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

Preliminary Version – Please do not circulate and do not cite

Torben Andersen[*]
Northwestern University

Gökhan Cebiroğlu[†]
University of Vienna

Nikolaus Hautsch[‡]
University of Vienna
Center for Financial Studies

January 10, 2016

Abstract

In this paper, we propose a generalization of the classical ”martingale-plus-noise” model for prices on a high-frequency level. In our framework, observed prices are driven by market microstructure noise and the prevailing discrepancy between observed prices and efficient prices. The speed by which observed prices adjust to the ”mis-pricing” component is governed by a parameter which naturally measures the price efficiency on the market. The framework provides a structural approach for the linkage between observed prices and efficient prices and the role of feedback effects in trading. It allows to capture a wide range of stochastic behavior in observed micro prices and captures the classical ”martingale plus i.i.d. noise" framework as a special case.

We illustrate that the variance of the mis-pricing components is naturally linked to a measure of the market efficiency and demonstrate that its interplay with the signal-to-noise ratio determines two major regimes in the market: when the signal-to-noise ratio is low relative to the variance of mis-pricing, observed returns become negatively autocorrelated and observed prices tend to revert towards efficient prices. In the opposite case, observed returns become positively correlated and observed prices reveal local trends in line with momentum behavior. By locally estimating the model based on NASDAQ limit order book data, we empirically identify the presence of these two regimes in terms of the model parameters and show that they are in line with autocorrelations in observed returns.

We moreover show that the two regimes imply different behavior of realized volatility estimators sampled on high frequencies. While in one regime resulting volatility-signature-plots are upward sloped for increasing sampling frequencies, a reverse pattern prevails in

[*]Kellog School of Management, *email*: t-andersen@kellog.northwestern.edu.

[†]Department of Statistics and Operations Research, Oskar-Morgenstern-Platz 1, A-1090 Vienna, *email*: goekhan.cebiroglu@univie.ac.at.

[‡]Department of Statistics and Operations Research, Oskar-Morgenstern-Platz 1, A-1090 Vienna, *email*: nikolaus.hautsch@univie.ac.at.

the other. We provide empirical evidence on local volatility-signature plots and demonstrate substantial time variations of the corresponding shapes. Our results therefore provide new insights for high-frequency based volatility estimation, provide new channels for the construction of improved estimators and establish a direct link to underlying market microstructure literature.

JEL classification: C58, C32, G14

Keywords: High-Frequency Data, Volatility Estimation, Market Microstructure Noise, Trade Reversals

1 Introduction

A major theme in the recent financial econometrics literature is how to (efficiently) back out estimators of daily or intra-daily volatility from financial high-frequency data. A common methodological starting point is to assume that micro-level prices (or quotes) stem from an underlying (unobservable) semi-martingale process which, however, is polluted by noise. The latter is due to market microstructure effects, such as the minimum tick size making prices to move on a fixed grid, the bid-ask bounce effect or liquidity effects causing observed micro-level prices to deviate from an underlying (semi-)martingale process. A wide range of studies considered the problem of optimally utilizing noisy high-frequency information to efficiently estimate the quadratic variation of the underlying ”efficient” price process. For an overview, see, e.g., the survey by Andersen et al. (2010) or the recent book by Ait-Sahalia and Jacod (2014).

While most studies treat market microstructure noise as a discrete-time process which is equipped with certain stochastic properties, only very few studies make the attempt to link the properties of noise more explicitly to the process of trading and the underlying microstructure. Notable exceptions are Diebold and Strasser (2013) and Chaker (2013).

In this paper, we propose a new model for high-frequency prices generalizing the framework utilized in the aforementioned literature. The major idea is to assume that changes in observed prices are caused by two different sources of randomness. First, in line with the classical literature, market prices are disturbed by random noise arising due to market microstructre frictions or noise trading behavior which is unrelated to information. Such ”non-informational” shocks cause ”mis-pricing” as observed prices and efficient prices deviate. This mis-pricing component, however, is the second force causing changes in observed prices. The underlying mechanism is that mis-pricing generates trading opportunities for informed traders in the spirit of Kyle (1985). Consequently, in the long run market prices revert back to the efficient level as informed traders reap trading benefits and impound their knowledge into observed prices. The speed by which observed prices react to inherent mis-pricing in the market is governed by a parameter which is associated with the rate of efficiency of the observed price process. See, e.g., Sheppard (2013) for a related concept to measure the speed of the market.

We study the properties of the model in a simple discrete-time framework of constant volatility. In such a setting, the model depends on three key parameters, i.e., the volatility of the efficient price process (so-called ”efficient volatility”), the ”noise level” corresponding to the volatility of the noise process and the market efficiency parameter governing the speed of reversals due to mis-pricing. We show that such a three-parameter setting is quite flexible and can capture a wide range of stochastic behavior typically observed in micro-level prices.

As an important special case, our framework contains the classical ”martingale plus i.i.d. noise” case as a benchmark, see Zhang et al. (2005). Due to the interplay between the underlying ”efficient” volatility, market microstructure noise and the rate of market inefficiency, the approach, however, allows for more general high-frequency asset price dynamics. A key parameter in the model is the speed of price reversion due to inherent mis-pricing in the market. Since prices are directly disturbed by this mis-pricing, the model features properties in the observed price process which in the classical setting can only be captured by assuming endogenous noise, see e.g., Kalnina and Linton (2008).

With this new framework, we aim to provide a more structural approach for linking observed prices to efficient prices and therefore to explicitly link the literature on high-frequency based volatility estimation to market microstructure analysis. We show that our framework allows for natural interpretations of the underying model components. In particular, we propose the variance of the mis-pricing process as a natural measure for the market inefficiency and show that it crucially interferes with the speed of price reversion and the signal-to-noise ratio, corresponding to the proportion of the efficient variance to the total noise variance. Depending on the relative magnitude of these components we can distinguish between two fundamental states of the market: In one state, ”mis-pricing” is removed by ”contrarian behavior causing negative autocorrelations in observed returns. In the other regime, ”mis-pricing” is enforced by ”momentum” behavior causing positive autocorrelations in observed price changes.

We provide empirical evidence for the existence of both of these regimes in typical continous trading. Utilizing high-frequency data from NASDAQ trading, processed via the data platform LOBSTER[1] and estimating the model over short sub-periods provides local estimates of the underlying model parameters and thus the local market ”inefficiency”. It turns out that these parameters change over time and thus reflect different states in which the stochastic behavior of observed prices significantly change. In line with our model, we show that the identified regimes are in accordance with the underlying (local) autocorrelation structure in observed returns.

Moreover, we document that the existence of the two underlying regimes has profound consequences for realized volatilities sampled on high frequencies, e.g., Andersen et al. (2003). In the regime of contrarian behavior, realized volatility sampled on high frequencies is shown to

[1] See https://lobsterdata.com/.

significantly over-estimate the underlying quadratic variation of the efficient price process causing 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 easily explained in traditional ”martingale-plus-noise” settings but are naturally motivated in our framework. These results are empirically supported by providing novel evidence on time-varying volatility-signature plots. It turns out that the volatility signature locally changes and is in line with the identified underlying market regimes. We show that regimes of ”reverted” signature plots indeed occur more often than commonly assumed.

In its basic form, the model can be conveniently represented in state-space form and thus allows for maximum likelihood estimation using the Kalman filter. We therefore straightforwardly produce an efficient volatility estimator in a generalized martingale-plus-noise setting which can be benchmarked to existing estimators. Our results therefore provide new insights for high-frequency based volatility estimation, provide new channels for the construction of improved estimators and establish a direct link to underlying market microstructure literature.

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. Section 5 gives the empirical results and Section 6 concludes.

2 A High-Frequency Asset Price Model with Information Feedback and Market Microstructure Noise

In this section, we propose a model for micro-level asset price formation that generalizes the classical ”martingale-plus-noise” framework in the literature of high-frequency based volatility estimation. In our model, feedback between the efficient price and the observed market price is central.

Feedback between the observed price and the efficient price arises as informed traders continuously trade upon trade opportunities. Hence, observed prices are subject to *information feedback* correcting inefficiency and guaranteeing that observed prices lie within certain bounds around the efficient price. We consider a model in discrete time, i.e., $t \in \{0, 1, 2, \ldots\}$. Observed prices p_t are assumed to be driven by the following model:

$$p_{t+1} = p_t - \alpha \underbrace{(p_t - p_t^*)}_{:=\mu_t} + \epsilon_{t+1}, \quad \alpha \geq 0, \tag{2.1}$$

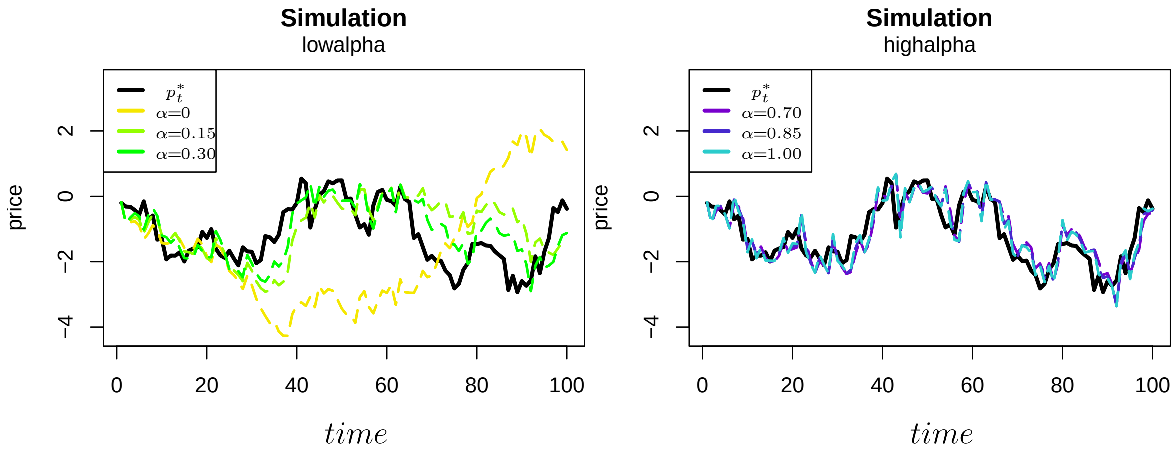
$$p_{t+1}^* = p_t^* + \epsilon_{t+1}^*, \tag{2.2}$$

where p_t^* denotes the efficient price at time t followinga standard martingale process with i.i.d. white noise innovations ϵ^*_{t+1}. Accordingly, the observed price p_t is subject to two different sources of randomness. First, market prices are disturbed by i.i.d. noise ϵ_t arising from non-informational sources. It may arise due to exogenous trading motifs or market imperfections associated with market microstructure frictions. For instance, liquidity or noise traders submit orders affecting observed prices without being connected to information. Consequently, observed prices face non-informational shocks which lead to mis-pricing (or price inefficiency), defined by $\mu_t := p_t - p^*_t \neq 0$.

Such mis-pricing generates trading opportunities for informed traders. Thus, in the long run, market prices p_t revert back to the efficient level p^*_t as informed traders reap trading benefits and impound their knowledge into observed prices. The larger the magnitude of the price inefficiency $|\mu_t|$, the stronger is the feedback that reverts market prices back to their efficient level.

Correspondingly, α is the speed of price reversion, and thus is a natural measure for the (in-)efficiency of the market. For instance, a large value of α suggests that market prices closely follow the efficient level, while a low value suggests a *sluggish* relationship. Figure 1 illustrates the relationship between p_t and p^*_t for different values of α. Low values of α represent a high market in-efficiency with observed prices and efficient prices deviating substantially. For high values of α, the observed price follows the efficient price more tightly.

Figure 1: Simulation of observed prices p_t (colored solid lines) and the efficient price p^*_t (black solid line) for different values of α. The left (right) figure shows the corresponding plots for low (high) values of α.



For the case $\alpha = 1$, we obtain the benchmark case where prices are driven by a martingale processes and are polluted by i.i.d. noise", see Zhang et al. (2005). Using (2.7), we straightforwardly obtain

$$\begin{aligned} p_{t+1} &= p_t - \alpha(p_t - p^*_t) + \epsilon_{t+1} \\ &= p^*_t + \epsilon_{t+1}. \end{aligned} \tag{2.3}$$

Under this specification, feedback perfectly mitigates prevailing mispricing. The ”classical” (i.i.d.) market microstructure noise framework therefore assumes a strong feedback. Accordingly, observed prices are only confronted with contemporaneous noise innovations ϵ_{t+1}.

The general case ($0 \leq \alpha \leq 1$), however, is more flexible, as it captures scenarios where observed prices follow the fundamental value sluggishly ($\alpha < 1$). This even includes the case $\alpha = 0$, where efficient and observed prices are completely independent (see the yellow graph in Figure 1).

