \tag{32}$$

Then exploit equation (3) to establish,

$$\begin{aligned}
\mathbb{C}ov(r_{i+h}, r_i) &= \mathbb{E}\Big[\big(-\alpha\mu_{i+h-1} + \underbrace{\varepsilon_{i+h}^{\mu} + \varepsilon_{i+h}^*}_{=\varepsilon_{i+h}}\big)\big(-\alpha\mu_{i-1} + \underbrace{\varepsilon_i^{\mu} + \varepsilon_i^*}_{=\varepsilon_i}\big)\Big] \\
&= \mathbb{E}\Big[\big(-\alpha\mu_{i+h-1} + \varepsilon_{i+h}\big)\big(-\alpha\mu_{i-1} + \varepsilon_i\big)\Big] \\
&= \alpha^2\,\mathbb{E}[\mu_{i+h-1}\mu_{i-1}] - \alpha\mathbb{E}[\mu_{i+h-1}\varepsilon_i] - \alpha\underbrace{\mathbb{E}[\varepsilon_{i+h}\mu_{i-1}]}_{=0} + \underbrace{\mathbb{E}[\varepsilon_{i+h}\varepsilon_i]}_{=0},
\end{aligned} \tag{33}$$

where we exploited Lemma A, and the fact that ε_i is i.i.d.

It follows that,

$$\mathbb{C}ov(r_{i+h}, r_i) = \alpha^2\,\mathbb{E}[\mu_{i+h-1}\mu_{i-1}] - \underbrace{\alpha\mathbb{E}[\mu_{i+h-1}\varepsilon_i]}_{=(1-\alpha)^{h-1}\sigma_\varepsilon^2}. \tag{34}$$

The first term is given in equation (25). The two terms may then be evaluated using equation (5) and Lemma A to yield,

$$\begin{aligned}
\mathbb{C}ov(r_{i+h}, r_i) &= (\sigma_{\varepsilon^*}^2 + \sigma_\varepsilon^2)\alpha^2\frac{(1-\alpha)^h}{1-(1-\alpha)^2} - \alpha(1-\alpha)^{h-1}(\sigma_{\varepsilon^*}^2 + \sigma_\varepsilon^2) + \alpha(1-\alpha)^{h-1}\sigma_{\varepsilon^*}^2 \\
&= \underbrace{\alpha(1-\alpha)^{h-1}\sigma_{\varepsilon^*}^2}_{=\psi(h-1)\cdot(2-\alpha)}(1+\lambda)\left(\underbrace{\left(\frac{\alpha-\alpha^2}{1-(1-\alpha)^2} - \frac{1-(1-\alpha)^2}{1-(1-\alpha)^2}\right)}_{=-\frac{1}{2-\alpha}} + \frac{1}{1+\lambda}\right) \\
&= \psi(h-1)\Big((1-\alpha) - \lambda\Big).
\end{aligned} \tag{35}$$

□

Proof of Theorem 1. Using the identity $p_i = \mu_i + p_i^*$ from equation (3), to substitute for p_i below, we obtain,

$$\begin{aligned}
\langle p\rangle_T^\Delta &= \mathbb{E}\left[\sum_{j=0}^{T/\Delta-1}(r_{j\Delta}^\Delta)^2\right] = \mathbb{E}\left[\sum_{j=0}^{T/\Delta-1}(p_{j\Delta+\Delta} - p_{j\Delta})^2\right]. \\
&= \mathbb{E}\Big[\sum_{j=0}^{T/\Delta-1}(\mu_{j\Delta+\Delta} - \mu_{j\Delta} + p_{j\Delta+\Delta}^* - p_{j\Delta}^*)^2\Big].
\end{aligned}$$

