# Adaptive Realized Kernels

Marine Carrasco

Associate Professor

Department of Economics - Université de Montreal

and

Rachidi Kotchoni

Visiting Scholar

Department of Economics - Université de Montréal

Address: C.P. 6128, succursale Centre-ville. Montréal, QC, H3C 3J7. Canada

Email: rachidi.kotchoni@umontreal.ca

**Abstract**

We design adaptive realized kernels to estimate the integrated volatility in a framework that combines a stochastic volatility model with leverage effect for the efficient price and a semiparametric microstructure noise model specified at the highest frequency. Some time dependence parameters of the noise model must be estimated before adaptive realized kernels can be implemented. We study their performance by simulation and illustrate their use with twelve stocks listed in the Dow Jones Industrial. As expected, we find that adaptive realized kernels achieves the optimal trade-off between the discretization error and the microstructure noise.

Keywords: Integrated Volatility, Method of Moment, Microstructure Noise, Realized Kernels.

JEL Codes: C13, C14, G10

To estimate the monthly variance of a financial asset return, Merton (1980) proposes to use "*the sum of the squares of the daily logarithmic returns on the market for that month with appropriate adjustments for weekends and holidays and for the no-trading effect which occurs with a portfolio of stocks*". Unfortunately, the daily data available to Merton does not span a long enough period for the purpose of his study. He circumvents this difficulty by using a moving average of monthly squared logarithmic return. In the same vein, French, Schwert and Stambaugh (1987) estimate the monthly variances by the sum of squared returns plus twice the sum of product of adjacent returns to correct for the first order autocorrelation bias. Andersen and Bollerslev (1998) are the first to support the empirical use of the realized volatility (RV) as an estimator of integrated volatility (IV) by a rigorous consistency argument taken from Karatzas and Shreve (1988). Since then, many authors including Jacod and Protter (1998) and Barndorff-Nielsen and Shephard (2002) have well established the consistency of the RV for the IV when prices are observed without error.

However, it is commonly admitted that recorded stock prices are contaminated with market microstructure noise. As pointed out by Andersen and Bollerslev (1998), "*... because of discontinuities in the price process and a plethora of market microstructure effects, we do not obtain a continuous reading from a diffusion process...*". Barndorff-Nielsen and Shephard (2002) show that in the presence of jumps that cause the price to exhibit discontinuities, the RV is consistent for the total quadratic variation of the price process. But the presence of noise in measured prices causes the RV computed with high frequency data to be a biased estimator of the object of interest. The sources of noise are discussed in Stoll (1989) and Hasbrouck (1993). In the words of Hasbrouck (1993), the *"pricing errors"* are mainly due to *"... discreteness, inventory control, the non-information based component of the bid-ask spread, the transient component of the price response to a block trade, etc."*.

Many approaches have been proposed in the literature to deal with the noise. One of them consists of sampling sparsely the high frequency returns so as to mitigate the impact of the noise sufficiently[1]. Zhou (1996) and Hansen and Lunde (2006) propose a bias correction approach while Bollen and Inder (2002) and Andreou and Ghysels (2002) advocate filtering techniques. Under the assumption that the volatility of the high frequency returns are constant within a day, Ait-Sahalia, Mykland and Zhang (2005) derive a maximum likelihood estimator of the IV that is robust to both IID noise and distributional misspecification. Zhang, Mykland, and Ait-Sahalia (2005) propose another consistent estimator in the presence of IID noise which they called the two scale realized volatility. This estimator has been adapted in Ait-Sahalia, Mykland and Zhang (2011) to deal with dependent noise. Since then, other consistent estimators have become available among which the well-known realized kernels of Barndorff-Nielsen et al. (2008a) and the pre-averaging estimator of Podolskij and Vetter (2009).[2]

This paper presents a general framework to design adaptive and efficient kernel-based estimators for the integrated volatility in accordance with the properties of the noise. First, we propose a semi-parametric microstructure noise model that is tied to the frequency at which the price data are recorded. The noise is specified as the sum of an endogenous term that is correlated with the efficient returns and an exogenous term that is uncorrelated with the efficient returns. Flexible restrictions are imposed on the exogenous noise so that it admits $L$-dependent and AR(1) dynamics as special cases. We superimpose the overall noise model to a stochastic volatility model with leverage effect for the efficient price.

Second, we examine the implications of the overall framework for common realized measures that are aimed at estimating the IV. The bulk of the MSE of IV estimators is dominated by the contribution of the exogenous noise. When price data are contaminated with the endogenous noise only, the bias of the standard RV is $O(1)$ while kernel-based estimators are unbiased and consistent. Under an MA(L) exogenous noise, realized Bartlett kernels with bandwidth larger than

$L$ are unbiased for the IV. When the exogenous noise is AR(1), an unbiased estimator of IV can be obtained only upon having a first step estimator of the noise autoregressive root in hand. If the first order autocorrelation of the exogenous noise converges to one as the record frequency goes to infinity, then a necessary condition for the realized kernels to be consistent for the IV is that its bandwidth diverge sufficiently fast as the record frequency goes to infinity.

Third, we examine the trade-off involved as one moves, on the one hand, from the standard RV to a bias-corrected RV, and on the other hand, from a bias-corrected RV to a consistent realized kernel. We show that unbiasedness and/or consistency are achieved by conceding more and more discretization error. Acting on this, we argue that the performance of any IV estimator at a given sampling frequency reflects the balance between the discretization error and the microstructure noise at that frequency. We propose an adaptive realized kernel that achieves the optimal trade-off bewteen both types of errors. As an optimal linear combination of initial estimators, the adaptive realized kernel provides us with an additional degree of freedom for tuning kernel-based estimators besides the bandwidth parameter.

Fourth, we propose two inference procedures for the microstructure noise. The first procedure is designed for AR(1) types of noise and it is based on an overidentified generalized method of moments. The second procedure is designed for MA(L) noises and it uses as many moment conditions as there are parameters to be estimated. The AR(1) assumption best suits for noise processes with infinite dependence lag while the MA(L) assumption is reasonable if the noise has finite dependence. Our simulations show that the inference procedure designed for AR(1) noises has good size and it has power against MA(L) alternatives. Hence, our best investigation strategy in practice consists of first testing whether the noise is AR(1) and next, applying the MA(L) inference procedure if the AR(1) specification is rejected. We apply this strategy to twelve stocks listed in the Dow Jones Industrial and find that the AR(1) noise model cannot be rejected for six of them. For the other six stocks, we apply the MA(L) noise inference procedure and find estimates of L that lie between 8 and 12 minutes.

The paper is organized as follows. In Section 1, we present our models for the frictionless price and for the microstructure noise. In section 2, we study the properties of common realized measures within our framework. In Section 3, we design adaptive realized kernels for the IV. Our inference procedures for microstructure noise are presented in Section 4. In Sections 5 we evaluate the performance of all estimators proposed in the paper by simulation. Section 6 presents the empirical application and Section 7 concludes. The proofs are collected in an appendix.

# 1 The Framework

First, we present a model for the efficient price that allows for leverage effect. Next, we present our model for the microstructure noise.