The model can be re-formulated in terms of a state-space representation that de-couples latent and observed variables. This is straightforwardly done by treating the mis-pring component $\mu_t = p_t - p_t^*$ as an additional latent state variable. Therefore, we denote X_t as the state vector at time t with

$$X_t = \begin{pmatrix} \mu_{t+1} \\ p_{t+1}^* \end{pmatrix}. \tag{2.4}$$

To specify the dynamics of X_t, we need to recover the dynamics of μ_t and p_t^*. With $\epsilon_{t+1}^* := p_{t+1}^* - p_t^*$, μ_t can be represented by

$$\begin{aligned} \mu_{t+1} &= p_{t+1} - p_{t+1}^* \\ &= p_t - \alpha(p_t - p_t^*) + \epsilon_{t+1} - p_{t+1}^* \\ &= (1-\alpha)\underbrace{(p_t - p_t^*)}_{=\mu_t} + \epsilon_{t+1} - \underbrace{(p_{t+1}^* - p_t^*)}_{=\epsilon_{t+1}^*} \\ &= (1-\alpha)\mu_t + \epsilon_{t+1}^\mu, \end{aligned} \tag{2.5}$$

with $\epsilon_{t+1}^\mu := \epsilon_{t+1} - \epsilon_{t+1}^*$ being i.i.d. white noise. The mis-pricing component μ_t thus follows an AR(1) process with autoregressive parameter $1 - \alpha$. Accordingly, the state space dynamics are given by

$$X_{t+1} = GX_t + w_t,$$

with

$$X_t = \begin{pmatrix} \mu_t \\ p_t^* \end{pmatrix}, \quad G = \begin{pmatrix} 1-\alpha & 0 \\ 0 & 1 \end{pmatrix}, \quad w_t = \begin{pmatrix} \epsilon_{t+1}^\mu \\ \epsilon_{t+1}^* \end{pmatrix}. \tag{2.6}$$

The observed price process p_t is then given by

$$p_t = FX_t \qquad \text{with} \qquad F = \begin{pmatrix} 1 & 1 \end{pmatrix}. \tag{2.7}$$

The error covariance matrix Σ_w is obtained by

$$\Sigma_w = \begin{pmatrix} \mathbb{E}[\epsilon_t^\mu \epsilon_t^\mu] & \mathbb{E}[\epsilon_t^\mu \epsilon_t^*] \\ \mathbb{E}[\epsilon_t^\mu \epsilon_t^*] & \mathbb{E}[\epsilon_t^* \epsilon_t^*] \end{pmatrix} = \begin{pmatrix} \sigma_\epsilon^2 + \sigma_{\epsilon^*}^2 & -\sigma_{\epsilon^*}^2 \\ -\sigma_{\epsilon^*}^2 & \sigma_{\epsilon^*}^2 \end{pmatrix}. \tag{2.8}$$

Define the signal-to-noise ratio as the proportion of efficient (informational) volatility,

$$\lambda := \frac{\sigma^2_{\epsilon^*}}{\sigma^2_{\epsilon^*} + \sigma^2_{\epsilon}} = \frac{\sigma^2_{\epsilon^*}}{\sigma^2_{\mu}} \tag{2.9}$$

with $\sigma^2_\mu = \sigma^2_{\epsilon^*} + \sigma^2_\epsilon$ denoting the total noise variance in the system. Below we will show that λ has a profound impact on the dynamics implied by the model. Using λ, Σ_w can be written as

$$\Sigma_w = \sigma^2_\mu \begin{pmatrix} 1 & -\lambda \\ -\lambda & \lambda \end{pmatrix}. \tag{2.10}$$

Observe that Σ_w is not a diagonal matrix implying endogeneity between efficient prices p^*_{t+1} and the mis-pricing component μ_{t+1}. Note that this endogeneity is directly built in the model without explicitly assuming dependence between the errors ϵ^*_t and ϵ_t. See, e.g., Kalnina and Linton (2008).

From (2.5), it follows that for $0 < \alpha < 2$, μ_t is covariance-stationary with first and second moments given by

$$\mathbb{E}[\mu_t] = 0, \qquad \mathbb{E}[\mu_{t+\Delta}\mu_t] = (1-\alpha)^\Delta \frac{\sigma^2_\mu}{\alpha}\frac{1}{2-\alpha}, \tag{2.11}$$

for $\Delta \geq 0$ and with $\sigma^2_\mu = \mathbb{E}[(\epsilon^\mu_t)^2] = \sigma^2_{\epsilon^*} + \sigma^2_\epsilon$ as of (2.8). Hence, $\mathbb{E}[\mu_t] = 0$ implies that $\mathbb{E}[p_t] = \mathbb{E}[p^*_t]$. This arises from the information feedback $\alpha > 0$, which constantly pulls observed prices to back to their efficient level. Note that in the extreme case of $\alpha = 0$, $\mathbb{E}[\mu_t] = 0$ does not hold and observed prices can systematically deviate from efficient prices.

While observed and efficient prices equal in expectation, they nevertheless deviate over time. Evidently, the extent of this deviation is captured by $Var(\mu_t)$. Therefore, in our framework, it is reasonable to use $Var(\mu_t)$ as a proxy for the extent of inefficiency in p_t.

Definition 1 (Market Inefficiency). *Given the model* (2.1) *and* (2.2), *the total market inefficiency κ and relative market inefficiency $\bar{\kappa}$ are defined as*

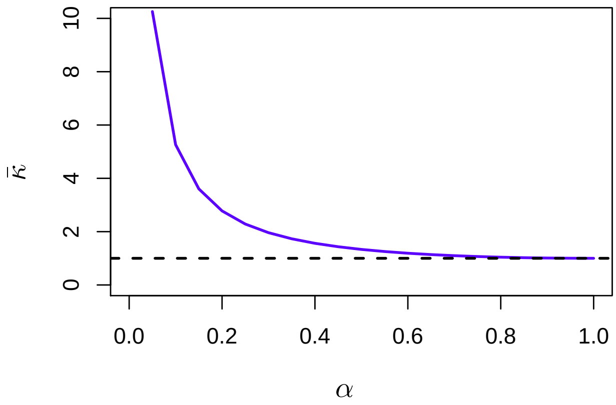
$$\kappa := Var(\mu_t), \qquad \bar{\kappa} := \frac{Var(\mu_t)}{\sigma^2_\mu}. \tag{2.12}$$

The total market inefficiency κ is a proxy for the total magnitude of deviations between efficient and observed price when the system is confronted with total noise level σ^2_μ. Correspondingly, the relative inefficiency $\bar{\kappa}$ is measured in terms a unit standard deviation of noise, i.e., $\sigma^2_\mu = 1$. Consequently, it only depends on the speed of price reversion α. Utilizing (2.11), κ and $\bar{\kappa}$ can thus be expressed as

$$\kappa := \frac{\sigma^2_\mu}{\alpha}\frac{1}{2-\alpha}, \qquad \bar{\kappa} := \frac{1}{\alpha}\frac{1}{2-\alpha}. \tag{2.13}$$

Figure 2 shows the relationship between $\bar{\kappa}$ and α. It is shown that $\bar{\kappa}$ is monotonously decreasing in α which is consistent with the intuition that a stronger (information) feedback α should make observed prices more efficient. Note, moreover, that $\bar{\kappa} > 1$. Thus, the total inefficiency κ cannot get arbitrarily small and is bounded from below by σ_μ^2.

Figure 2: Illustration of $\bar{\kappa}$ depending on α. The solid blue line corresponds to $\bar{\kappa}$, the dashed black line corresponds to the level 1.



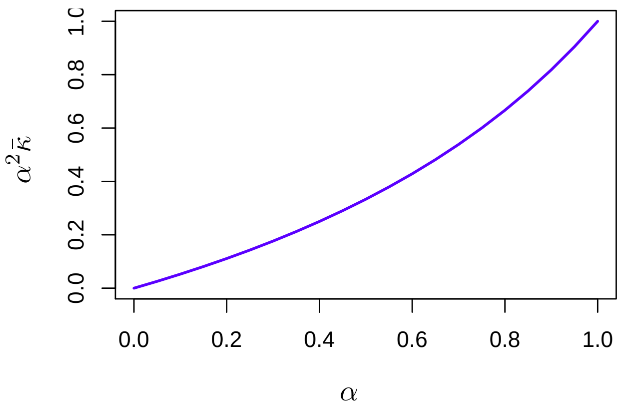
3 Informational Noise and the Return Autocovariance: Two Regimes

Define $r_{t+1} := p_{t+1} - p_t$ and $r^*_{t+1} := p^*_{t+1} - p^*_t$ as the observed and efficient return, respectively. Note that the observed return obeys $r_{t+1} = -\alpha\mu_t + \epsilon_{t+1}$. Taking into account that ϵ_{t+1} and μ_t are independent, the return variance is thus given by

$$\begin{aligned} Var(r_t) &= \sigma_\epsilon^2 + \alpha^2 Var(\mu_t) \\ &= \sigma_\epsilon^2 + \alpha^2 \kappa. \end{aligned} \tag{3.1}$$

Thus, information feedback triggered by α induce excess variation in observed returns on top of variations due to the non-informational noise component ϵ_t. This contribution is amplified by the variance of the mis-pricing process, $Var(\mu_t)$. Hence, innovations in the underlying efficient price are channeled forward to the market price by the speed α. Thus, the higher α, the more informational variations arrive at the observed return process, amplifying their variations. It is easily checked that, despite κ being monotonously decreasing in α, $Var(r_t)$ is monotonously increasing in α, see Figure 3. Hence, when feedback is strong, market inefficiency κ decreases, but variations in the price return increase. This is natural as mis-pricing μ_t triggers stronger reactions when the feedback magnitude α is high.

Figure 3: Relationship between $\alpha^2\bar{\kappa}$ and α.



We moreover show that feedback effects do not necessarily imply that excess variations in observed return variations exceed variations of the efficient price. Observe that (3.1) implies

$$\sigma_\epsilon^2 \leq Var(r_t) \leq 2\sigma_\epsilon^2 + \sigma_{\epsilon^*}^2, \tag{3.2}$$

where the lower and upper bound are sharp bounds, i.e., they are attained for $\alpha = 0$ and $\alpha = 1$, where in the latter case, we obtain $Var(r_t) = 2\sigma_\epsilon^2 + \sigma_{\epsilon^*}^2$. Thus, in fact, depending on α, σ_ϵ^2 and $\sigma_{\epsilon^*}^2$, both scenarios may occur, i.e., either observed or efficient variance can be higher. To see this, observe that by construction $\sigma_\epsilon^2 = (1-\lambda)\sigma_\mu^2$ and $\sigma_{\epsilon^*}^2 = \lambda\sigma_\mu^2$ holds. Using (3.1) the variances of the efficient return and the observed price return are given by

$$Var(r_t) = \sigma_\mu^2(2\alpha\bar{\kappa} - \lambda), \qquad Var(r_t^*) = \lambda\sigma_\mu^2. \tag{3.3}$$

This allows to identify regimes where the observed return variance exceeds the efficient return variance and vice versa. The result is stated in the following lemma.

Lemma 1. *The return variance of the observed price process, $Var(r_t)$, and the efficient price process, $Var(r_t^*)$, obey*

$$Var(r_t) < Var(r_t^*) \qquad \textit{if} \qquad \alpha\bar{\kappa} < \lambda, \tag{3.4}$$

$$Var(r_t) \geq Var(r_t^*) \qquad \textit{otherwise}, \tag{3.5}$$

with $\frac{1}{2} \leq \alpha\bar{\kappa} = \frac{1}{2-\alpha} \leq 1$.

If $\lambda \leq 1/2$, *then*

$$Var(r_t) \geq Var(r_t^*). \tag{3.6}$$

Thus, when informational noise dominates ($\lambda \geq \bar{\kappa}$), the observed return variation is lower than the return variation of the efficient price. The result is intuitive. If λ is sufficiently high, the proportion of the ”informational variance” is high, i.e., price innovations are dominantly driven by information. Then, p_t sluggishly follows p_t^* because $\alpha \leq 1$, implying that changes in the underlying efficient price are not passed over to the observed price one-to-one but in a mitigated way.

In the second regime, however, λ is low and price innovations are mostly driven by non-informational noise. In this case, p_t tend to continuously correct non-informational disturbances ϵ_t. Then, information feedback tends to act in a contrarian way (i.e., implying contrarian trading) and the observed price constantly mean reverts in order to mitigate the impact of non-informational shifts. This mechanism induces excess return variance. It is interesting to note that the regime $\alpha\bar{\kappa} < \lambda$ only exists for $\lambda \geq 1/2$. In other words, if the signal to noise ratio is not at least 50%, the observed return variance always exceeds the return variance of the efficient price.

An important implication of these mechanisms are that the amount of informational noise affects the nature of autocovariances in observed returns. More precisely, when informational noise is dominant, observed returns are positively autocorrelated and prices tend to follow local trends. In the opposite case, observed returns become negatively autocorrelated and prices reveal mean-reverting behavior. These effects are stated in the following lemma.

Lemma 2 (Return Autocovariance)**.** *Assume* $0 < \alpha < 2$ *and* $\Delta \geq 1$*. Then,*

$$Cov(r_{t+\Delta}, r_t) = -\psi(\Delta - 1)\Big(\alpha\bar{\kappa} - \lambda\Big), \tag{3.7}$$

with $\psi(\Delta - 1) = \alpha(1-\alpha)^{\Delta-1}\Big(\sigma^2_{\epsilon^*} + \sigma^2_{\epsilon}\Big) \geq 0.$

Proof. See Appendix. □

Hence, high informational noise ($\lambda > \alpha\bar{\kappa}$) induces positive serial correlation, while low informational noise ($\lambda < \alpha\bar{\kappa}$) causes negative serial correlation. The corollary characterizes the different regimes in terms of the parameters λ, α and $\bar{\kappa}$.