Now, exploiting the recursive structure of μ_i via the identity $\mu_{j\Delta+\Delta} = (1-\alpha)^\Delta\mu_{j\Delta} + \sum_{k=0}^{\Delta-1}(1-\alpha)^k\varepsilon_{j\Delta+\Delta-k}^\mu$, along with equation (3) and $p_{j\Delta+\Delta}^* - p_{j\Delta}^* = \sum_{k=1}^{\Delta}\varepsilon_{j\Delta+k}^*$, we rewrite this equation as,

$$\begin{aligned}
\langle p\rangle_T^\Delta &= \sum_{j=0}^{T/\Delta-1}\mathbb{E}\Big[(\tilde{A}_j + \tilde{B}_j + \tilde{C}_j)^2\Big] \\
&= \sum_{j=0}^{T/\Delta-1}\Big\{\mathbb{E}[\tilde{A}_j^2] + \mathbb{E}[\tilde{B}_j^2] + \mathbb{E}[\tilde{C}_j^2] + 2\underbrace{\mathbb{E}[\tilde{A}_j\tilde{B}_j]}_{=0} + 2\underbrace{\mathbb{E}[\tilde{A}_j\tilde{C}_j]}_{=0} + 2\mathbb{E}[\tilde{B}_j\tilde{C}_j]\Big\},
\end{aligned}$$

where the terms satisfy $\tilde{A}_j = \big((1-\alpha)^\Delta - 1\big)\mu_{j\Delta}$, $\tilde{B}_j = \sum_{k=0}^{\Delta-1}(1-\alpha)^k\varepsilon_{j\Delta+\Delta-k}^\mu$ and $\tilde{C}_j = \sum_{k=1}^{\Delta}\varepsilon_{j\Delta+k}^*$. We then exploit, from Lemma A, that $\mu_{j\Delta}$ is independent of future noise

terms, implying $\mathbb{E}[\tilde{A}_j\tilde{B}_j]=0$ and $\mathbb{E}[\tilde{A}_j\tilde{C}_j]=0$, i.e.,

$$\langle p\rangle_T^{\Delta} = \sum_{j=0}^{T/\Delta-1}\Big\{\mathbb{E}[\tilde{A}_j^2]+\mathbb{E}[\tilde{B}_j^2]+\mathbb{E}[\tilde{C}_j^2]+2\mathbb{E}[\tilde{B}_j\tilde{C}_j]\Big\}.$$

The first three terms can be rewritten, using a couple of definitions and equation (5), $\mathbb{E}[(\varepsilon_i^{\mu})^2]=\sigma_{\mu}^2$, $\mathbb{E}[\varepsilon_i^2]=\sigma_{\varepsilon}^2$ and $\mathbb{E}[\mu_{j\Delta}^2]=\sigma_{\mu}^2/(\alpha(2-\alpha))$. Moreover, since $\mathbb{E}[\varepsilon_k^*\varepsilon_l^{\mu}]=-\sigma_{\varepsilon^*}^2\delta_{kl}$, the cross-term satisfies $\mathbb{E}[\tilde{B}_j\tilde{C}_j]=\mathbb{E}[\big(\sum_{k=0}^{\Delta-1}(1-\alpha)^k\varepsilon_{j\Delta+\Delta-k}^{\mu}\big)\big(\sum_{k=1}^{\Delta}\varepsilon_{j\Delta+k}^*\big)\Big]=-\sigma_{\varepsilon^*}^2\sum_{k=1}^{\Delta}(1-\alpha)^{k-1}$, where δ_{kl} denotes the Kronecker-symbol, i.e., $\delta_{kl}=1$ if $k=l$ and $\delta_{kl}=0$ otherwise. Thus, our equation takes the form,