## 1.1 A Model for the Efficient Price

Let $p_s^*$ denote the latent (or efficient) log-price of an asset and $p_s$ its observable counterpart. Assume that $p_s^*$ obeys the following stochastic differential equation:

$$dp_s^* = \mu\left(s, p_s^*\right) ds + \sigma_s d\widetilde{W}_s; \quad p_0^* = 0, \tag{1}$$

where $\mu\left(.,.\right)$ is a deterministic and smooth function, $\sigma_s$ is the spot volatility and $\widetilde{W}_s$ is a standard Brownian motion. In turn, assume that $\sigma_s$ satisfies:

$$d\sigma_s = f\left(\sigma_s\right) ds + g\left(\sigma_s\right) dB_s, \tag{2}$$

where $f\left(.\right)$ and $g\left(.\right)$ are deterministic and smooth functions, $B_s$ is a Brownian motion such that $\widetilde{W}_s = \rho B_s + \sqrt{1-\rho^2} W_s$, $W_s$ is another Brownian motion that is independent of $B_s$ and $\rho \in (0,1)$ is the leverage effect parameter.

It is assumed that Equation (2) admits a continuous solution with infinite lifetime so that any power function of the spot volatility process $\left\{\sigma_s\right\}_{s\geq 0}$ is locally integrable with respect to the Lebesgue Measure. Also, the processes $\mu\left(s, p_s^*\right)$, $f\left(\sigma_s\right)$ and $g\left(\sigma_s\right)$ are assumed adapted to the filtration generated by $\left\{W_u, B_u, u < s\right\}$. Throughout this paper, it is maintained that there is no jump in the efficient price. However, the conclusions of our analysis remain valid if a jump component that is uncorrelated with all other randomness is added to the efficient price. In this case, the estimators that we consider for the IV are designed for the total quadratic variation of the efficient price process.[3] Without loss of generality, we condition all our analysis on the volatility path $\left\{\sigma_s\right\}_{s\geq 0}$ but the conditioning is often removed from the notation for simplicity. Accordingly, all deterministic transformations of the spot volatility process are treated as constants.

We assume that there exists a twice differentiable deterministic function $p^*_{(1)}(.)$ that satisfies $\frac{\partial p^*_{(1)}(\sigma_s)}{\partial \sigma_s} = \frac{\rho\sigma_s}{g(\sigma_s)}$ so that the stochastic process $p^*_{(2),s} = p^*_s - p^*_{(1)}(\sigma_s)$ follows a diffusion without leverage effect[4]. Indeed, by the Itô Lemma, we have:

$$dp^*_{(1),s} = \mu_{(1),s}ds + \sigma_{(1),s}dB_s, \text{ and} \tag{3}$$

$$dp^*_{(2),s} = \mu_{(2),s}ds + \sigma_{(2),s}dW_s. \tag{4}$$

where $p^*_{(1),s} \equiv p^*_{(1)}(\sigma_s)$, $\sigma_{(1),s} = \rho\sigma_s$, $\sigma_{(2),s} = \sqrt{1-\rho^2}\sigma_s$, $\mu_{(2),s} = \mu_s - \mu_{(1),s}$ and

$$\mu_{(1),s} = \frac{\rho\sigma_s}{g(\sigma_s)} f(\sigma_s) + \frac{1}{2}\frac{\partial^2 p^*_{(1)}(\sigma_s)}{\partial\sigma^2} g^2(\sigma_s).$$

By construction, $p^*_s = p^*_{(1),s} + p^*_{(2),s}$ and $p^*_{(1),s}$ and $p^*_{(2),s}$ are uncorrelated. Hence, $IV_t = \int_{t-1}^{t}\sigma_s^2 ds$ is equal to the sum of the quadratic variations of $p^*_{(1),s}$ and $p^*_{(2),s}$.

We consider a sampling scheme where the unit period is normalized to one day. By definition, the microstructure noise is the difference between the observed log-price and the efficient log-price, that is, $u_s = p_s - p^*_s$. Thus, let $r^*_t$ denote the latent log-return at day $t$ and $r_t$ its observable counterpart. We have:

$$r_t \equiv p_t - p_{t-1} = r^*_{(1),t} + r^*_{(2),t} + u_t - u_{t-1} \tag{5}$$

where $r^*_{(i),t} = \int_{t-1}^{t}\mu_{(i),s}ds + \int_{t-1}^{t}\sigma_{(i),s}dW_s$. The drifts of the diffusions (1), (3) and (4) are irrelevant for their quadratic variations. Acting on this, we treat these diffusions as though they had no drift ($\mu_s = \mu_{(1),s} = \mu_{(2),s} = 0$).

Suppose that we observe to a large number $m$ of intradaily returns $r_{t,1}, r_{t,2}, ..., r_{t,m}$ for $t = 1, ..., T$ days. We have:

$$r_{t,j} = r^*_{(1),t,j} + r^*_{(2),t,j} + u_{t,j} - u_{t,j-1} \text{ for all } t \text{ and } j, \tag{6}$$

where $u_{t,j} \equiv u_{t-1+j/m}$ and $r^*_{(i),t,j} \equiv \int_{t-1+(j-1)/m}^{t-1+j/m}\sigma_{(i),s}dW_s$. It is maintained that the high frequency observations are equidistant in calendar time. The noise-contaminated (observed) and true (latent)

RV computed at frequency $m$ are:

$$RV_t^{(m)} = \sum_{j=1}^{m} r_{t,j}^2 \text{ and } RV_t^{*(m)} = \sum_{j=1}^{m} r_{t,j}^{*2}.$$

Barndorff-Nielsen and Shephard (2002) show that $RV_t^{*(m)}$ converges to $IV_t$ and derived the asymptotic distribution:

$$\frac{RV_t^{*(m)} - IV_t}{\sqrt{\frac{2}{3}\sum_{j=1}^{m} r_{t,j}^{*4}}} \to N(0,1), \tag{7}$$

as $m$ goes to infinity. In the presence of microstructure noise, the estimator $RV_t^{*(m)}$ is not feasible.

### 1.2 A Semiparametric Model for the Microstructure Noise

To model the microstructure noise, we posit that the frequency at which the price data are recorded determines the time series properties of the microstructure noise. This idea is acknowledge by Barndorff-Nielsen et al. (2008a, *Section 5.4*) who considered "*a situation where the serial dependence is tied to the sampling frequency [...], as opposed to calendar time*". Here, we follow a semiparametric approach that consists of specifying how the correlation structure of the noise changes as the record frequency increases.

To motivate this approach, let us consider an MA(1) process $\varepsilon_{t,j}$ at the highest frequency with $E(\varepsilon_{t,j}^2) = \omega_0$ and $E(\varepsilon_{t,j}\varepsilon_{t,j-1}) = \omega_1$. The time elapsed between $\varepsilon_{t,j}$ and $\varepsilon_{t,j-h}$ is $\frac{h}{m}$ when the record frequency is $m$. By letting $\omega\left(\frac{h}{m}, m\right)$ denote the $h^{th}$ order autocovariance of $\varepsilon_{t,j}$, we have:

$$\omega(0,m) = \omega_0,\ \omega\left(\frac{1}{m}, m\right) = \omega_1 \text{ and } \omega\left(\frac{h}{m}, m\right) = 0,\ h \geq 2. \tag{8}$$

If we posit that $\varepsilon_{t,j}$ remains an MA(1) with constant parameters whatever the record frequency, then we can assert that:

$$\omega(0,km) = \omega_0,\ \omega\left(\frac{1}{km}, km\right) = \omega_1 \text{ and } \omega\left(\frac{h}{km}, km\right) = 0,\ h \geq 2, \tag{9}$$

as $k \to \infty$ and $m$ is fixed. However, if we assume that $\varepsilon_{t,j}$ obeys an MA(1) model at the record frequency $m$ but its first order autocorrelation is not invariant with respect to $m$, then (8) cannot be used to infer (9). By contrast, the autocorrelation structure of the sparsely sampled noise process can always be inferred from the properties of the noise at the highest frequency.