Corollary 1. *Assume* $0 < \alpha < 2$*. Then, the following holds.*

(i) *If* $\alpha = 0$*, then* $Cov(r_{t+\Delta}, r_t) = 0$*.*

(ii) *If* $0 < \alpha \leq 1$*, and*

   (a) *if* $\lambda < \alpha\bar{\kappa}$*, then* $Cov(r_{t+\Delta}, r_t) < 0$*,*

   (b) *if* $\lambda \geq \alpha\bar{\kappa}$*, then* $Cov(r_{t+\Delta}, r_t) \geq 0$*.*

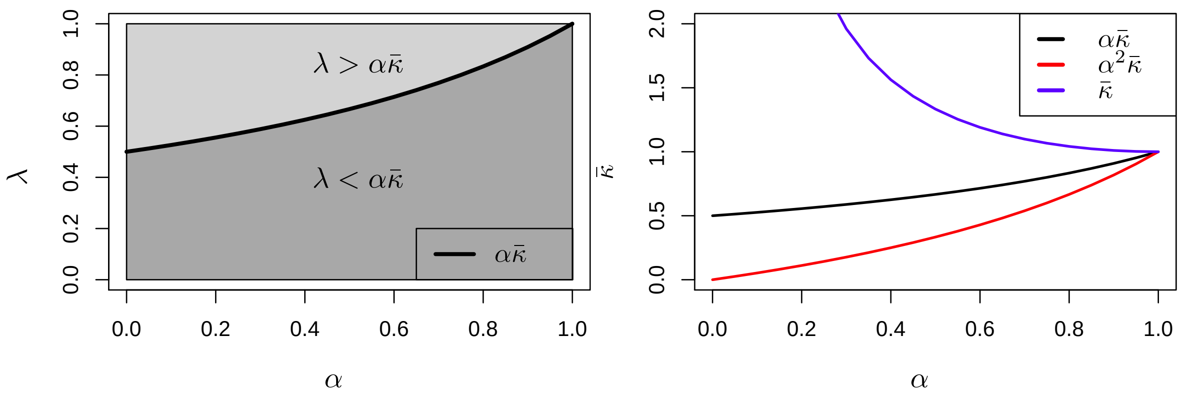
(iii) *If* $1 < \alpha < 2$*, then* $Cov(r_{t+\Delta}, r_t) < 0$*.*

Proof. Follows directly from Lemma 2. □

In the case $1 < \alpha < 2$, observed price changes over-shoot as a response to mis-pricing. Hence, the forces that make prices efficient in the long run induce inefficiency themselves. Therefore, this case is a rather pathological (but not necessarily unrealistic) setting.

While $\alpha = 0$ is the trivial case, our major focus therefore is on the generic case (ii). The two regimes, characterized by λ and the threshold $\alpha\bar{\kappa}$, are illustrated in Figure 4. When observed prices p_t are disturbed by non-informational noise ϵ_t, then market makers or informed traders act as contrarian traders and revert prices back to their efficient level inducing negative serial correlation. Conversely, when the informational variance innovation $\sigma^2_{\epsilon^*}$ is high, informed traders act as momentum traders and cause positive return correlation. Consequently, prices reveal local trends. Overall, we can expect that the correlation structure of observed returns is driven by the combination of both regimes.

Figure 4: Illustration of the two regimes $\lambda \geq \alpha\bar{\kappa}$ and $\lambda < \alpha\bar{\kappa}$ (left figure) and a comparsion of the components $\bar{\kappa}$, $\alpha\bar{\kappa}$ and $\alpha^2\bar{\kappa}$.



To illustrate the effects, we present the two regimes in terms of two numerical examples.

Example 1 (Positive serial correlation)**.** *Assume* $\alpha = 0.5$ *and* $\lambda = 1$. *Then* $\xi = -1/3$. *Assume also that the efficient and the market price initially co-incide, i.e.,* $p_0 = p^* = 0$ *and that* $\alpha = 0.5$. *Now assume hat at time* $t = 1$, *information arrival (*$\epsilon_1 = 2$*) changes the efficient level to* $p^*_1 = 2$ *and that it remains constant thereafter, i.e.,* $\epsilon^*_t = 0$ *for* $t \geq 2$. *Then, because of* (2.7), *market returns take the form* $r_{t+1} = -\alpha\mu_{t+1} = -\alpha(p_t - p^*_t)$. *Therefore, the returns at times* $t > 0$ *obey*

$$r_1 = -\frac{1}{2}(0-0) = 0, \qquad p_1 = p_0 + r_1 = 0 \tag{3.8}$$

$$r_2 = -\frac{1}{2}(0-2) = 1 > 0, \qquad p_2 = p_1 + r_2 = 1 \tag{3.9}$$

$$r_3 = -\frac{1}{2}(1-2) = 1/2 > 0, \qquad p_3 = p_2 + r_3 = 3/2 \tag{3.10}$$

Hence, innovations in the efficient market price at $t = 1$, *generate a sequence of positively correlated returns at* $t > 1$.

Example 2 (Negative serial correlation)**.** *Consider the case* $\alpha = 1/2$ *and* $\lambda = 0$*. Then,* $\xi = 2/3 > 0$*. Assume also* $p_0 = p^* = 0$ *and that the system is shocked by a single non-informational innovation at* $t = 1$*, i.e.* $\epsilon_1 = 2$ *and* $\epsilon_t = 0$ *for* $t > 1$ *and that there is no information arrival, i.e.* $\epsilon_t^* = 0$ *for* $t \geq 0$*. Market returns obey* $r_{t+1} = -\alpha(p_t - p_t^*) + \epsilon_{t+1}$*. Therefore, the returns at times* $t > 0$ *obey*

$$r_1 = -\frac{1}{2}(0-0) + 0 = 0, \qquad p_1 = p_0 + r_1 = 0 \tag{3.11}$$

$$r_2 = -\frac{1}{2}(0-0) + 2 = 2 > 0, \qquad p_2 = p_0 + r_2 = 2 \tag{3.12}$$

$$r_3 = -\frac{1}{2}(2-0) + 0 = -1 < 0, \qquad p_3 = p_1 + r_2 = 1 \tag{3.13}$$

Hence, non-informational innovations at $t = 1$ *induce negative dependence in* $t = 2$*.*

4 Implications for the Realized Variance

In this section, we analyze the quadratic variation of the observed price process implied by (2.6) and (2.7). We will show that the identified regimes translate in an intriguing way in distinct properties of the quadratic variation process. This has clear implications for realized variances sampled on high frequencies.

Let $r_t^\Delta := p_{t+\Delta} - p_t$, denote the $\Delta-$ return sampled at time t. Moreover, define the time-T *quadratic variation* of the price process p_t, sampled over subsequent sub-intervals of size Δ, as

$$\langle p \rangle_T^\Delta := \mathbb{E}\left[\sum_{i=0}^{\frac{T}{\Delta}} (r_{i\Delta}^\Delta)^2\right]. \tag{4.1}$$

Now, our main result follows.

Theorem 1. *Assume* $0 < \alpha < 1$*. Then, the time-*T *quadratic variation* $\langle p \rangle_T^\Delta$ *sampled over the interval size* Δ *obeys*

$$\langle p \rangle_T^\Delta = \sigma_{\epsilon^*}^2 T + \sigma_\mu^2 T \phi(\Delta)(\alpha\bar{\kappa} - \lambda), \tag{4.2}$$

$$\textit{with} \qquad \phi(\Delta) = \frac{2}{\alpha\Delta}\left(1 - (1-\alpha)^\Delta\right). \tag{4.3}$$

The mapping $\Delta \mapsto \phi(\Delta)$ *is strictly monotonously decreasing and obeys*

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

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

(iii) with $0 \leq \phi(\Delta) \leq -\frac{2}{\alpha}\log(1-\alpha)$.

Hence, according to Theorem 1, the quadratic variation consists of two terms. The first term is the quadratic variation of the efficient price which in our simplified setting just equals $\sigma^2_{\epsilon^*}T$ and is associated with the "fundamental" variance measured over the interval from 0 to T. The second term originates from the information transmission feedback between efficient and observed prices and is a function of the sampling interval Δ (or of the sampling frequency Δ^{-1}, respectively, for $T = 1$). Excluding the trivial case $\alpha = 0$, we can again identify two regimes triggered by the signal-to-noise ratio λ and the relative market inefficiency $\bar{\kappa}$. In the regime of high informational noise ($\lambda > \alpha\bar{\kappa}$), the quadratic variation of the observed price process is systematically below the quadratic variation of the underlying efficient price process. In the opposite case, however, the quadratic variation of the observed price process systematically exceeds the fundamental variance.

The following corollary show that resulting volatility-signature plots (depicting the relationship between $\langle p\rangle^{\Delta}_{T}$ and Δ) are monotonically increasing or decreasing, respectively, in the two regimes.

Corollary 2 (Monotonicity of quadratic variation and noise regimes)**.** *Assume* $0 \leq \alpha \leq 1$*, a fixed time period* $T > 0$*, and the sub-sampling interval* $\Delta \leq T$*. Then, the following holds.*

(i) If $\alpha = 0$*, then the mapping* $\Delta \mapsto \langle p\rangle^{\Delta}_{T}$ *is constant.*

(ii) If $< \alpha \leq 1$*, then*

(a) if $\lambda < \alpha\bar{\kappa}$*, the mapping* $\Delta \mapsto \langle p\rangle^{\Delta}_{T}$ *is strictly monotonically decreasing,*

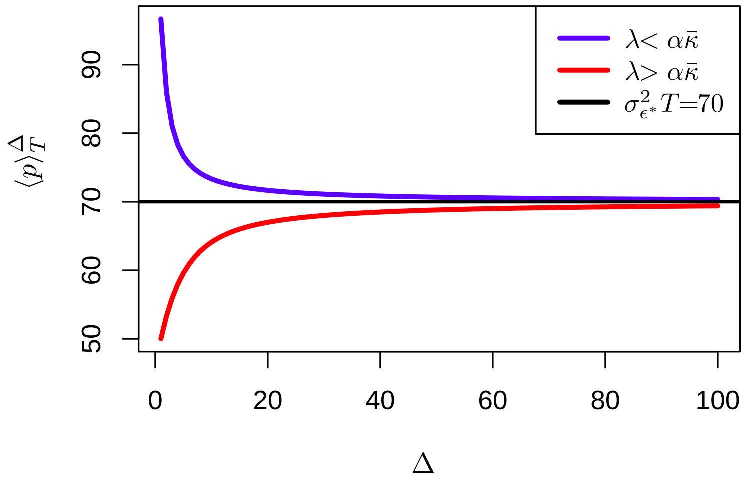
(b) if $\lambda > \alpha\bar{\kappa}$*, the mapping* $\Delta \mapsto \langle p\rangle^{\Delta}_{T}$ *is strictly monotonically increasing.*

Proof. Direct consequence of Theorem 1. □

Hence, when the signal-to-noise ratio is low ($\lambda < \alpha\bar{\kappa}$), the quadratic variation of the observed price process decreases with the sample size and systematically overestimates the quadratic variation of the efficient price. Consequently, realized variances sampled on high frequencies reveal the well-know positive bias documented in the literature, see, e.g., Hansen and Lunde (2006). At extremely high sampling frequencies ($\Delta \to 0$), the divergence between efficient and observed quadratic variation is maximal and equals $|\phi(0)T\sigma^2_{\mu}(\alpha\bar{\kappa} - \lambda)|$.

In the opposite regime, when informational noise is high ($\lambda < \alpha\bar{\kappa}$), the quadratic variation of the observed price under-estimates the efficient quadratic variation for $\Delta \to 0$. An illustration of both regimes are shown in Figure 5.

Figure 5: Simulated quadratic variation $\langle p \rangle_T^\Delta$ as a function of the sampling interval Δ for the two different regimes. Exemplary parameters are $T = 100$, $\sigma_\mu^2 = 1$, $\lambda = 7/10$. The blue line corresponds to $\alpha = 4/5$ implying that informational noise is low, i.e., $\alpha\bar{\kappa} > \lambda$. The red line corresponds to $\alpha = 1/3$ corresponding high informational noise, i.e., $\alpha\bar{\kappa} < \lambda$. The black line corresponds to the efficient level, i.e., $\sigma_{\epsilon^*}^2 T = \sigma_\mu^2 \lambda = 70$.



The underlying reasoning for the two regimes is in analogy to the discussion above and is essentially induced by the return autocovariance structure as of Lemma 2. When informational noise is sufficiently high, observed returns are positively autocorrelated returns and thus local drifts accumulate on longer time horizons and do not average out. Consequently, the quadratic variation increases monotonously with Δ. As $\alpha < 1$, observed prices however do not fluctuate as much as the efficient price process since only a fraction of these fluctuations are channeled back to the observed price.

On the other hand, when informational noise is low, observed returns are negatively serially correlated and long-run returns get effectively averaged out and decay over the sampling interval. Hence, increasing volatility signatures are mainly due to local drifts, while decreasing volatility signature arise from mean-reverting behavior.