$$\begin{aligned}
\langle p\rangle_T^{\Delta} &= \sum_{j=0}^{T/\Delta-1}\Big\{\big((1-\alpha)^{\Delta}-1\big)^2\frac{\sigma_{\mu}^2}{\alpha(2-\alpha)}+\underbrace{\sum_{k=0}^{\Delta-1}(1-\alpha)^{2k}}_{=\frac{1-(1-\alpha)^{2\Delta}}{1-(1-\alpha)^2}}\sigma_{\mu}^2+\Delta\sigma_{\varepsilon^*}^2-2\sigma_{\varepsilon^*}^2\underbrace{\sum_{k=1}^{\Delta}(1-\alpha)^{k-1}}_{=\frac{1-(1-\alpha)^{\Delta}}{\alpha}}\Big\}\\
&= T\,\sigma_{\varepsilon^*}^2+\frac{T}{\Delta}\Big\{\big((1-\alpha)^{\Delta}-1\big)^2\frac{\sigma_{\mu}^2}{\alpha(2-\alpha)}+\frac{1-(1-\alpha)^{2\Delta}}{1-(1-\alpha)^2}\sigma_{\mu}^2-2\sigma_{\varepsilon^*}^2\frac{1-(1-\alpha)^{\Delta}}{\alpha}\Big\}\\
&= T\,\sigma_{\varepsilon^*}^2+T\Big\{2\sigma_{\mu}^2\frac{1-(1-\alpha)^{\Delta}}{\alpha(2-\alpha)\Delta}-2\sigma_{\varepsilon^*}^2\frac{1-(1-\alpha)^{\Delta}}{\alpha\Delta}\Big\}\\
&= T\,\sigma_{\varepsilon^*}^2+T\underbrace{\frac{2(1-(1-\alpha)^{\Delta})}{\alpha\Delta}}_{=\phi(\Delta)}\frac{\big(\sigma_{\mu}^2-\sigma_{\varepsilon^*}^2(2-\alpha)\big)}{(2-\alpha)}\\
&= T\,\sigma_{\varepsilon^*}^2+T\phi(\Delta)\Big\{\sigma_{\varepsilon}^2+\sigma_{\varepsilon^*}^2-\sigma_{\varepsilon^*}^2(2-\alpha)\Big\}.
\end{aligned}$$

Now, substitute the noise-to-signal ratio $\lambda=\sigma_{\varepsilon}^2/\sigma_{\varepsilon^*}^2$ into the equation,

$$\langle p\rangle_T^{\Delta} = T\,\sigma_{\varepsilon^*}^2\;+\;T\sigma_{\varepsilon^*}^2\cdot\phi(\Delta)\cdot\Big\{\lambda-(1-\alpha)\Big\}.$$

□

Proof of Theorem 2. The observed price at time $i+\Delta$ satisfies,

$$p_{i+\Delta}=p_{i+\Delta-1}-\alpha(\mu_{i+\Delta-1})+\varepsilon_{i+\Delta}, \tag{36}$$

with $\mu_{i+\Delta-1}=p_{i+\Delta-1}-p_{i+\Delta-1}^*$. Then, by recursion we obtain,

$$p_{i+\Delta}=p_i-\alpha\sum_{j=0}^{\Delta-1}\mu_{i+j}+\sum_{j=1}^{\Delta}\varepsilon_{i+j}. \tag{37}$$

Now, we use the AR recursion property of μ_i, i.e., $\mu_{i+j} = (1-\alpha)^j \mu_i + \sum_{k=0}^{j-1}(1-\alpha)^k \varepsilon^{\mu}_{i+j-k}$ and substitute it into (37) to obtain,

$$\begin{aligned} p_{i+\Delta} &= p_i - \underbrace{\alpha \sum_{j=0}^{\Delta-1}(1-\alpha)^j}_{:=\alpha_\Delta} \cdot \underbrace{\mu_i}_{=p_i - p_i^*} - \underbrace{\alpha \sum_{j=0}^{\Delta-1}\sum_{k=0}^{j-1}(1-\alpha)^k \varepsilon^{\mu}_{i+j-k} + \sum_{j=1}^{\Delta}\varepsilon_{i+j}}_{:=\varepsilon_{i+\Delta,\Delta}} \\ &= p_i - \alpha_\Delta (p_i - p_i^*) + \varepsilon_{i+\Delta,\Delta}. \end{aligned} \tag{38}$$

This gives the observed price dynamics of the *aggregated* model, sampled on step sizes Δ. It is formally equivalent with our initial model, where,

$$\alpha_\Delta = \alpha \sum_{j=0}^{\Delta-1}(1-\alpha)^j = 1 - (1-\alpha)^\Delta. \tag{39}$$