With this in mind, we postulate the following microstructure noise model at the record frequency:

$$u_{t,j} = a_{t,j} r^*_{(2),t,j} + \varepsilon_{t,j}, \; j = 1, 2, ..., m, \text{ for all } t, \tag{10}$$

where $a_{t,j}$ is a time varying coefficient that depends on the spot volatility process and $\varepsilon_{t,j}$ is independent of the efficient returns. In the words of Hasbrouck (1993), $\varepsilon_{t,j}$ is the information uncorrelated or exogenous pricing error while $a_{t,j} r^*_{(2),t,j}$ is the information correlated or endogenous pricing error. We assume that time dependence in the noise process is only due to its information uncorrelated part. The following assumptions are further made:

**Assumption E0.** $a_{t,j} = \beta_0 + \frac{\beta_1}{\sqrt{m}\sigma^*_{t,j}}$, where $\beta_0$ and $\beta_1$ are constants and:

$$\sigma^{*2}_{t,j} \equiv \int_{t-1+(j-1)/m}^{t-1+j/m} \sigma^2_s ds.$$

**Assumption E1.** For fixed $m$, $\varepsilon_{t,j}$ is a zero mean, discrete time and stationary process that is independent of $\{\sigma_s\}$ and $r^*_{t,j}$.

**Assumption E2.**

E2(a) $E(\varepsilon_{t,j}\varepsilon_{t,j-h}) \equiv \omega\left(\frac{h}{m}, m\right) = \omega_{m,h}$, $0 \leq \frac{h}{m} \leq 1$ and $\omega_{m,h} = 0$ for all $h > m$, where $\omega\left(\frac{h}{m}, m\right)$ is bounded.

E2(b) $\omega(0, m) = \omega_0$ for all $m$, and $\frac{\omega_{m,h} - \omega_{m,h+1}}{\omega_0} = O(m^{-\alpha})$ for some $\alpha \geq 0$, $h = 0, ..., m-1$.

**Assumption E3.** For fixed $m$, we have:

E3(a) $E\left|u_{t,j}u_{t,j-h}\right|^{4+\epsilon} < \infty$, for some $\epsilon > 0$, for all $h$.

E3(b) $Var\left(n^{-1/2}m^{-1/2}\sum_{t=t'+1}^{t'+n}\sum_{j=1}^{m} r_{t,j}r_{t,j-h}\right) \to q_h$, uniformly in any $t'$, as $n \to \infty$.

Assumption E0 is a convenient way to depart from the constant coefficient case ($a_{t,j} = \beta_0$) and it implies that the variance of the endogenous part of the noise goes to zero at rate $m$:

$$Var\left(a_{t,j} r^*_{(2),t,j}\right) = \left(1 - \rho^2\right)\left(\beta_0 \sigma^{*2}_{t,j} + 2\beta_0\beta_1\sqrt{\frac{\sigma^{*2}_{t,j}}{m}} + \frac{\beta_1^2}{m}\right).$$

Assumption E1 is quite standard in the literature. Assumption E2 stipulates general restrictions on the autocovariance structure of $\varepsilon_{t,j}$ rather a parametric distribution. Assumption E2(a) imposes that $\varepsilon_{t,j}$ be autocorrelated across $j$ within the same day $t$, but not across days. This approximation is reasonable if the market closes at 4:30pm and re-opens the next day at 9:00am. Overall, Assumption E2 is consistent with several parametric models. An $L$-dependent model with fix lag $L$ corresponds to $\alpha = 0$ and $\omega_{m,h} = 0$ for all $h > L$. This includes the IID and uncorrelated noise as special cases. One may also think of an MA(L) noise such that $L = Cm^{\delta}$ for some constants $C > 0$ and $\delta \geq 0$. The case $\delta = \alpha = 0$ brings us back to the MA(L) model with constant lag $L$ whilst $\delta \in (0, 1)$ describes a situation where a higher market activity generates a noise with longer dependence lag.[5] Assumption E2 also accommodates an AR(1) models. Indeed, assume that $\varepsilon_{t,j}$ satisfies $\omega_{m,h} = \omega_0 \left(\phi_m\right)^h, h \geq 0$ and $\phi_m = \frac{\omega_{m,1}}{\omega_0}$. This model fits into E2(b) if $\omega_{m,h} - \omega_{m,h+1} = \omega_0 \left(\phi_m\right)^h (1 - \phi_m) = O(m^{-\alpha})$, for all $h > 1$. Hence, $\alpha = 0$ accommodates an AR(1) with constant autoregressive root whilst $\alpha \in (0, 1)$ implies that $\phi_m$ converges either to zero or to one as $m \to \infty$.[6]

Finally, Assumption E3 replicates Assumption (1c) and Assumption 2 of Ubukata and Oya (2009) and it is needed for the central limit theorem of Politis, Romano and Wolf (1997) to hold. This assumption is satisfied if the squared return process $r_{t,j}^2$ is stationary and strong-mixing.

# 2 Properties of Common Realized Measures

In this section, we study the traditional realized variance, the kernel-based estimator of Hansen and Lunde (2006) and the realized kernels of Barndorff-Nielsen et al. (2008a) under our microstructure noise model.

## 2.1 The Realized Volatility

Under an IID noise, $RV_t^{(m)}$ is biased and inconsistent for $IV_t$ and its bias and variance increase linearly in $m$, see e.g. Hansen and Lunde (2006). Here, we consider the sparsely sampled realized variance given by:

$$RV_t^{(m_q)} = \sum_{k=1}^{m_q} \widetilde{r}_{t,k}^2, \;\; k = 1, ..., m_q,$$

with $m_q = \frac{m}{q}, q \geq 1$ and $\widetilde{r}_{t,k} = \sum_{j=qk-q+1}^{qk} r_{t,j}$ being the sum of $q$ consecutive returns.[7] Note that Equation (6) implies that $\widetilde{r}_{t,k} = \widetilde{r}_{(1),t,k} + \widetilde{r}_{(2),t,k}$ where $\widetilde{r}_{(1),t,k} = \sum_{j=qk-q+1}^{qk} r^*_{(1),t,j}$ and:

$$\begin{aligned}\widetilde{r}_{(2),t,k} &= \left(1 + \beta_0 + \frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right) r^*_{(2),t,qk} + \sum_{j=qk-q+1}^{qk-1} r^*_{(2),t,j} \\ &\quad - \left(\beta_0 + \frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk-q}}\right) r^*_{(2),t,qk-q} + \left(\varepsilon_{t,qk} - \varepsilon_{t,qk-q}\right),\end{aligned}$$

with the convention that $\sum_{j=qk-q+1}^{qk-1} r^*_{(2),t,j} = 0$ when $q = 1$. The next theorem gives the bias and variance of $RV_t^{(m_q)}$.