Thus, our theory provides a unified framework to explain both regimes and tie them to fundamental informational properties of the underlying market, such as the signal-to-noise ratio λ and the relative market inefficiency $\bar{\kappa}$. Particularly the regime of downward sloped signature plots (for $\Delta \to 0$) is interesting and are implied by effects which are not discussed in the literature yet. In the empirical section, we will provide evidence for the presence of both regimes in high-frequency trading processes.

The next corollary provides some results on interesting limit cases.

Corollary 3. *The following limits hold*

$$(i) \qquad \lim_{\alpha\to 0} \langle p \rangle_T^\Delta = \sigma_\epsilon^2 T, \qquad \lim_{\alpha\to 1} \langle p \rangle_T^\Delta = \sigma_{\epsilon^*}^2 T + \sigma_\mu^2 T \frac{2-\lambda}{2\Delta}, \tag{4.4}$$

$$(ii) \qquad \lim_{\lambda\to 0} \langle p \rangle_T^\Delta = \sigma_\epsilon^2 T \phi(\Delta) \frac{1}{2-\alpha}, \quad \lim_{\lambda\to 1} \langle p \rangle_T^\Delta = \sigma_{\epsilon^*}^2 T \phi(\Delta) \left(\frac{1}{2-\alpha} - 1 \right). \tag{4.5}$$

Proof. Direct consequence of Theorem 1. □

Finally, we consider the diffusion limit of the discrete-time model (2.6) and (2.7). In the continuous time limit, the inefficiency process μ_t becomes a continuous time Ornstein-Uhlenbeck process and the efficient price process P_t^* becomes a standard Brownian motion, i.e.,

$$d\mu_t = -\alpha\mu_t + \sigma dW_t - \sigma^* dW_t^*, \qquad dP_t^* = \sigma^* dW_t^*, \tag{4.6}$$

with the first two moments given by

$$\mathbb{E}[\mu_t] = 0, \qquad \mathbb{E}[\mu_t^0] = \frac{(\sigma^*)^2 + \sigma^2}{2\alpha}. \tag{4.7}$$

Consequently, (4.6) is a system of linear stochastic differential equations with correlated random error terms with dW_t and dW_t^* being the differential increments of two uncorrelated standard Wiener process with $\mathbb{E}[dW_t, dW_t^*] = 0$ and $\mathbb{E}[dW_t] = 0$. The parameter σ^* denotes the (constant) volatility of the efficient price process P_t^* and σ denotes the (constant) noise volatility. Recall that the observed price process P_t satisfies

$$P_t = P_t^* + \mu_t. \tag{4.8}$$

By Ito's Lemma, the integrated Ornstein-Uhlenbeck process μ_t then satisfies

$$\mu_t = \mu_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW_s - \sigma^* \int_0^t e^{-\alpha(t-s)} dW_s^*. \tag{4.9}$$

For the observed market price process p_t, we consider the quadratic variation over the period T with sampling interval Δ, i.e.,

$$\langle p \rangle_T^\Delta = \sum_{i=0}^N \mathbb{E}[r_{i\Delta}^2], \tag{4.10}$$

where $r_{i\Delta} = p_{(i+1)\Delta} - p_{i\Delta}$. Using (4.6) and (4.9) , one can decompose the high-frequency returns $r_{i\Delta}$ into three mutually independent components

$$r_{i\Delta} = A_i + B_i + C_i, \tag{4.11}$$

with

$$A_i = \mu_0 e^{-\alpha i\Delta}\Big(e^{-\alpha\Delta} - 1\Big), \tag{4.12}$$

$$B_i = \sigma e^{-\alpha i\Delta}\left(\int_0^{(i+1)\Delta} e^{-\alpha(\Delta-s)}dW_s - \int_0^{i\Delta} e^{\alpha s}dW_s\right), \tag{4.13}$$

$$C_i = \sigma^*\left(\int_0^{(i+1)\Delta}(1-e^{-\alpha((i+1)\Delta-s)})dW_s^* - \int_0^{i\Delta}(1-e^{-\alpha(i\Delta-s)})dW_s^*\right). \tag{4.14}$$

Observe that A_i, B_i and C_i are zero-centered and Gaussian and so is the return $r_{i\Delta}$.The mutual independence of A_i, B_i and C_i derives from the fact that the Wiener increments dW_t and dW_t^* are independent, i.e., $\mathbb{E}[dW_t, dW_s^*] = 0$ for any $s, t \geq 0$. Therefore, the quadratic variation can be decomposed as follows

$$\langle p\rangle_T^\Delta = \sum_{i=0}^N \mathbb{E}[r_{i\Delta}^2] = \sum_{i=0}^N\left(\mathbb{E}[A_i^2] + \mathbb{E}[B_i^2] + \mathbb{E}[C_i^2]\right). \tag{4.15}$$

The following lemma provides the quadratic variation $\langle p\rangle_T^\Delta$ for the observed market price in the diffusion limit as of (4.6). We show that similar to the discrete case, the quadratic variation can be decomposed into an efficient and a noise component, where the latter depends on the sampling interval Δ.

Lemma 3. *Let σ^2 and $(\sigma^*)^2$ be the variance of the market microstructure noise and informational noise; Δ and be the sampling interval for the quadratic variation of the observed price process, $\langle p\rangle_T^\Delta$. Then, $\langle p\rangle_T^\Delta$, satisfies*

$$\langle p\rangle_T^\Delta = T\sigma_\epsilon^2 + T\widetilde{\Phi}(\Delta)(\sigma^2 - (\sigma^*)^2), \tag{4.16}$$

with $\Phi(\Delta) = \frac{1-e^{-\alpha\Delta}}{\alpha\Delta}$ such that the mapping $\Delta \mapsto \widetilde{\Phi}(\Delta)$ is strictly monotonically decreasing and strictly convex. Therefore, $\langle p\rangle_T^\Delta$ satisfies the following:

(i) If $\sigma > \sigma^*$, then $\langle p\rangle_T^\Delta > T(\sigma^*)^2$ and $\langle p\rangle_T^\Delta$ is strictly monotonically decreasing in Δ.

(ii) If $\sigma < \sigma^*$, then $\langle p\rangle_T^\Delta < T(\sigma^*)^2$ and $\langle p\rangle_T^\Delta$ is strictly monotonically increasing in Δ.

(iii) If $\sigma = \sigma^*$, then $\langle p\rangle_T^\Delta = T(\sigma^*)^2$ and $\langle p\rangle_T^\Delta$ is constant with respect to Δ.

Moreover, $T\min(\sigma^2, (\sigma^*)^2) \leq \langle p\rangle_T^\Delta \leq T\max(\sigma^2, (\sigma^*)^2)$ and the following limits hold.

$$\lim_{\Delta\to 0}\langle p\rangle_T^\Delta = T(\sigma^*)^2, \qquad \lim_{\Delta=T\to\infty}\langle p\rangle_T^\Delta = T\sigma^2 \tag{4.17}$$

5 Empirical Evidence

In this section, we provide empirical support for the proposed model and demonstrate that high-frequency trading processes reveal the two regimes suggested by our model. We first provide novel evidence on time variability of volatility-signature plots. For a given sampling interval Δ, we consider high-frequency midquote returns measured over the interval Δ,

$$r_{t_i}^{\Delta} = p_{t_i+\Delta} - p_{t_i} \qquad \text{with} \qquad t_i \in \mathbb{T} = \{0, \Delta, 2\Delta, \ldots m\Delta\}, \tag{5.1}$$

with $m := \sup_{l \leq T/\Delta} l$ and p_{t_i} denoting the midquote at time t_i. Then, the realized variance RV_T^{Δ} for a fixed sampling interval Δ and time period T is given as

$$RV_T^{\Delta} := \sum_{t_i \in \mathbb{T}} (r_{t_i}^{\Delta})^2. \tag{5.2}$$

Figures 14 to 18 show the signature plots of daily realized variances averaged over ten consecutive trading days for different stocks selected from the NASDAQ 100 index in January 2014. To illustrate that dynamic evolution of the realized variances, we consider 6 different 10-day periods between January 2nd, 2014 to March 28th 2014. Non-trading days are excluded.

The results are as follows. First, in line with our theoretical model, we observe that volatility signatures reveal both upward and downward sloping behavior. Particularly, the presence of upward sloped (for Δ increasing) signature plots is striking and clearly present in several cases, see e.g., for Apple (Figure 14), Amazon (Figure 15) and Google (see Figure 18). We checked that these effects are *not* driven by the presence by a local drift in price levels.[2] These results therefore indicate that these effects occur more frequently than currently believed.

Second, the slope of the signature plots are highly time dependent. For instance, in the case of Google (Figure 18), quadratic variation tends to increase in the second and third trading period. Then, in the forth trading period the volatility signature inverts and jumps back to monotonic decay in the fifth and sixth trading period. These effects suggest that market conditions and thus the stochastic properties of the underlying processes may change rapidly.

Third, overall, the generic patterns can be classified into three distinct regimes or superpositions thereof. 1) monotonic decay, 2) monotonic increase or 3) constancy during the sampling period. Given that the quadratic variation is regime-dependent and we average over an extended time period of ten trading days, we may naturally estimate the combined effect of two regimes.

[2]Re-doing the analysis based on de-meaned high-frequency returns yields virtually identical results, see Appendix. The reason is that the second moment dominates the first moment of the observed return by orders of magnitude. Therefore, the quadratic variation is not affected by drift terms, even if we average over comparably short periods.

Figure 6: Averaged daily realized variance for six different trading periods for the stock Apple (AAPL). Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for a given sampling interval Δ.

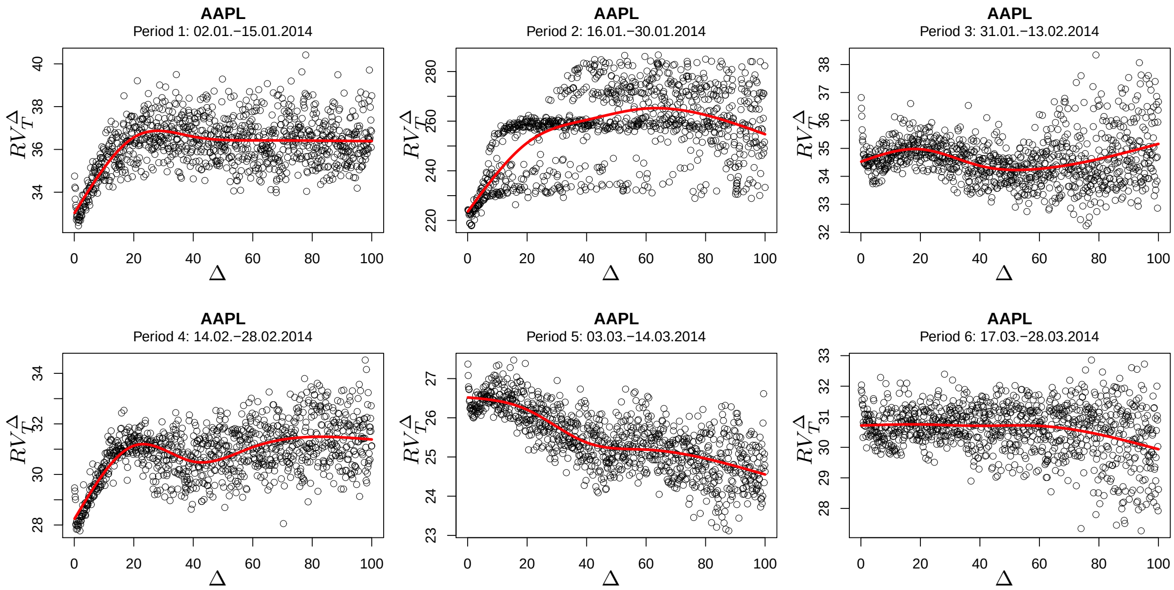


Figure 7: Averaged daily realized variance for six different trading periods for the stock Amazon (AMZN). Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for a given sampling interval Δ.