The error terms $\varepsilon_{i+\Delta,\Delta}$ are i.i.d. normally distributed error terms with zero mean. Independence originates from the fact that $\varepsilon_{i+\Delta,\Delta}$ is based on non-overlapping partial sums of ε_i and ε_i^*, thus preserving independence. The distribution properties derive directly from the corresponding properties of the primal noise terms ε_i and ε_i^*.
To compute $\sigma^2_{\varepsilon,\Delta} := \mathbb{V}\text{ar}(\varepsilon_{i+\Delta,\Delta})$, we re-write $\varepsilon_{i+\Delta,\Delta}$ as,

$$\begin{aligned} \varepsilon_{i+\Delta,\Delta} &= -\alpha \sum_{j=1}^{\Delta-1} \underbrace{\sum_{k=0}^{j-1}(1-\alpha)^k}_{=\frac{1}{\alpha}(1-(1-\alpha)^j)} \varepsilon^{\mu}_{i+\Delta-j} + \sum_{j=1}^{\Delta}\varepsilon_{i+j} \\ &= -\sum_{j=1}^{\Delta-1}\left\{1-(1-\alpha)^j\right\}\varepsilon^{\mu}_{i+\Delta-j} + \sum_{j=1}^{\Delta}\varepsilon_{i+j} \\ &= -\sum_{j=0}^{\Delta-1}\left\{1-(1-\alpha)^j\right\}(\varepsilon_{i+\Delta-j} - \varepsilon^*_{i+\Delta-j}) + \sum_{j=0}^{\Delta-1}\varepsilon_{i+\Delta-j} \\ &= \sum_{j=0}^{\Delta-1}\left\{(1-\alpha)^j\right\}\varepsilon_{i+\Delta-j} + \sum_{j=0}^{\Delta-1}\left\{1-(1-\alpha)^j\right\}\varepsilon^*_{i+\Delta-j}. \end{aligned} \tag{40}$$

Thus, exploiting the i.i.d. property of the primal noise terms, we obtain the *aggregated* noise variance as,

$$\begin{aligned} \sigma^2_{\varepsilon,\Delta} &= \sigma^2_\varepsilon \underbrace{\sum_{j=0}^{\Delta-1}(1-\alpha)^{2j}}_{=g_\varepsilon} + \sigma^2_{\varepsilon^*} \underbrace{\sum_{j=0}^{\Delta-1}(1-(1-\alpha)^j)^2}_{=g_{\varepsilon^*}} \\ &= g_\varepsilon \sigma^2_\varepsilon + g_{\varepsilon^*}\sigma^2_{\varepsilon^*}, \end{aligned} \tag{41}$$

with

$$g_{\varepsilon} = \frac{1}{\alpha(\alpha-2)}\Big\{(1-\alpha)^{2\Delta}-1\Big\}, \tag{42}$$

$$g_{\varepsilon^*} = \Delta + \frac{1}{\alpha(\alpha-2)}\Big\{3-2\alpha-2(1-\alpha)^{\Delta+1}+(1-\alpha)^{2\Delta}\Big\}. \tag{43}$$

The aggregated efficient variance $\sigma^2_{\varepsilon^*,\Delta} := \mathbb{V}\text{ar}(\varepsilon^*_{i+\Delta,\Delta})$ is $\sigma^2_{\varepsilon^*,\Delta} = \mathbb{V}\text{ar}(\sum_{j=1}^{\Delta}\varepsilon^*_{i+j}) = \Delta\sigma^2_{\varepsilon^*}$. □

Tables and Figures

Figure 15: Empirical distribution of parameter estimates for the different quote-revision groups for $\Delta = 2sec$. Results are based on $\Delta = 2sec$ and varying intra-day estimation period T. The left column represents $T = 5min$, the middle column represents $T = 10min$ and the right column represents $T = 30min$.