**Theorem 1** *Assume that the noise process evolves according to equation (10). Then we have:*

$$\begin{aligned} E\left(RV_t^{(m_q)}\right) &= IV_t + \underbrace{2m_q\left(\omega_0 - \omega_{m,q}\right)}_{\textit{bias due to exogenous noise}} \\ &\quad \underbrace{+2\left(1-\rho^2\right)\left(\frac{\beta_1^2}{q} + \frac{\beta_1\left(2\beta_0+1\right)}{\sqrt{m}}\sum_{k=1}^{m_q}\sigma^*_{t,qk} + \beta_0\left(\beta_0+1\right)\sum_{k=1}^{m_q}\sigma^{*2}_{t,qk}\right)}_{\textit{bias due to endogenous noise}} \\ &\quad \underbrace{+\left(1-\rho^2\right)\left(\beta_0^2\left(\sigma^{*2}_{t,0} - \sigma^{*2}_{t,m}\right) + \frac{2\beta_0\beta_1}{\sqrt{m}}\left(\sigma^*_{t,0} - \sigma^*_{t,m}\right)\right)}_{\textit{end effects}}, \textit{ and} \\ Var\left(RV_t^{(m_q)}\right) &= O(m_q). \end{aligned}$$

The bias of $RV^{(m_q)}$ is comprised of three terms. The dominant term, $2m_q\left(\omega_0 - \omega_{m,q}\right)$, is due to the exogenous noise. According to Assumption E2(b), this term is $O(m^{1-\alpha})$ and hence. The second term of the bias is due to the endogenous noise and it is $O(1)$. The latter term does not diverge as $m \to \infty$ and hence, a volatility signature plot may not be able to detect its presence. The third term of the bias is $O(m^{-1})$ and it is due to end effects.

Gloter and Jacod (2001) considered an exogenous noise whose variance depends on the sampling frequency and they show that this noise is irrelevant if $m$ times the noise variance is bounded for all $m$. Our endogenous noise satisfies this condition as it is $O_p(m^{-1/2})$. However, because it is endogenous, it causes a bias term of magnitude $O(1)$.

### 2.2 The Estimator of Hansen and Lunde

Hansen and Lunde (2006) proposed the following flat kernel estimator:

$$RV_t^{(AC,m,L+1)} = \sum_{j=1}^{m} r_{t,j}^2 + \sum_{h=1}^{L+1} \sum_{j=1}^{m} r_{t,j} \left( r_{t,j+h} + r_{t,j-h} \right), \tag{11}$$

where $L$ is the dependence lag of the noise. When $L = 0$ so that $\varepsilon_{t,j}$ is IID, $RV_t^{(AC,m,L+1)}$ coincides with the estimator of French and al. (1987) and Zhou (1996):

$$RV_t^{(AC,m,1)} = \sum_{j=1}^{m} r_{t,j}^2 + 2 \sum_{j=1}^{m} r_{t,j} r_{t,j-1} + \underbrace{\left( r_{t,m+1} r_{t,m} - r_{t,1} r_{t,0} \right)}_{\text{end effects}}. \tag{12}$$

The estimator $RV_t^{(AC,m,L+1)}$ is unbiased for $IV_t$ under a general MA(L) noise. However, it is biased if the exogenous noise is AR(1). An unbiased estimator under AR(1) exogenous noise is given by:

$$RV_t^{(AC,m,\infty)} = \sum_{j=1}^{m} r_{t,j}^2 + \sum_{j=1}^{m} r_{t,j} \left( r_{t,j+1} + r_{t,j-1} \right) + \frac{1}{1 - \phi_m} \sum_{j=1}^{m} r_{t,j} \left( r_{t,j+2} + r_{t,j-2} \right) \tag{13}$$

where $\phi_m$ is the autoregressive root of the noise.[8] In order to gain some insights on the properties of the estimators above, we specialize the exogenous noise to the IID case and derive the mean and variance of $RV_t^{(AC,m,1)}$.

**Theorem 2** *Assume that the noise process evolves according to Equation (10). If $\varepsilon_{t,j}$ is IID, then we have:*

$E\left(RV_t^{(AC,m,1)}\right) = IV_t + \left(1 - \rho^2\right) \left[ \left(\beta_0^2 + 2\beta_0\right) \left(\sigma_{t,m}^{*2} - \sigma_{t,0}^{*2}\right) - \frac{2\beta_1(1+\beta_0)}{\sqrt{m}} \left(\sigma_{t,m}^{*} - \sigma_{t,0}^{*}\right) \right]$, *and*
$Var\left(RV_t^{(AC,m,1)}\right) = O\left(m\right).$

When the endogenous noise is absent and the exogenous noise is IID, Theorem 2 yields a well-known result derived by Hansen and Lunde (2006, Lemma 3). In this particular case, $RV_t^{(AC,m,1)}$ is unbiased for $IV_t$ while its variance increases linearly in $m$ and consequently, $RV_t^{(AC,m,L+1)}$ and $RV_t^{(AC,m,\infty)}$ are not consistent for $IV_t$.

When the exogenous noise is absent so that the noise is purely endogenous, it can be shown that the bias and variance of $RV_t^{(AC,m,1)}$ are both $O\left(m^{-1}\right)$. This result is obtained by specializing the formulas derived in the proof of Theorem 2 to the "*no leverage and no exogenous noise*" case. Hence, $RV_t^{(AC,m,1)}$ is consistent for $IV_t$ in the presence of the endogenous noise. This property extends to $RV_t^{(AC,m,L+1)}$ since $RV_t^{(AC,m,L+1)} - RV_t^{(AC,m,1)}$ converges to zero in the absence of exogenous noise. The same can be said for $RV_t^{(AC,m,\infty)}$ as long as an estimator of $\phi_m$ that converges sufficiently fast is available.

The realized kernel considered in the next section is a version of $RV_t^{(AC,m,L+1)}$ where the higher order covariance terms are weighted by a kernel function. Hence, this estimator is also robust to endogenous noise and leverage effect. Acting on this, we study the realized kernel below by assuming that $\beta_0 = \beta_1 = \rho = 0$.

### 2.3 The Realized Kernels

Barndorff-Nielsen et al. (2008a) proposed the following estimator for $IV_t$ which they named "*realized kernel*":

$$K_t^{BNHLS} = \gamma_{t,0}(r) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right), \tag{14}$$

for a positive kernel function $k(.)$ such that $k(0) = 1$ and $k(1) = 0$, where $\gamma_{t,h}(x) = \sum_{j=1}^{m} x_{t,j}x_{t,j-h}$ for any variable $x$. Equation (14) is reminiscent of the long run variance estimators of Newey and West (1987) and Andrews and Monahan (1992).