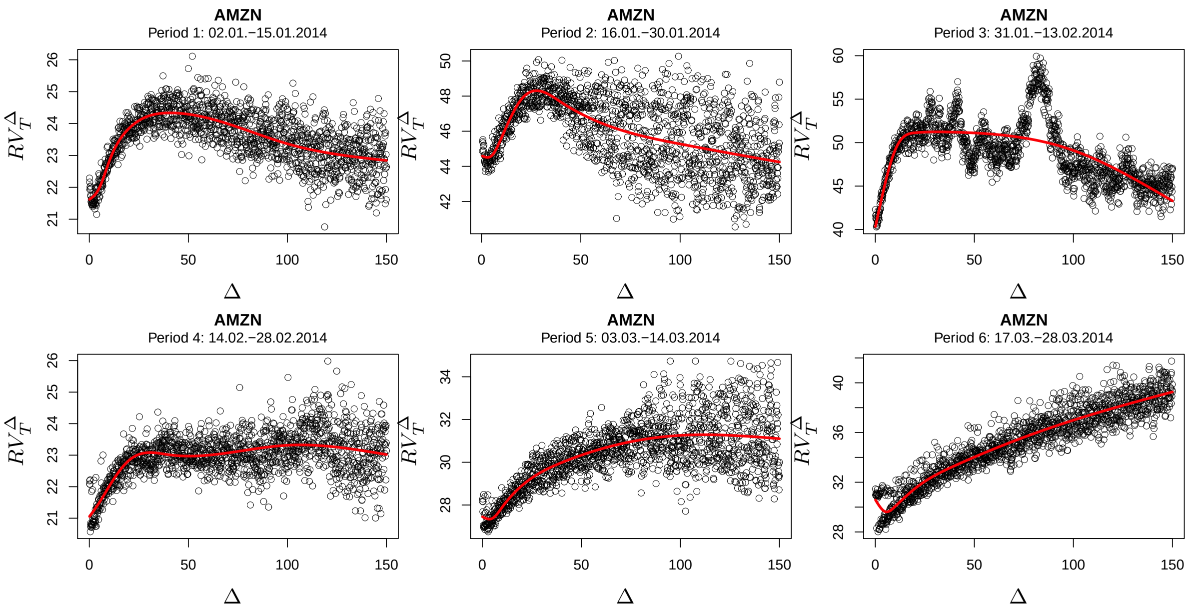


Figure 8: Averaged daily realized variance for six different trading periods for the stock Cisco (CSCO). Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for fixed sub-sampling interval Δ.

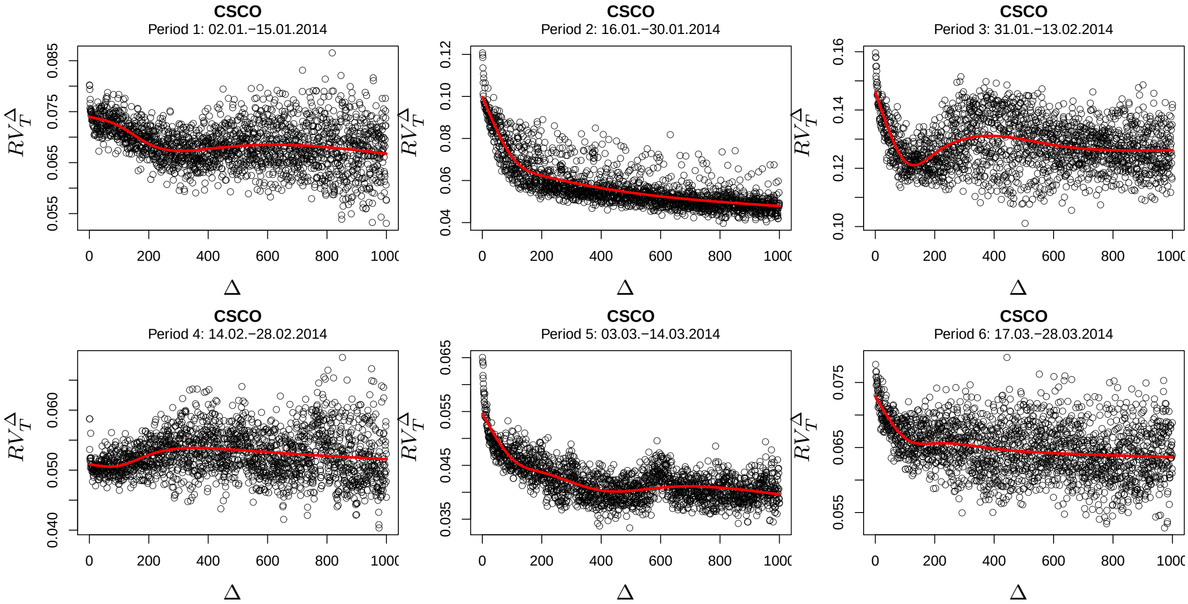


Figure 9: Averaged daily realized variance for six different trading periods for the stock Ebay (EBAY). Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for a given sampling interval Δ.

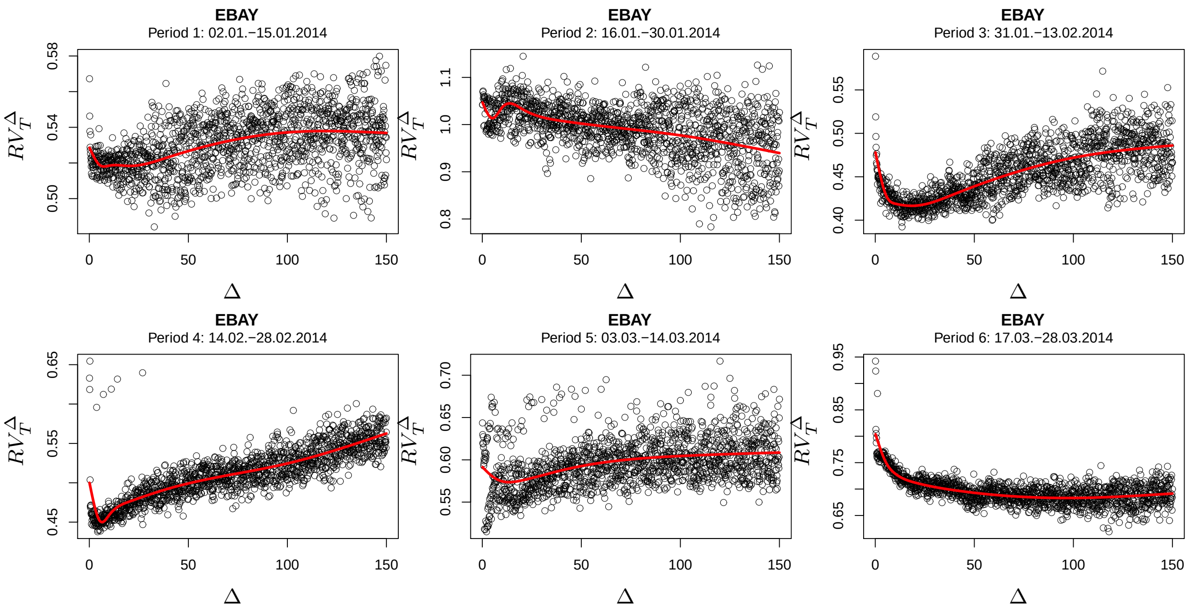


Figure 10: Averaged daily realized variance for six different trading periods for the stock Google (GOOG). Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for a given sampling interval Δ.

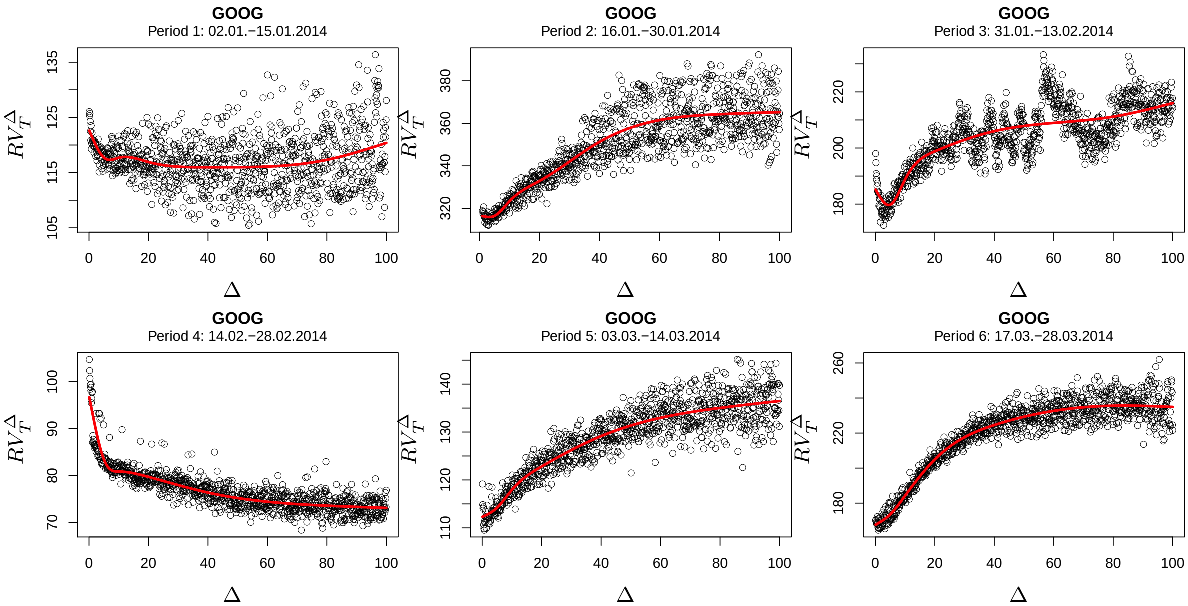


Figure 11: Averaged daily realized variance for six different trading periods for the stock Intel (INTC). Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for a given sampling interval Δ.

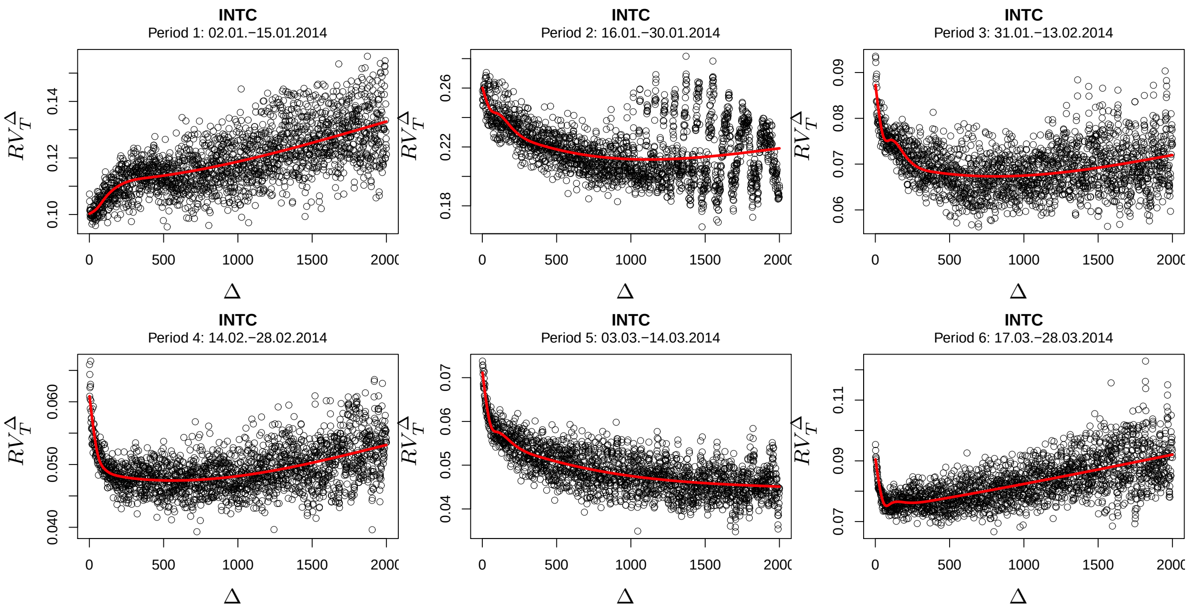
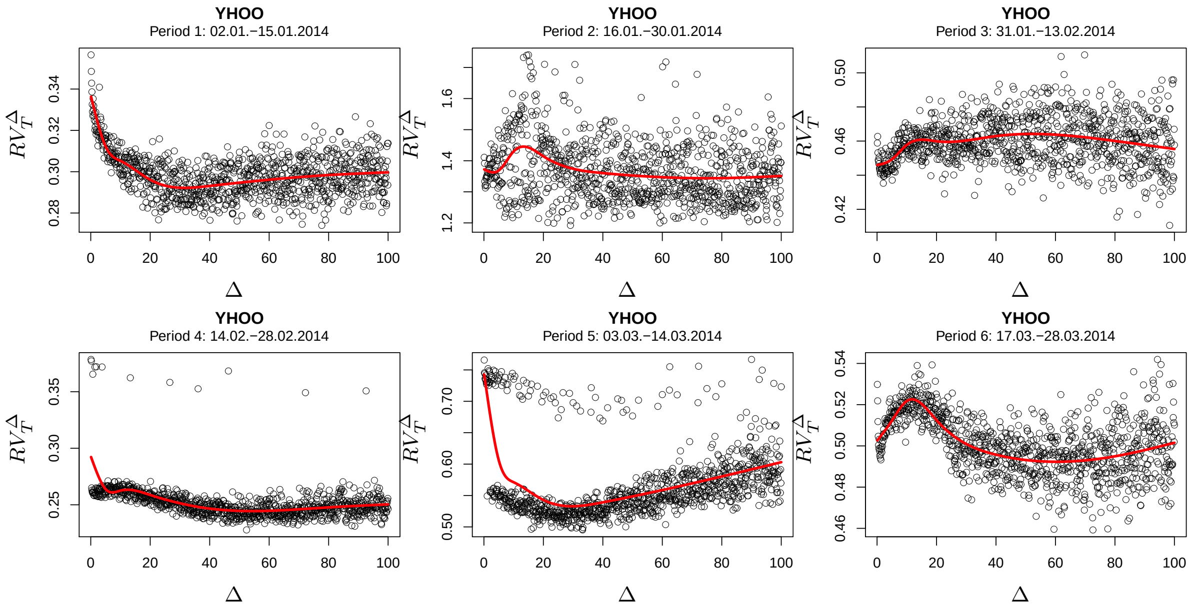


Figure 12: Averaged daily realized variance for six different trading periods for the stock Yahoo (YHOO). Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for a given sampling interval Δ.