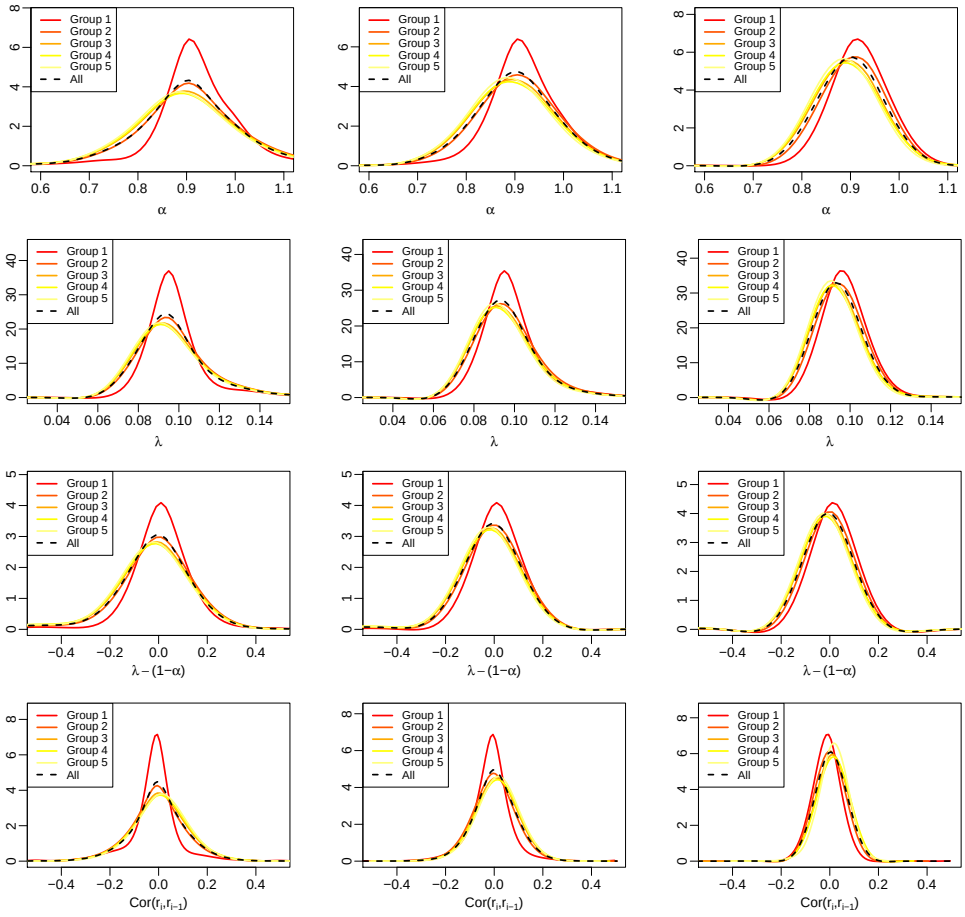


Figure 16: Empirical distribution of parameter estimates for the different quote-revision groups for $\Delta = 5sec$. Results are based on $\Delta = 5sec$ and varying intra-day estimation period T. The left column represents $T = 5min$, the middle column represents $T = 10min$ and the right column represents $T = 30min$.