Barndorff-Nielsen et al. (2008a, *Proposition 3*) show that $K_t^{BNHLS}$ is consistent for $IV_t$ for some choices of bandwidth if the covariance between $\varepsilon_{t,h}$ and $\varepsilon_{t,h-j}$ vanishes for any $h$ and $j$ as $m \to \infty$. This condition is satisfied for an AR(1) noise with $\phi_m \to 0$ as $m \to \infty$ or for an MA(L) noise with constant lag $L$. It is not satisfied for an AR(1) noise with $\phi_m \to 1$ as $m \to \infty$ nor for an MA(L) noise with $L = Cm^\delta$ and $\delta > 0$. For the latter MA(L) noise, Assumption E2(b) implies that:

$$\omega_{m,j} - \omega_0 = -\sum_{h=0}^{j-1}(\omega_{m,h} - \omega_{m,h+1}) = O(jm^{-\alpha}).$$

Hence, $\omega_{m,j} - \omega_0 = O(m^{-\alpha})$ and $\lim\limits_{m\to\infty} \frac{\omega_{m,j}}{\omega_0} = 1$ for any fixed $j$.

Further, Barndorff-Nielsen et al. (2008a, *Proposition 4*) show that $K_t^{BNHLS}$ is consistent for

$IV_t$ when $\varepsilon_{t,j}$ is AR(1) with constant autoregressive root. Below, we study the estimator given by:

$$K_{t,Lead}^{BNHLS} = \gamma_{t,0}(r) + 2\sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\gamma_{s,h}(r), \tag{15}$$

when the first order autocorrelation of the noise converges to one as $m \to \infty$ and we use the results to infer some properties of $K_t^{BNHLS}$. Note that $K_{t,Lead}^{BNHLS}$ can be decomposed as:

$$K_{t,Lead}^{BNHLS} = K_t(r^*) + K_t(r^*, \Delta u) + K_t(\Delta u, r^*) + K_t(\Delta u),$$

where

$$\begin{aligned} K_t(x) &= \gamma_{t,0}(x) + 2\sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\gamma_{t,h}(x), \\ K_t(x,y) &= \gamma_{t,0}(x,y) + 2\sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\gamma_{t,h}(x,y), \end{aligned}$$

and $\gamma_{t,h}(x,y) = \sum_{j=1}^{m} x_{t,j} y_{t,j-h}$. We have the following consistency result under IID exogenous noise.

**Theorem 3** *Assume that $\beta_0 = \beta_1 = \rho = 0$, $k(x) = 1 - x$ (the Bartlett kernel) and $\varepsilon_{t,j}$ is IID. Then, we have:*

$$K_{t,Lead}^{BNHLS} - IV_t = -\varepsilon_{t,0}^2 + \varepsilon_{t,m}^2 + O_p(m^{-1/6}).$$

*as $m \to \infty$ and $H$ is proportional to $m^{2/3}$.*

Barndorff-Nielsen et al. (2008a, Theorem 4) gives the same rate of convergence for $K_t^{BNHLS}$ under the conditions of Theorem 3. Here, Theorem 3 focuses on $K_{t,Lead}^{BNHLS}$. Note that $K_t^{BNHLS} = \frac{1}{2}\left(K_{t,Lead}^{BNHLS} + K_{t,Lag}^{BNHLS}\right)$ where $K_{t,Lag}^{BNHLS}$ is the twin estimator given by:

$$K_{t,Lag}^{BNHLS} = \gamma_{t,0}(r) + 2\sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\gamma_{s,-h}(r)$$

Hence, $Var\left(K_{t,Lead}^{BNHLS}\right)$ is always larger than $Var\left(K_t^{BNHLS}\right)$ although both estimators enjoy the same rate of convergence.

When the noise is not IID but remains purely exogenous, $K_t^{BNHLS}$ and $K_{t,Lead}^{BNHLS}$ have the same expectation as $K_t(\Delta u)$. We have the following results.

**Theorem 4** *Assume that $\beta_0 = \beta_1 = \rho = 0$, $k(x) = 1 - x$ (the Bartlett kernel) and $\varepsilon_{t,j}$ and that Assumptions E1 and E2 hold.*

*(i) If the noise is AR(1) with autoregressive root $\phi_m$, then we have:*

$$E(K_t(\Delta u)) = -2mH^{-1}(\phi_m)^H(2 + \phi_m)\omega_0.$$

*If further $\phi_m \simeq 1 - Dm^{-\alpha}$ $(\alpha \geq 0)$ so that $\phi_m \to 1$ as $m \to \infty$ and $H = Cm^{\gamma}$ $(\gamma \in (0,1))$ then we have:*

$$|E(K_t(\Delta u))| \simeq 2(2 + \phi_m)\omega_0 C m^{1-\gamma} \exp\left(-CDm^{\gamma-\alpha}\right)$$

*(ii) If the noise is MA(L), $E(K_t(\Delta u)) = 0$ as long as $H \geq L + 1$.*

The first result of Theorem 4 stipulates that under an AR(1) exogenous noise with $\phi_m \to 1$ as $m \to \infty$, $E(K_t(\Delta u))$ converges to zero if and only if $H = Cm^{\gamma}$ with $\gamma > \alpha$. Otherwise, the bias diverges to infinity. Hence, a necessary condition for the MSE of $K_t^{BNHLS}$ to be finite is that $H$ diverges to infinity sufficiently fast as $m \to \infty$. Under MA(L) noise, $E(K_t(\Delta u)) = 0$ if $H$ is greater than the dependence lag of the noise. Again, if $L = Cm^{\delta}$ as assumed, a sufficient condition for $K_t^{BNHLS}$ to be unbiased is that $H$ diverges to infinity faster than $L$ as $m \to \infty$.

# 3 Adaptive Realized Kernels

The results of a simulation study performed by Gatheral and Oomen (2007) suggests that inconsistent estimators like $RV_t^{(AC,m,1)}$ often outperform some theoretically consistent estimators like $K_t^{BNHLS}$ at record frequencies commonly encountered in practice (e.g. one to five minutes). From a theoretical point of view, one can think of at least three situations where the MSE of $RV_t^{(AC,m,1)}$ can be lower than that of $K_t^{BNHLS}$. The first situation is the one in which the variance of the microstructure noise is so small that it contributes very little to the MSE. The second situation may happen because the bandwidth $H$ is not optimally selected for $K_t^{BNHLS}$. The third situation corresponds to the case where the sampling frequency is not large enough to make the asymptotic results for $K_t^{BNHLS}$ useful. All three situations are related to the fact that the performance of an IV estimator at a given sampling frequency reflects the trade-off between the discretization error and the microstructure noise at that frequency. Indeed, $RV_t^{(AC,m,1)}$ is exempted of the bias of its ancestor $RV_t^{(m)}$ at the expense of a higher discretization error (i.e. the MSE in the absence of noise). Also, $K_t^{BNHLS}$ brings consistency upon conceding a higher discretization error than $RV_t^{(AC,m,1)}$. Given that the discretization error $K_t^{BNHLS}$ increases with the bandwidth $H$, the optimal selection of $H$ involves a trade-off between the MSE due to discretization and the MSE due to the microstructure noise. Below, we propose an adaptive estimator that is aimed at achieving this optimal trade-off.