According to our model, the different slopes of volatility-signature plots should go hand in hand with the presence of positive and negative autocorrelations in observed high-frequency returns during the respective intervals. Moreover, the corresponding regimes should be alternatively identifiable in terms of the relationship between λ, α and $\bar{\kappa}$.

The model is straightforwardly estimated using the state-space representation provided in Section2. Given normality of the underlying model errors, the parameters are then estimated by Gaussian maximum likelihood using the Kalman filter. Alternatively, the model might be estimated by the method of moments utilizing the autocovariance function derived in Section 3.

According to theory, there are two fundamental regimes: First, when the proportion of informational noise is high, i.e., $\alpha\kappa - \lambda < 0$, then markets exhibit positive serial return correlation and an upward sloping volatility signature. Second, when informational noise is low, $\alpha\kappa - \lambda > 0$, markets exhibit negative serial correlation and downward sloping volatility signatures.

To test our theory, we proceed as follows. From the volatility signature plots in Figure 14 to 20, we select those stocks and trading periods for which the volatility signature is either upward or downward sloping and estimate the return correlation of the observed price process and construct the ML estimators for α and λ, denoted by $\hat{\alpha}$ and $\hat{\lambda}$. Table 1 reports ML estimates

$\hat{\alpha}$ and $\hat{\lambda}$, the corresponding critical value

$$\frac{1}{2-\hat{\alpha}} - \hat{\lambda}, \tag{5.3}$$

separating between the two regimes and the first order autocorrelation of observed returns r_t, ρ.

Table 1: ML estimates for $\hat{\alpha}$ and $\hat{\lambda}$, $\frac{1}{2-\alpha} - \hat{\lambda}$ and the first order autocorrelation $\hat{\rho}$ for different stocks and trading periods. Estimation based on Kalman filter.

				ML Estimates			
Stock	Period	Trading Dates	V-Regime	$\hat{\alpha}$	$\hat{\lambda}$	$\frac{1}{2-\hat{\alpha}} - \hat{\lambda}$	$\hat{\rho}$
AMZN	5	03.03.-14.03.2014	up	0.769	1.000	-0.188	0.030
CSCO	2	16.01.-30.01.2014	down	0.445	0.523	0.120	-0.022
EBAY	6	17.03.-28.03.2014	down	0.440	0.612	0.029	-0.005
GOOG	4	14.02.-28.02.2014	down	0.489	0.536	0.126	-0.026
GOOG	5	03.03.-14.03.2014	up	0.728	0.999	-0.213	0.006
GOOG	6	17.03.-28.03.2014	up	0.066	0.701	-0.184	0.006
INTC	1	02.01.-15.01.2014	up	0.249	0.597	-0.026	-0.032
INTC	5	03.03.-14.03.2014	down	0.352	0.492	0.115	-0.017

The results are as follows. First, in line with our theory, all regimes with downward sloped signature plots coincide with a negative value of $\frac{1}{2-\hat{\alpha}} - \hat{\lambda}$ and thus a ”momentum regime”. This is moreover confirmed by positive autocorrelations $\hat{\rho}$. This accordance holds in all except one case. The only exception occurs for the Intel stock during the trading period $02.01. - 15.01.2014$, where the sign of $\frac{1}{2-\hat{\alpha}} - \hat{\lambda}$ is consistent with downward sloping signature plots but reveal negative serial return correlation. Hence, our sample selection broadly confirms our theory.

Third, the speed of market efficiency, α, varies significantly over time and the cross-section. For instance, we observe strong price feedback and a comparably high market efficiency for Amazon and Google during the first two weeks of March 2014. In these cases, the observed price more closely follows the efficient price. Conversely, we observe that in the second half of March, the observed price and efficient price are almost de-coupled and independent of each other ($\hat{\alpha} = 0.066$). Nevertheless, in all cases, α is significantly below 1 which supports our claim that our model captures effects that are not accounted for in the classical ”martingale-plus-i.i.d. noise” model. In fact, adjustments of observed prices due to mis-pricing are more sluggish than implied by the latter setting. Our results suggest that the market mechanisms that induce feedback to pull back price to their efficient level play a significant role in describing the price variation of observed price returns.

Similarly, the signal-to-noise-ratio and market microstructure noise varies strongly over the cross-section. For instance, we find that there is literally no market microstructure noise contamination in the first two weeks of March for Amazon and Google as $\hat{\lambda}$ is effectively one. On the other hand, in the same period market, microstructure noise accounts for 50% of all changes in the observed price process for Intel.

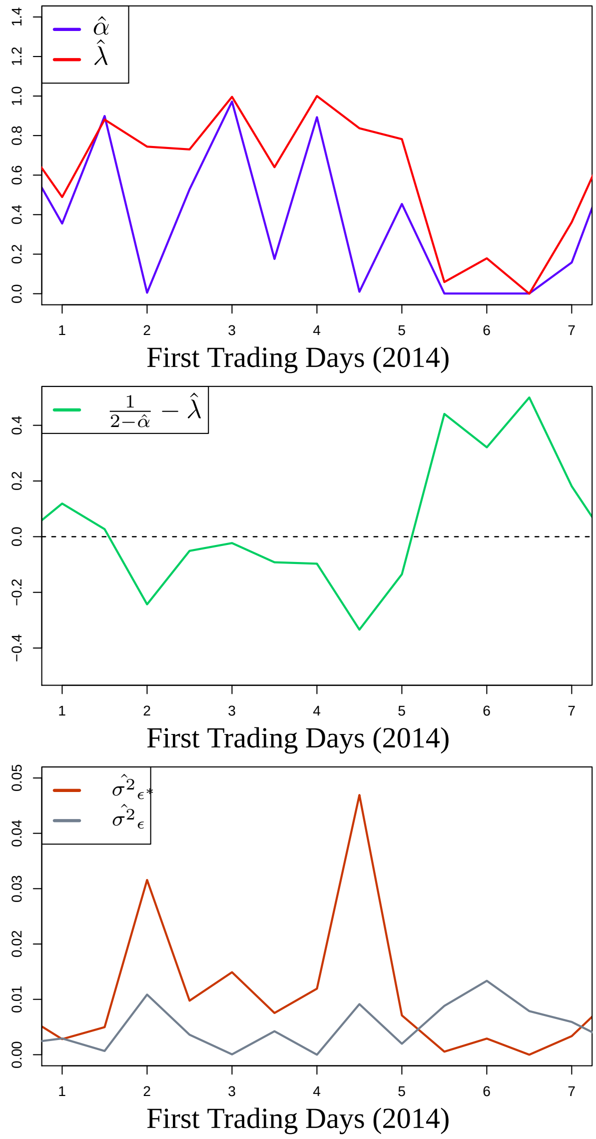
As a representative example, Figure 13 illustrates the time-varying nature of the model components for the Google stock for the first seven trading days of 2014. To reduce the computational burden, estimation is based on consecutive, non-overlapping trading intervals, with interval size being half a day and the underlying sampling interval being $\Delta = 2sec$. The first plot illustrates the time evolution of the estimates $\hat{\alpha}$ and $\hat{\lambda}$, while the second plot illustrates the time evolution of the estimated threshold $\frac{1}{2-\hat{\alpha}} - \hat{\lambda}$. Finally, the third plot illustrates the variance of the microstructure component $(\hat{\sigma^2}_{\epsilon})$ and the informational innovation in the efficient price $(\hat{\sigma^2}_{\epsilon^*})$.

Note that $\hat{\sigma^2}$ and $\hat{\sigma^2}_{\epsilon}$ are not estimated directly, but derive from the ML estimates $\hat{\lambda}$ and $\hat{\sigma^2}_{\mu}$. This is because, by definition, $\sigma^2_{\mu} = \sigma^2_{\epsilon^*} + \sigma^2_{\epsilon}$ and $\lambda = \frac{\sigma^2_{\epsilon^*}}{\sigma^2_{\mu}}$. Therefore, $\hat{\sigma^2}^*_{\epsilon}$ and $\hat{\sigma^2}_{\epsilon}$ are given by

$$\hat{\sigma^2}^*_{\epsilon} := \hat{\lambda}\hat{\sigma^2}_{\mu}, \qquad \hat{\sigma^2}_{\epsilon} := \left(1 - \hat{\lambda}\right)\hat{\sigma^2}_{\mu}. \tag{5.4}$$

The results support our idea that the key market characteristics, the noise levels and the efficiency of the market strongly vary even on an intraday level. Second, market efficiency α and the signal-to-noise ratio λ are strongly co-moving. In the case of Google, we see that markets tend to be more efficient when the signal-to-noise-ratio is high. Third, when informational innovations strongly dominate microstructure noise, the threshold $\frac{1}{2-\alpha} - \lambda$ becomes negative and the underlying regime changes. In the selected trading period, this happens at the end of the second trading day and during the fourth trading day of January 2014. Indeed, around these times, the informational noise level $\hat{\sigma^2}^*_{\epsilon}$ significantly overshoots the microstructure noise level $\hat{\sigma^2}_{\epsilon}$.

Figure 13: ML estimates ($\hat{\alpha}$, $\hat{\lambda}$, $\hat{\sigma^2}_{\epsilon^*}$ and $\hat{\sigma^2}_{\epsilon}$) for the stock Google, for non-overlapping, consecutive trading intervals of half a trading day throuth the first 7 trading days of 2014, starting at January 2nd. Estimation is based on the discrete state space model (2.6) and (2.7) with sampling interval $\Delta = 2sec$. The second picture below illustrates the time evolution of the threshold $\frac{1}{2-\hat{\alpha}} - \hat{\lambda}$. The dashed line in the second picture separates the two regimes: $\frac{1}{2-\hat{\alpha}} - \hat{\lambda} > 0$ and $\frac{1}{2-\hat{\alpha}} - \hat{\lambda} < 0$.



6 Conclusions

We propose a model establishing a novel link between observed high-frequency prices and the underlying (efficient) price process following a (semi-)martingale. The essential connection betwen both price processes is channeled through a feedback of observed prices as a response to prevailing mis-pricing in the market. The latter is given by the discrepancy between observed prices and ”efficient” prices. The rate by which observed prices react to mis-pricing is governed by a parameter which is naturally interpreted as a measure for market (in)efficiency and plays a crucial role for the resulting dynamics in observed returns.

Despite its simplicity, the proposed model yields a flexible framework to model high-frequency asset price behavior and to establish the interplay between the underlying ”efficient” volatility, market microstructure noise and the rate of market inefficiency. While the classical ”martingale-plus-i.i.d. noise” setting is contained as a special case, the proposed model captures many more settings and provides a novel view on commonly observed high-frequency phenomena. In this sense, the approach links together aspects from high-frequency based volatility estimation and market microstructure analysis.

Most importantly, the model allows to identify different regimes where the dynamic properties of observed high-frequency returns are significantly different. We show that these regimes depend on the magnitude of the signal-to-noise relative to the market inefficiency and provide a natural link to underlying trading behavior and the way how information is channeled through the market.

We moreover show that the existence of these regimes has a profound effect on realized volatilities sampled on high frequencies. We demonstrate that the sampling limit of the latter behave very differently depending on the underlying regime. In particular, we illustrate that volatility signature plots can be upward sloped as well as downward sloped, respectively. Our model thus provides a very natural explanation for the presence of downward sloped signature plots (for increasing sampling frequencies) which is not easily explained in the existing literature.

We moreover show strong empirical support for the model and illustrate that the regimes postulated by the model can be identified based on typical high-frequency processes. Utilizing high-frequency data from NASDAQ trading we demonstrate the dynamics of high-frequency returns can locally change and are in line with the regimes predicted by the model parameters. Likewise, we provide novel empirical evidence for locally changing volatility signature plots. We particularly demonstrate that downward sloped signature plots occur much more frequently than commonly believed.

We discuss a simplified version of the model assuming discrete time and constant underlying volatility. In such a setting, all model parameters (including volatility) are straightforwardly estimated by maximum likelihood using the Kalman filter or moment matching. An interesting remaining question is whether our method of backing out volatility estimates from high-frequency data is superior if benchmarked against existing approaches. If we stick to the simplified (fully parametric) framework, we should expect gains as our approach is more general as the classical ”martingale-plus-noise” setting and thus should be able to identify underlying volatility more precisely.

Likewise, even more efficiency gains can be expected if we impose even more structure on the microstructure noise and the modeled feedback in response to mispricing in the market. Here, the link to market microstructure literature could be exploited bringing this literature together with the statistical literature on high-frequency volatility estimation.

Finally, the proposed framework is straightforwardly extended in various directions. For instance, relaxing the assumption of constant volatility and/or augmenting the model by a jump component could bring up interesting challenges for the statistical literature on volatility estimation.