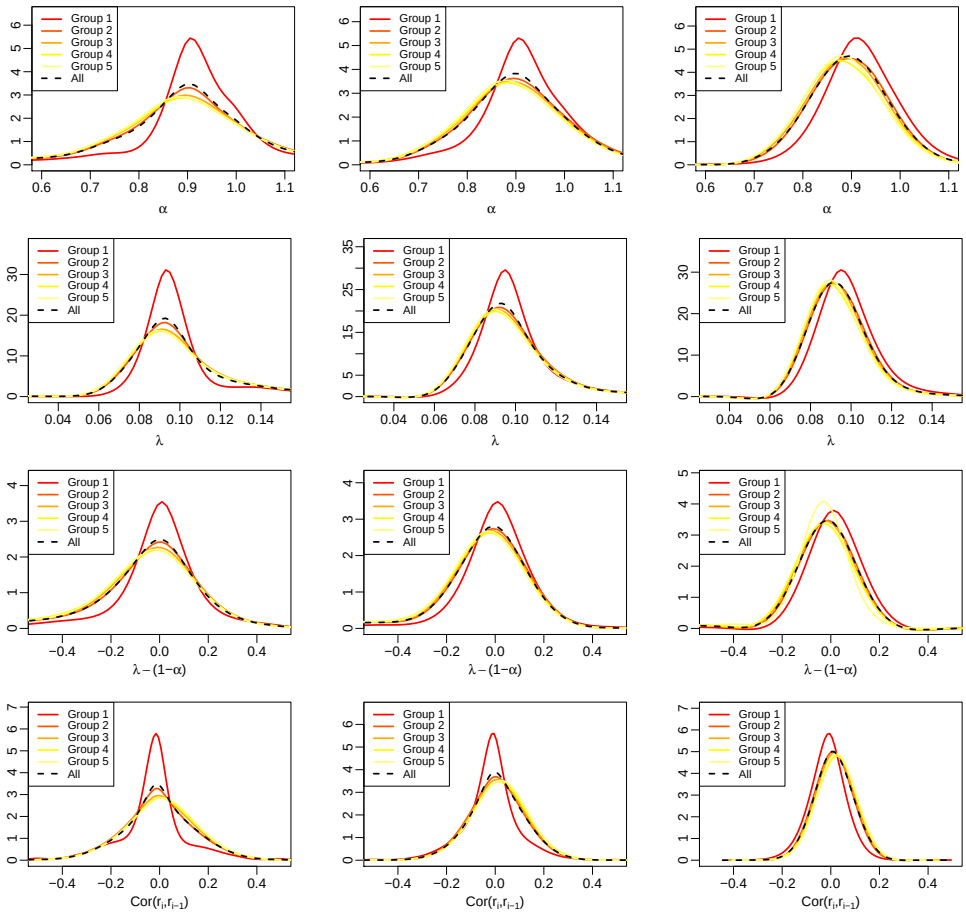


Figure 17: Empirical distribution of parameter estimates for the different quote-revision groups for $\Delta = 10sec$. Results are based on $\Delta = 10sec$ and varying intra-day estimation period T. The left column represents $T = 5min$, the middle column represents $T = 10min$ and the right column represents $T = 30min$.