Consider $N$ kernel-based estimators of $IV_t$ given by:

$$\widehat{IV}_t^{(i)} = \gamma_{t,0}(r) + \sum_{h=1}^{H} k_i\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right), i = 1, 2, ..., N, \tag{16}$$

where $k_i(.), i = 1, ..., N$ are distinct kernel functions. Alternatively, one may consider using the same kernel function but different bandwidths, as in the following example:

$$\widehat{IV}_t^{(i)} = \gamma_{t,0}(r) + \sum_{h=1}^{H_i} k\left(\frac{h-1}{H_i}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right), i = 1, 2, ..., N.$$

By letting $H = \max_{1\leq i\leq N} H_1$, the latter equation may be re-written as (16) with $k_i(x) \equiv k\left(\frac{H}{H_i}x\right), 0 \leq x \leq 1$ and $k_i(x) = 0$ otherwise.

We consider selecting the estimator with smallest MSE within the class defined by:

$$K_t^{\varpi} = \sum_{i=1}^{N} \varpi_i \widehat{IV}_t^{(i)} \text{ subject to } \sum_{i=1}^{N} \varpi_i = 1,$$

where $\varpi = (\varpi_1, ..., \varpi_N)'$ is a vector of weights. Note that $K_t^{\varpi}$ is also a realized kernel, as we have:

$$K_t^{\varpi} = \gamma_{t,0}(r) + \sum_{h=1}^{H} k_{\varpi}\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right), \tag{17}$$

with $k_{\varpi}(x) = \sum_{i=1}^{N} \varpi_i k_i(x)$.

To illustrate the idea, suppose that the exogenous noise is $L$-dependent. Then, we may define:

$$K_t^{\varpi} = (1 - \varpi)\, \widehat{IV}_t^{(1)} + \varpi \widehat{IV}_t^{(2)},$$

where $\widehat{IV}_t^{(1)}$ is $K_t^{BNHLS}$ at bandwidth $L$ and $\widehat{IV}_t^{(2)}$ is the same estimator at bandwidth $H$. We have:

$$K_t^{\varpi} = \gamma_{t,0}(r) + \sum_{h=1}^{L+1} k\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right) + \varpi \sum_{h=L+2}^{H} k\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right)$$

We see that $K_t^{\varpi}$ exploits the $L$-dependence of the noise by discounting the kernel windows assigned to the covariance terms beyond lag $L+1$. The optimal weight $\varpi^*$ that minimizes the MSE of $K_t^{\varpi}$ mitigates the impact of the discretization error induces by the higher order covariance terms while guaranteeing that $K_t^{\varpi^*}$ inherits the consistency of $K_t^{BNHLS}$. The standard realized kernel includes the covariance terms of higher displacements in order to control the variance, but it does not exploit the life of a dependent noise. A theoretical importance of the estimator $K_t^{\varpi^*}$ resides in that it introduces an extra degree of freedom ($\varpi$) besides the bandwidth parameter ($H$) and hence, it provides an adaptive approach for tuning realized kernel. Subsequently, we refer to $K_t^{\varpi^*}$ as the "*adaptive realized kernels*". Note that $K_t^{\varpi^*}$ has the flavor of a model averaging estimator (see Hansen, 2007) and it shares some similarities with the estimator proposed in Ghysels, Mykland and Renault (2008).

Let $\widehat{V}_t = \left(\widehat{IV}_t^{(1)}, ..., \widehat{IV}_t^{(N)}\right)'$ so that $K_t^{\varpi} = \varpi' \widehat{V}_t$ with $\varpi' \iota = 1$, where $\iota$ is a vector of ones. The unconditional MSE of $K_t^{\varpi}$ is $E\left(K_t^{\varpi} - IV_t\right)^2 = \varpi' \Omega \varpi$, where $\Omega = E\left[\left(\widehat{V}_t - IV_t \iota\right)\left(\widehat{V}_t - IV_t \iota\right)'\right]$ is the MSE matrix of the vector $\widehat{V}_t$. The optimal vector of weights is given by:

$$\varpi^* = \left(\iota' \Omega^{-1} \iota\right)^{-1} \Omega^{-1} \iota \qquad (18)$$

A feasible vector of weights is obtained by plugging an empirical counterpart of $\Omega$ into (18), as illustrated in Section 5.2. By construction, the MSE of $K_t^{\varpi^*}$ is necessarily smaller than the MSE of each of the initial estimators $\widehat{IV}_t^{(i)}, i = 1, ..., N$.

# 4 Inference on the Microstructure Noise Parameters

In order to implement the realized kernels efficiently, one needs to know whether the noise has finite dependence lag or infinite dependence lag. In this section, we consider estimating the correlogram of the noise by assuming that the noise is either AR(1) or MA(L) at a given record frequency $m$.[9] The AR(1) assumption targets noises with infinite dependence lag while the MA(L) assumption provides a reasonable approximation if the noise has finite dependence.

From Theorem 2, we can infer that:

$$\begin{aligned} E\left[\gamma_{t,1}\right] &= -\left(1-\rho^2\right)\sum_{j=1}^{m}\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,j-1}}\right)\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,j-1}}\right)\sigma^{*2}_{t,j-1} \\ &\quad +m\left(-\omega_0+2\omega_{m,1}-\omega_{m,2}\right), \end{aligned}$$

where $\gamma_{t,h}$ is used as shorthand for $\gamma_{t,h}(r)$. Let $b_t^{(m)} = E\left[RV_t^{(m)} - IV_t\right]$ denote the bias of the realized volatility computed at the record frequency. When $q=1$, it follows from Lemma 5 in appendix that:

$$\begin{aligned} b_t^{(m)} &= 2\left(1-\rho^2\right)\sum_{j=1}^{m}\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,j-1}}\right)\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,j-1}}\right)\sigma^{*2}_{t,j-1} \\ &\quad +2m\left(\omega_0-\omega_{m,1}\right)+\left(1-\rho^2\right)\left(\beta_0^2\left(\sigma^{*2}_{t,0}-\sigma^{*2}_{t,m}\right)+\frac{2\beta_0\beta_1}{\sqrt{m}}\left(\sigma^*_{t,0}-\sigma^*_{t,m}\right)\right). \end{aligned}$$

Hence, the following unconditional moment conditions hold:

$$E\left(RV_t^{(m)}\right) = IV_t + b_t^{(m)}, \tag{19}$$

$$E\left(\gamma_{t,1}+\gamma_{t,-1}\right) = -b_t^{(m)} + 2m\left(\omega_{m,1}-\omega_{m,2}\right) \text{ and} \tag{20}$$

$$E\left(\gamma_{t,h+1}+\gamma_{t,-h-1}\right) = -2m\left(\omega_{m,h}-2\omega_{m,h+1}+\omega_{m,h+2}\right),\ h\geq 1, \tag{21}$$

Below, we consider the AR(1) and MA(L) cases separately.[10]

## 4.1 Inference with an AR(1) Microstructure Noise

Under an AR(1) model, the noise autocovariances satisfy $\omega_{m,h} = \omega_0 (\phi_m)^h$ and thus, Equation (21) implies that $E(\widehat{g}_h(\omega_0, \phi_m)) = 0$ with:

$$\widehat{g}_h(\omega_0, \phi_m) = \frac{1}{2mT} \sum_{t=1}^{T} \left(\gamma_{t,h+1} + \gamma_{t,-h-1}\right) + \omega_0 (1 - \phi_m)^2 (\phi_m)^h, \ h \geq 1 \tag{22}$$

Let $\widehat{g} = (\widehat{g}_1, ..., \widehat{g}_n)$ be a vector of $n$ selected moments conditions, with $\widehat{g}_h \equiv \widehat{g}_h(\omega_0, \phi_m)$. The GMM estimators of $(\omega_0, \phi_m)$ are given by:

$$\begin{pmatrix} \widehat{\omega}_0 \\ \widehat{\phi}_m \end{pmatrix} = \arg\min \widehat{g}' \widehat{S}^{-1} \widehat{g},$$

where $\widehat{S}$ is a consistent first step estimator of the long run covariance matrix of the moment conditions, that is, $\widehat{S} = \lim_{T \to \infty} Var\left(\sqrt{T}\widehat{g}\right)$.