References

AIT-SAHALIA, Y. AND J. JACOD (2014): *High-Frequency Financial Econometrics*, Princeton University Press. 2

ANDERSEN, T. G., T. BOLLERSLEV, F. DIEBOLD, AND P. LABYS (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625. 3

ANDERSEN, T. G., T. BOLLERSLEV, AND F. X. DIEBOLD (2010): "Parametric and Nonparametric Volatility Measurement," in *Handbook of Financial Econometrics*, ed. by Y. Ait-Sahalia and L. P. Hansen, Elsevier, 67–137. 2

CHAKER, S. (2013): "Volatility and Liquidity Costs," Tech. Rep. 2013-29, Bank of Canada. 2

DIEBOLD, F. AND G. STRASSER (2013): "On the Correlation Structure of Microstructure Noise: A Financial Economic Approach," *Review of Economic Studies*, 80, 1304337. 2

HANSEN, P. R. AND A. LUNDE (2006): "Realized Variance and Market Microstructure Noise," *Journal of Business and Economic Statistics*, 24, 127–161. 4, 13

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. 3, 7

KYLE, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53 (6), 1315–1335. 2

SHEPPARD, K. (2013): "Measuring Market Speed," Tech. rep., University of Oxford. 3

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, 1394411. 3, 5

Appendices

Appendix A Empirical Results

A.1 Robustness: Quadratic Variation Estimates Based on De-trended Returns

Figure 14: Averaged daily realized variance for six different trading periods for the stock Apple (AAPL) based on de-trended returns. Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for fixed sub-sampling interval Δ.

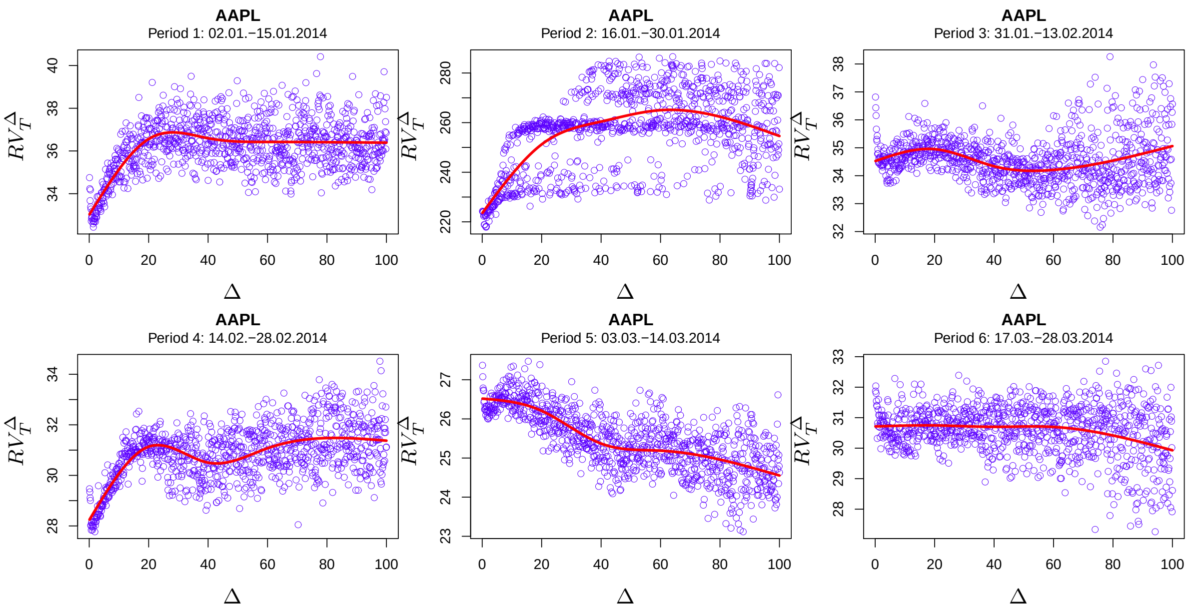


Figure 15: Averaged daily realized variance for six different trading periods for the stock Amazon (AMZN) based on de-trended returns. Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for fixed sub-sampling interval Δ.

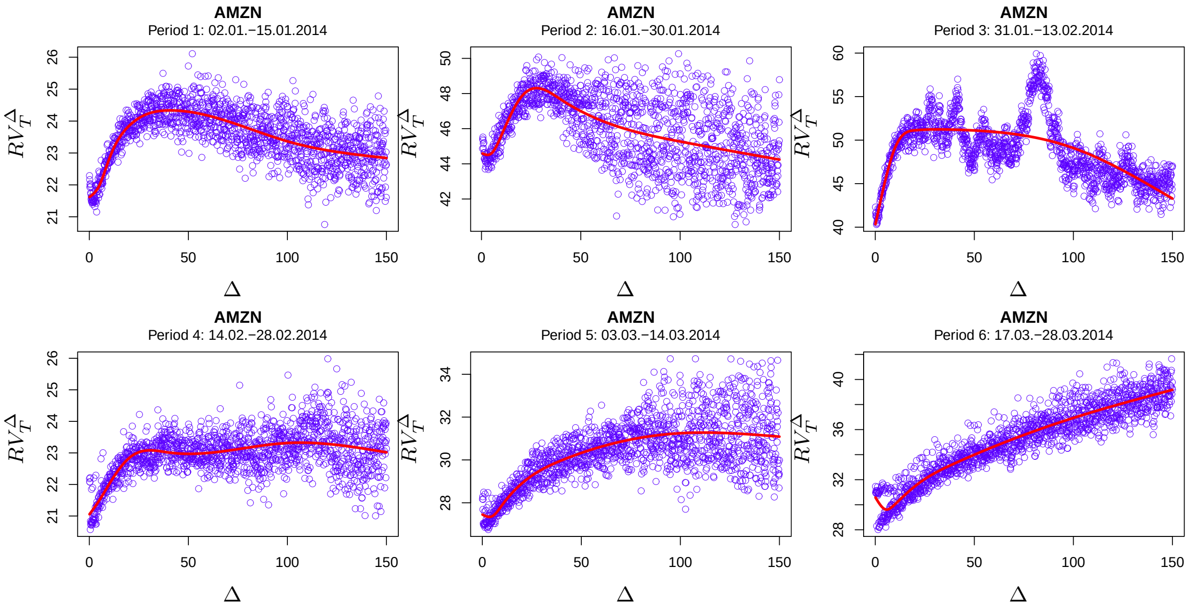


Figure 16: Averaged daily realized variance for six different trading periods for the stock Cisco (CSCO) based on de-trended returns. Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for fixed sub-sampling interval Δ.

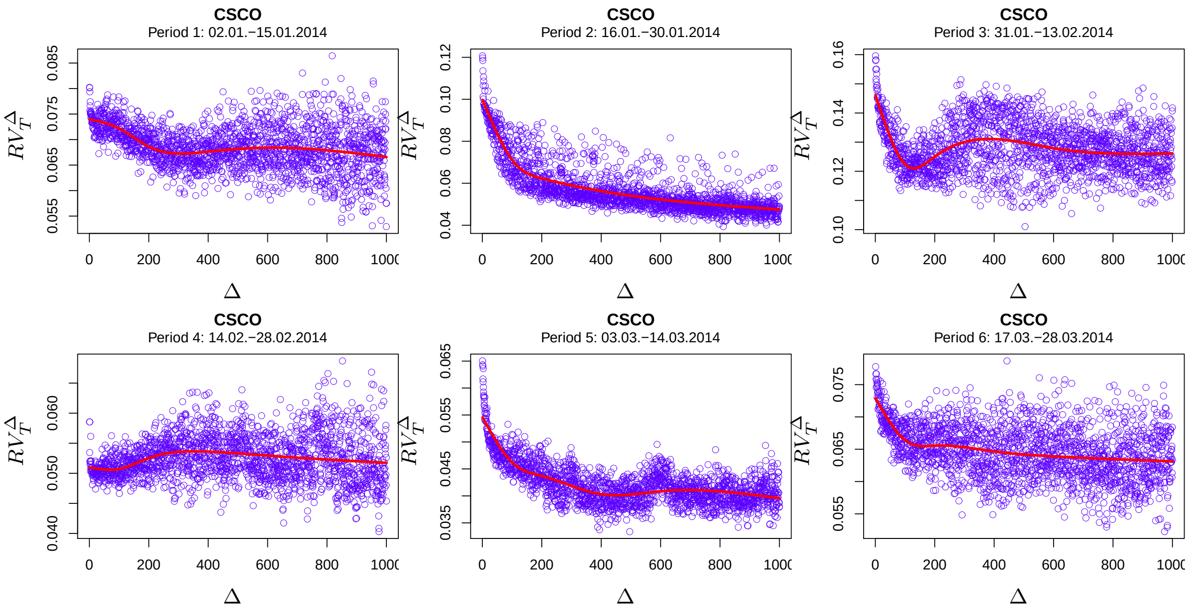


Figure 17: Averaged daily realized variance for six different trading periods for the stock Ebay (EBAY) based on de-trended returns. Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for fixed sub-sampling interval Δ.

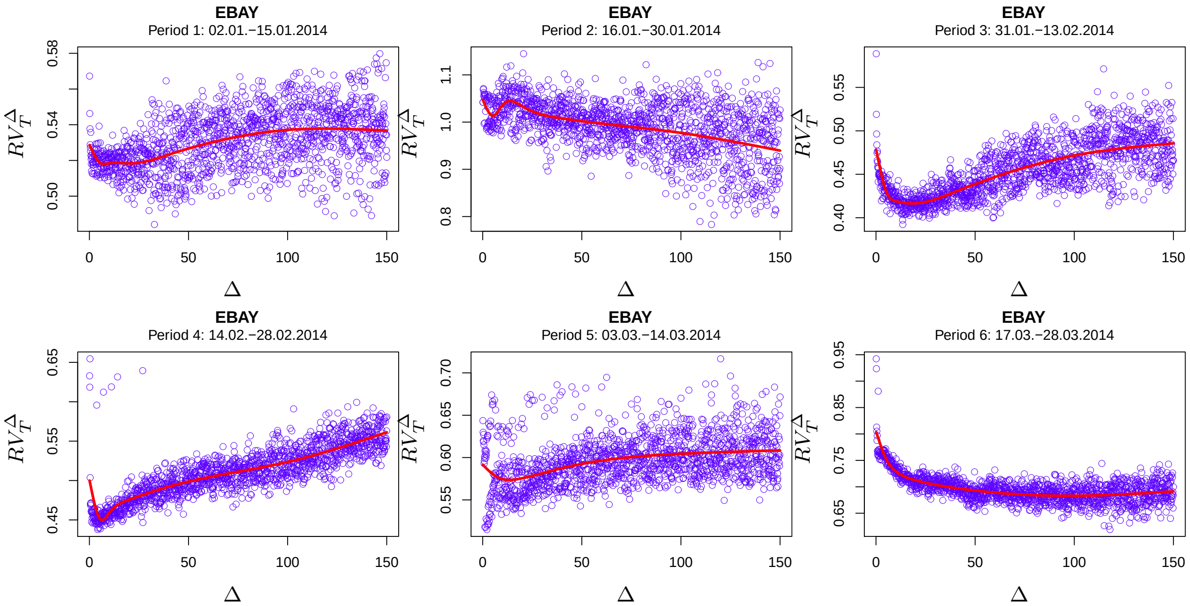


Figure 18: Averaged daily realized variance for six different trading periods for the stock Google (GOOG) based on de-trended returns. Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for fixed sub-sampling interval Δ.

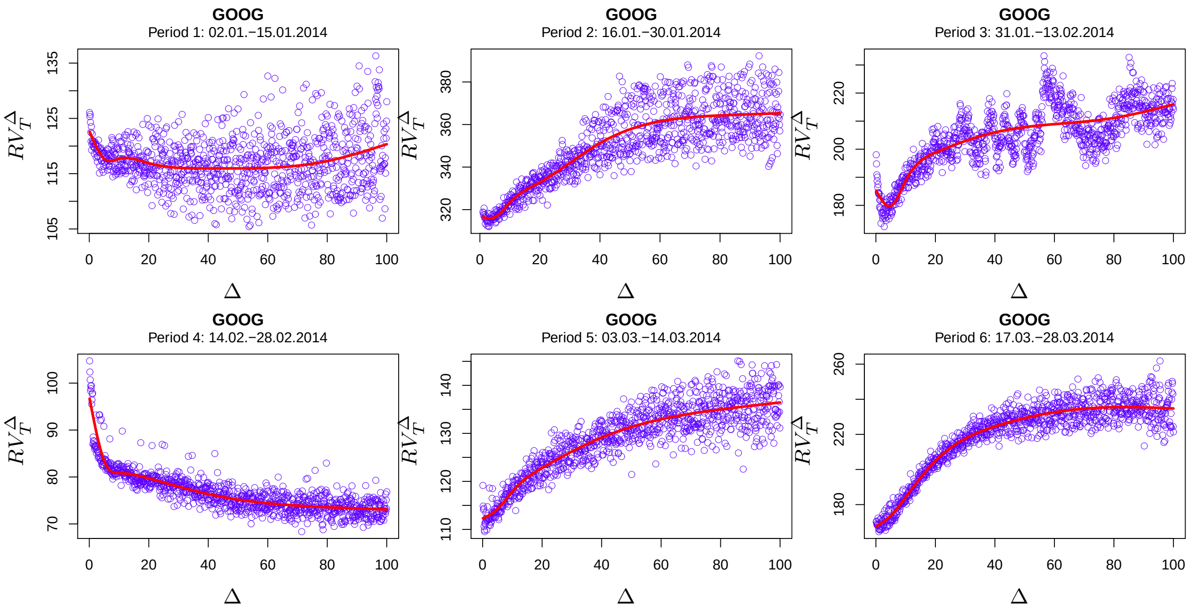


Figure 19: Averaged daily realized variance for six different trading periods for the stock Intel (INTC) based on de-trended returns. Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for fixed sub-sampling interval Δ.