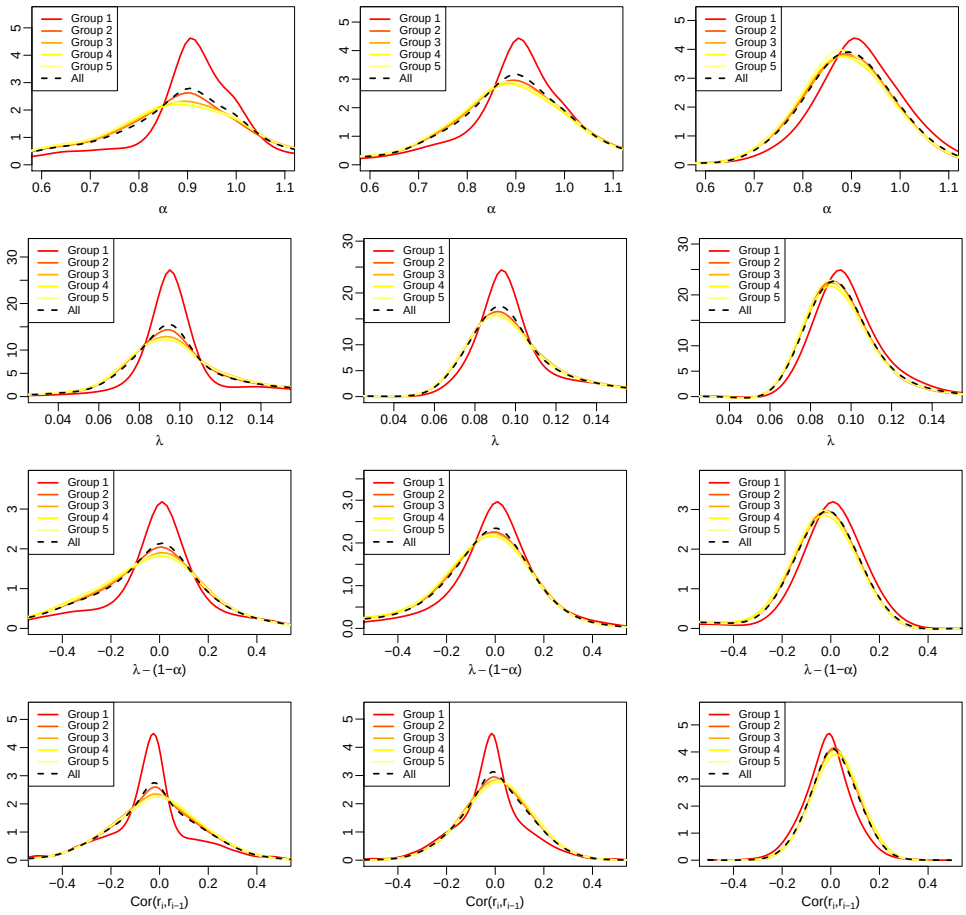


Table 4: The table shows group-specific averages of stock-specific time averages. The groups are built based on quintiles on the average number of daily mid-quote updates. We record time averages of the mid-quote price in dollars ($Price$), the relative tick size in basis points ($Tick$, computed as the inverse of $Price$), the spread ($Spread$) in cents, the spread in basis points, the depth on top of the book ($Depth$) in shares, and the daily trade volume ($TrVol$) in shares and dollar. The table also provides group-specific averages of daily mid-quote revisions (Nrv), the average number of daily order messages ($Nmsg$), and the average first-order return autocorrelation.

	Stock Characteristics							Stock Activity		Corr
Revision Activity Groups	Price (in $)	Tick (in bps)	Spread (in cents)	Spread (in bps)	Depth (in shares)	TrVol (in mio. shares)	TrVol (in mio USD)	Nrv (in 10^4)	$Nmsg$ (in 10^5)	ρ
Group 1	26	5.3	1.2	6.5	20357	3	63.4	0.211	2.367	0.001
Group 2	59.2	1.9	3	4.5	988	1	39.8	0.613	1.414	0.027
Group 3	68.4	1.6	3.4	4.4	575	0.8	50.1	0.871	1.477	0.026
Group 4	99.9	1.1	7.8	6.3	385	0.7	56.8	1.118	1.22	0.031
Group 5	280.1	0.8	19.3	6.1	368	1.5	241.2	1.994	1.609	0.03
All	106.7	2.1	7	5.6	4535	1.4	90.3	0.961	1.617	0.023

CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

Recent Issues

All CFS Working Papers are available at *www.ifk-cfs.de*.

No.	Authors	Title
568	Joshua R. Goldstein, Christos Koulovatianos, Jian Li, Carsten Schröder	*Evaluating How Child Allowances and Daycare Subsidies Affect Fertility*
567	Roman Kräussl, Thorsten Lehnert, Denitsa Stefanova	*The European Sovereign Debt Crisis: What Have We Learned?*
566	Luiz Felix, Roman Kräussl, and Philip Stork	*Single stock call options as lottery tickets*
565	Luiz Felix, Roman Kräussl, and Philip Stork	*Implied Volatility Sentiment: A Tale of Two Tails*
564	Roman Kräussl and Elizaveta Mirgorodskaya	*The Winner's Curse on Art Markets*
563	Christiane Baumeister, Reinhard Ellwanger and Lutz Kilian	*Did the Renewable Fuel Standard Shift Market Expectations of the Price of Ethanol?*
562	Thierry Mayer, Marc Melitz and Gianmarco Ottaviano	*Product Mix and Firm Productivity Responses to Trade Competition*
561	Jeronimo Carballo, Gianmarco Ottaviano and Christian Volpe Martincus	*The Buyer Margins of Firms' Exports*
560	Gianmarco Ottaviano, Giovanni Peri and Greg Wright	*Immigration, Trade and Productivity in Services: Evidence from U.K. Firms*

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de