After estimation, the overidentification test of Hansen (1982) may be used to check whether the AR(1) model fits the data reasonably well. This test is based on the following asymptotic distribution under the null hypothesis that the AR(1) model is true:

$$J = T\widehat{g}' \widehat{S}^{-1} \widehat{g} \to \chi^2 (N-2) \text{ as } T \to \infty, \tag{23}$$

After performing this test, and if the null hypothesis is not rejected, we may then perform a standard t-test for the significance of the parameters $(\omega_0, \phi_m)$. The distribution of the estimators under the null hypothesis is:

$$\sqrt{T} \begin{pmatrix} \widehat{\omega}_0 - \omega_0 \\ \widehat{\phi}_m - \phi_m \end{pmatrix} \to N\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \left(G' \widehat{S}^{-1} G\right)^{-1} \right). \tag{24}$$

where $G$ is the $(n \times 2)$ Jacobian matrix of the moment conditions. The $h^{th}$ row of $G$ is given by:

$$\begin{aligned} G_h &= \left( \frac{\partial \widehat{g}_h}{\partial \omega_0}, \frac{\partial \widehat{g}_h}{\partial \phi_m} \right) \\ &= \left( (1-\phi_m)^2 (\phi_m)^h, \ \ -2\omega_0 (1-\phi_m)(\phi_m)^h + h\omega_0 (1-\phi_m)^2 (\phi_m)^{h-1} \right) \end{aligned}$$

Note that $G$ is a deterministic matrix.

### 4.2 Inference with an MA(L) Microstructure Noise

Under the MA(L) model, the noise autocovariances satisfy $\omega_{m,h} = 0$ for $h > L$. Thus, Equations (19)-(21) provide $L + 2T$ moment conditions that can be used to estimate $L + 2T$ parameters, namely $\left\{ b_t^{(m)}, IV_t \right\}_{t=1}^{T}$ and $\{\omega_{m,h}\}_{h=1}^{L}$. Estimating these parameters by the method of moments is straightforward. First solving for $\omega_{m,L}$ and then proceeding by backward substitution into (21) to (19) yields the following unbiased estimators for $\omega_{m,h}$, $b_t^{(m)}$ and $IV_t$ respectively:

$$
\begin{aligned}
\widehat{\omega}_{m,h} &= -\frac{1}{2Tm} \sum_{s=1}^{T} \sum_{l=1}^{L-h+1} l \left( \gamma_{s,h+l} + \gamma_{s,-h-l} \right), \; h = 1, ...L, && (25) \\
\widehat{b}_t^{(m)} &= -\gamma_{t,1} - \gamma_{t,-1} - \frac{1}{T} \sum_{s=1}^{T} \sum_{l=2}^{L+1} \left( \gamma_{s,l} + \gamma_{s,-l} \right) \text{ and} && (26) \\
\overline{RV}_t^{(AC,m,L+1)} &= RV_t^{(m)} - \widehat{b}_t^{(m)} && (27) \\
&= \gamma_{t,0} + \gamma_{t,1} + \gamma_{t,-1} + \frac{1}{T} \sum_{s=1}^{T} \sum_{l=2}^{L+1} \left( \gamma_{s,l} + \gamma_{s,-l} \right).
\end{aligned}
$$

Hence, $\overline{RV}_t^{(AC,m,L+1)}$ is an unbiased method-of-moment estimator of $IV_t$. Note that $\overline{RV}_t^{(AC,m,L+1)}$ specializes to the estimator of Hansen and Lunde when $T = 1$.

To estimate the noise variance $\omega_0$, we use the expression of the bias of the RV sampled at the highest frequency. We have:

$$
\widehat{\omega}_0 = \frac{1}{2mT} \sum_{t=1}^{T} \widehat{b}_t^{(m)} + \widehat{\omega}_{m,1} \tag{28}
$$

All the noise autocovariance estimates can be written as:

$$
\widehat{\omega}_{m,h} = \frac{1}{mT} \sum_{t=1}^{T} \sum_{j=1}^{m} \widehat{\omega}_{t,j,h}, h = 0, 1, ..., L \tag{29}
$$

where $\widehat{\omega}_{t,j,h}$, $h = 0, ..., L$ are defined as follows:

$$
\begin{aligned}
\widehat{\omega}_{t,j,0} &= -\frac{1}{2} \sum_{h=1}^{L+1} \left( \gamma_{t,j,h} + \gamma_{t,j,-h} \right) + \left( P^{-1} \gamma_{t,j,(2,L+1)} \right)_1, \\
\left( \widehat{\omega}_{t,j,1}, ..., \widehat{\omega}_{t,j,L} \right)' &= P^{-1} \gamma_{t,j,(2,L+1)} \text{ and} \\
\gamma_{t,j,(2,L+1)} &= \left( \gamma_{t,j,2}, ..., \gamma_{t,j,L+1} \right)'.
\end{aligned}
$$

with $\gamma_{t,j,h} = \frac{1}{2} r_{t,j} \left( r_{t,j-h} + r_{t,j+h} \right)$ for all $t$ and $h$, $P$ being the $L \times L$ matrix with elements: $P_{i,i} = -1$, $P_{i,i+1} = 2$, $P_{i,i+2} = -1$, $P_{i,j} = 0$ otherwise $1 \leq i, j \leq L$, and $\left( P^{-1} \overline{\gamma}_{t,(2,L+1)} \right)_1$ being the first element of the vector $P^{-1} \overline{\gamma}_{t,(2,L+1)}$.