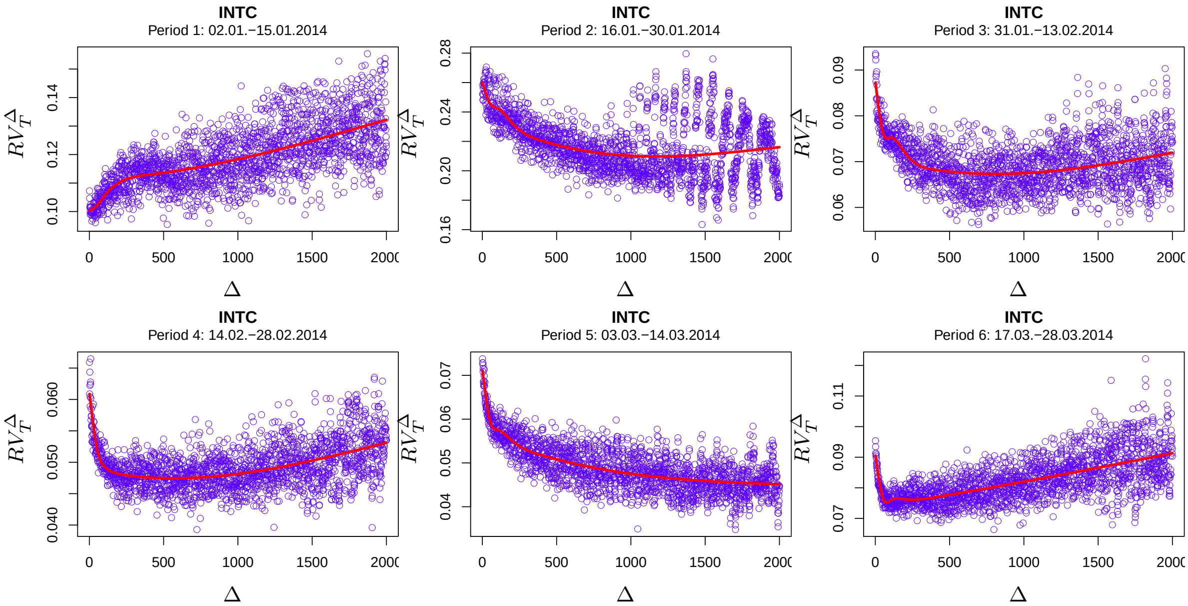
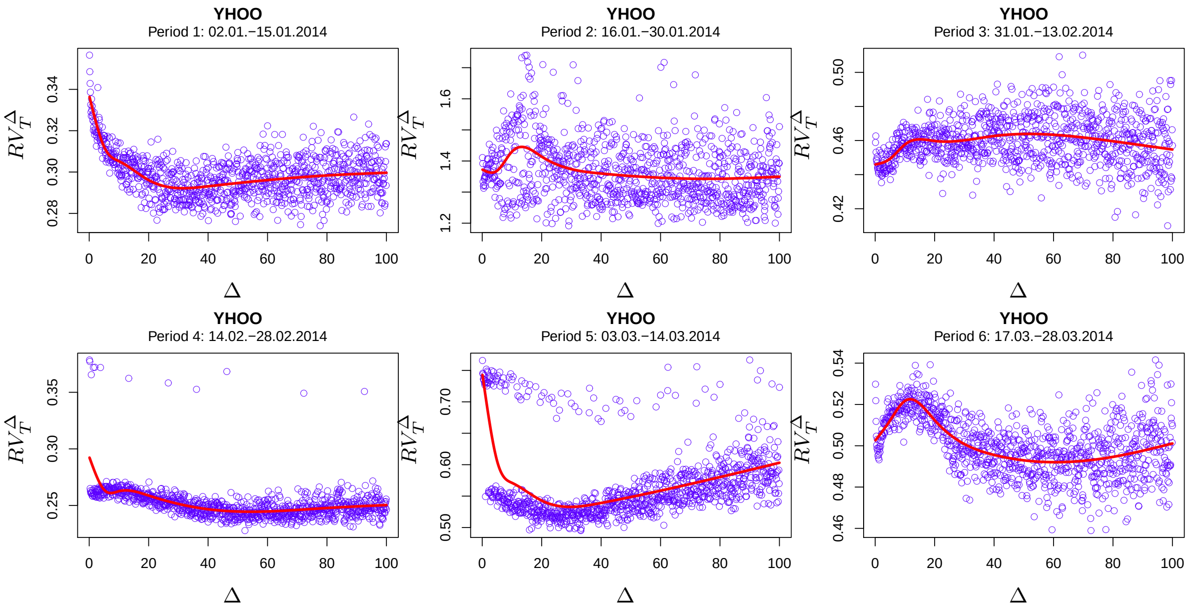


Figure 20: Averaged daily realized variance for six different trading periods for the stock Yahoo (YHOO) based on de-trended returns. Each trading period consists of 10 trading days, starting January 2nd 2014 and ending March 3rd 2014. The estimated realized variance is shown for fixed sub-sampling interval Δ.



Appendix B 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 4 (Support Lemma)**.** *The cross-moments of ϵ_t and μ_t obey*

(i) $\mathbb{E}[\mu_{t-\Delta}\epsilon_t] = 0 \qquad \forall\Delta > 0,$

(ii) $\mathbb{E}[\mu_{t+\Delta}\epsilon_t] = (1-\alpha)^{\Delta}\sigma_\epsilon^2 \qquad \forall\Delta \geq 0.$

Proof. Observe that because of (2.5), we have $\mu_{t+1} = -\alpha\mu_t + \epsilon^\mu_{t+1}$. Hence, through backward recursion, we obtain the moving average representation for $\mu_{t+\Delta}$ and μ_t, i.e.

$$\mu_{t+\Delta} = (1-\alpha)^{\Delta+1}\mu_{t-1} + \sum_{i=0}^{\Delta}(1-\alpha)^i\epsilon^\mu_{t+\Delta-i}, \tag{B.1}$$

$$\mu_t = (1-\alpha)^t\mu_0 + \sum_{i=0}^{t}(1-\alpha)^i\epsilon^\mu_{t-i}. \tag{B.2}$$

Also note that the inefficiency error term is defined as $\epsilon^\mu_t = \epsilon_t - \epsilon^*_t$ and that ϵ_t and ϵ^*_t are mutually independent and are iid random noise. Therefore

$$\mathbb{E}[\epsilon_t\epsilon^\mu_k] = \delta(t-k)\sigma^2_\epsilon, \tag{B.3}$$

where δ denotes the Dirac-Delta function. With this preparation, we prove the lemma.

(i) Using (B.2), we have for $t > 0$ and $\Delta > 0$

$$\mathbb{E}[\mu_{t-\Delta}\epsilon_t] = (1-\alpha)^{t-\Delta}\underbrace{\mathbb{E}[\epsilon^\mu_0\epsilon_t]}_{=0} + \sum_{i=0}^{t-\Delta}(1-\alpha)^i\mathbb{E}[\epsilon^\mu_{t-\Delta-i}\epsilon_t] \tag{B.4}$$

The last terms vanisesh because of(B.3) and the fact that $t-\Delta- \leq t-\Delta < t$.

(ii) With (B.1), we obtain

$$\mathbb{E}[\mu_{t+\Delta}\epsilon_t] = (1-\alpha)^{\Delta+1}\underbrace{\mathbb{E}[\mu_{t-1}\epsilon_t]}_{=0} + \sum_{i=0}^{\Delta}(1-\alpha)^i\mathbb{E}[\epsilon_t\epsilon^\mu_{t+\Delta-i}]$$

The first term is zero because of (iii) that have proven. For the second term, we use (B.3) and obtain

$$= (1-\alpha)^{\Delta}\sigma^2_\epsilon. \tag{B.5}$$

□

Appendix C Main Results

Proof of Lemma 2.

$$Cov(r_{t+\Delta}, r_t) = \mathbb{E}[(r_{t+\Delta} - \underbrace{\mathbb{E}[r_{t+\Delta}]}_{=0})(r_t - \underbrace{\mathbb{E}[r_t]}_{=0})] = \mathbb{E}[r_{t+\Delta} r_t] = \mathbb{E}[(p_{t+\Delta} - p_{t+\Delta-1})(p_t - p_{t-1})]$$

Using $p_t = \mu_t + p_t^*$, *we replace* p_t *and get*

$$= \mathbb{E}[(\mu_{t+\Delta} - \mu_{t+\Delta-1} + \hat{p}_{t+\Delta} - \hat{p}_{t+\Delta-1})(\mu_t - \mu_{t-1} + \hat{p}_t - \hat{p}_{t-1})]$$

Now we use (2.5) *and obtain*

$$\begin{aligned}
&= \mathbb{E}\Big[\Big(-\alpha\mu_{t+\Delta-1} + \underbrace{\epsilon^{\mu}_{t+\Delta} + \hat{\epsilon}_{t+\Delta}}_{=\epsilon_{t+\Delta}}\Big)\Big(-\alpha\mu_{t-1} + \underbrace{\epsilon^{\mu}_t + \hat{\epsilon}_t}_{=\epsilon_t}\Big)\Big] \\
&= \mathbb{E}\Big[\Big(-\alpha\mu_{t+\Delta-1} + \epsilon_{t+\Delta}\Big)\Big(-\alpha\mu_{t-1} + \epsilon_t\Big)\Big] \\
&= \alpha^2\mathbb{E}[\mu_{t+\Delta-1}\mu_{t-1}] - \alpha\mathbb{E}[\mu_{t+\Delta-1}\epsilon_t] - \alpha\underbrace{\mathbb{E}[\epsilon_{t+\Delta}\mu_{t-1}]}_{=0} + \underbrace{\mathbb{E}[\epsilon_{t+\Delta}\epsilon_t]}_{=0}
\end{aligned}$$

Where we used Lemma 4 and the fact that ϵ_t *is iid.*

$$= \alpha^2\mathbb{E}[\mu_{t+\Delta-1}\mu_{t-1}] - \underbrace{\alpha\mathbb{E}[\mu_{t+\Delta-1}\epsilon_t]}_{=(1-\alpha)^{\Delta-1}\sigma^2_\epsilon}$$

The first term is given in (2.11)*, while the last term derives from Lemma 4. Thus, we finally obtain*

$$\begin{aligned}
&= (\sigma^2_{\epsilon^*} + \sigma^2_\epsilon)\alpha^2\frac{(1-\alpha)^\Delta}{1-(1-\alpha)^2} - \alpha(1-\alpha)^{\Delta-1}(\sigma^2_{\epsilon^*} + \sigma^2_\epsilon) + \alpha(1-\alpha)^{\Delta-1}\sigma^2_{\epsilon^*} \\
&= \underbrace{\alpha(1-\alpha)^{\Delta-1}(\sigma^2_{\epsilon^*} + \sigma^2_\epsilon)}_{:=\psi(\Delta-1)}\Bigg(\underbrace{\Bigg(\frac{\alpha-\alpha^2}{1-(1-\alpha)^2} - \frac{1-(1-\alpha)^2}{1-(1-\alpha)^2}\Bigg)}_{=-\frac{1}{2-\alpha}} + \lambda\Bigg) \\
&= \psi(\Delta-1)\left(\lambda - \frac{1}{2-\alpha}\right).
\end{aligned} \tag{C.1}$$

□

Proof of Theorem 1. The proof to Theorem 1 is derived in a similar fashion as in the proof to Lemma 2. For sake brevity, we leave the explicit proof to the reader. □

Proof of Lemma 3. There is direct proof of Lemma 3, exploiting the fact that the terms A_i, B_i and C_i are independent and Gaussian. Then, the quadratic variation the quadratic variation is obtained by calculating the expectation of each term separately and sum them up..

A more elegant proof is by considering a diffusion limit of the discrete time case as of Lemma 2.

To this end, re-write the discrete-time quadratic variation in terms of infinitesimal time increments dt as follows

$$\alpha = \widetilde{\alpha} dt, \qquad \Delta = \frac{\widetilde{\Delta}}{dt}. \tag{C.2}$$

Now, the proof follows from the fact that (for fixed $\widetilde{\alpha}$ and $\widetilde{\Delta}$), the following limits hold

$$\lim_{dt \to 0} \alpha = 0, \qquad \lim_{dt \to 0} \alpha\Delta = \widetilde{\alpha}\widetilde{\Delta}. \qquad \lim_{dt \to 0} (1-\alpha)^{\Delta} = e^{-\widetilde{\Delta}\widetilde{\alpha}}. \tag{C.3}$$

Applying these limits to (4.2) and – by slight abuse of notation – using the identification that in the diffusion limit α is $\widetilde{\alpha}$ and Δ is $\widetilde{\Delta}$, gives the result. $\square$