Based on Equation (29), we consider the subsampled variance $\widehat{Q}_h$ given by:

$$\widehat{Q}_h = \frac{m}{T} \sum_{t=1}^{T} \left( \frac{1}{m} \sum_{j=1}^{m} \widehat{\omega}_{t,j,h} - \widehat{\omega}_{m,h} \right)^2 . \tag{30}$$

Under Assumptions E1, E2 and E3, we have:

$$\frac{(mT)^{1/2} \left( \widehat{\omega}_{m,h} - \omega_{m,h} \right)}{\sqrt{\widehat{Q}_h}} \rightarrow N(0,1) \tag{31}$$

as $T$ goes to infinity and $m$ is fixed. See Ubukata and Oya (2009) for the proof.[11]

The knowledge of $L$ is required to estimate the correlogram of the microstructure noise. A simple way to estimate $L$ is to perform significance tests for $\omega_{m,h}$ by using autocovariance estimates that rely on an initial guess $L_{\max}$.[12] Under the null hypothesis that $\omega_{m,h} = 0$, we have:

$$\widehat{\tau}_h = \frac{(mT)^{1/2} \widehat{\omega}_{m,h}}{\sqrt{\widehat{Q}_h}} \rightarrow N(0,1) \tag{32}$$

The statistics $\widehat{\tau}_h$ diverges under the alternative. The estimator $\widehat{L}$ is the maximum lag at which the null is rejected. Provided that the initial guess $L_{\max}$ exceeds the true value of $L$, the estimator $\hat{L}$ will not underestimate the true $L$ asymptotically.

## 5 Monte Carlo Simulations

The simulation study is organized as follows. First, we apply the AR(1) noise inference procedure to a correctly specified model. Second, we verify the power of this procedure by applying it to an MA(3) noise. Third, we study the performance of the MA(L) noise inference procedure when the model is correctly specified. Finally, we assess the quality of the IV estimators under either type of noise.

## 5.1 The Data Generating Processes

We assumed that the efficient log-price process evolves according to the model of Heston (1993):

$$dp_t^* = \sigma_t dW_{1,t} \text{ and} \tag{33}$$

$$d\sigma_t^2 = \kappa\left(\alpha - \sigma_t^2\right)dt + \gamma\sigma_t\left(\rho dW_{1,t} + \sqrt{1-\rho^2}dW_{2,t}\right), \tag{34}$$

where $W_{1,t}$ and $W_{2,t}$ are independent Brownian motions and the parameter $\rho$ captures the leverage effect. Following Zhang and al. (2005), we set the annualized parameters values as follows:

$$\kappa = 5, \alpha = 0.04, \gamma = 0.5, \rho \in \{0, -0.5\},$$

Using the Poisson-Mixing-Gamma characterization of Devroye (1986) for the spot volatility process (34), we simulate the efficient price data at five seconds[13] but we assume that the record frequency is one minute.

To start with, we simulate once and for all a sample of $T = 500$ days of efficient price data. Next, we contaminate iteratively this sample with a microstructure noise that is simulated according to:

$$u_{t,j} = \left(\beta_0 + \frac{\beta_1}{\sqrt{m}\sigma_{t,j}^*}\right) r_{t,j}^* + \varepsilon_{t,j},\ j = 1, ..., m,$$

where $\beta_0 = 0.5$, $\beta_1 = 0.5$ and the exogenous noise $\varepsilon_{t,j}$ is either an AR(1) or an MA(3). For the AR(1) exogenous noise, we use $\varepsilon_{t,j} = \phi_m \varepsilon_{t,j-1} + v_{t,j}$, with $v_{t,j} \overset{IID}{\sim} N(0,\ \alpha_0)$, $\phi_m \in \{-0.9, 0, 0.9\}$ and $\alpha_0$ varying so as to match $\omega_0 = \frac{\alpha_0}{1-(\phi_m)^2}$ with the following values:[14]

$$\omega_0 \in \left\{2.5 \times 10^{-7}, 2.25 \times 10^{-6}, 2.5 \times 10^{-5}\right\}. \tag{35}$$

The variance $\omega_0 = 2.5 \times 10^{-7}$ has been used in Zhang and al. (2005) at five minute sampling frequency while $\omega_0 = 2.25 \times 10^{-6}$ has served in Ait-Sahalia and al. (2005) at frequencies ranging from one to thirty minutes.

For the MA(1) exogenous noise, we use $\varepsilon_{t,j} = v_{t,j} + \alpha_1 v_{t,j-1} + \alpha_2 v_{t,j-2} + \alpha_3 v_{t,j-3}$, with $v_{t,j} \overset{IID}{\sim}$

$N(0,\,\alpha_0)$, $\alpha_1 = 0.5$, $\alpha_2 = 0.2$ and $\alpha_3 = 0.05$. This implies:

$$\begin{aligned}
\omega_0 &\equiv E\left(\varepsilon_{t,j}^2\right) = \alpha_0\left(1+\alpha_1^2+\alpha_2^2+\alpha_3^2\right) = 1.2925\alpha_0, \\
\omega_{m,1} &\equiv E\left(\varepsilon_{t,j}\varepsilon_{t,j-1}\right) = \alpha_0\left(\alpha_1+\alpha_1\alpha_2+\alpha_2\alpha_3\right) = 0.61\alpha_0, \\
\omega_{m,2} &\equiv E\left(\varepsilon_{t,j}\varepsilon_{t,j-2}\right) = \alpha_0\left(\alpha_2+\alpha_1\alpha_3\right) = 0.225\alpha_0, \\
\omega_{m,3} &\equiv E\left(\varepsilon_{t,j}\varepsilon_{t,j-2}\right) = \alpha_0\alpha_3 = 0.05\alpha_0 \text{ and} \\
\omega_{m,h} &\equiv E\left(\varepsilon_{t,j}\varepsilon_{t,j-h}\right) = 0 \text{ for all } h \geq 4,
\end{aligned}$$

where $\alpha_0$ varies so as to match the variances in (35).

### 5.2 Simulation Results

Table 1 presents the estimation results for a correctly specified AR(1) noise model. The simulations are performed with and without the leverage effect.

Table 1: Estimation of a well-specified AR(1) noise model by GMM

$\phi_m = 0.9$, $m = 390$, 1000 replications

| | no leverage effect | | | | with leverage effect | | | |
|---|---|---|---|---|---|---|---|---|
| | T=250 | | T=500 | | T=250 | | T=500 | |
| | $\widehat{\omega}_0$ $\left(\times 10^{-6}\right)$ | $\widehat{\phi}_m$ | $\widehat{\omega}_0$ | $\widehat{\phi}_m$ | $\widehat{\omega}_0$ | $\widehat{\phi}_m$ | $\widehat{\omega}_0$ | $\widehat{\phi}_m$ |
| true | 2.25 | 0.900 | 2.25 | 0.900 | 2.25 | 0.900 | 2.25 | 0.900 |
| mean | 2.20 | 0.894 | 2.15 | 0.895 | 2.03 | 0.887 | 1.85 | 0.883 |
| median | 2.14 | 0.894 | 2.13 | 0.895 | 1.98 | 0.887 | 1.83 | 0.883 |
| std. dev. | 0.43 | 0.014 | 0.28 | 0.009 | 0.36 | 0.014 | 0.21 | 0.009 |
| mean std. dev. | 0.48 | 0.016 | 0.33 | 0.011 | 0.40 | 0.015 | 0.25 | 0.011 |
| rejection rate | 2.70% | | 4.00% | | 1.30% | | 4.00% | |

We see that the estimators $\widehat{\omega}_0$ and $\widehat{\phi}_m$ are slightly biased downward. The bias is more pronounced in the presence of leverage effect and it is more visible for $\widehat{\omega}_0$.[15] The standard deviation (std. dev.) of the empirical distribution of the estimates is quite close to the mean of the standard deviations (mean std. dev.) implied by the analytical formula (24). The last row of the table gives the rate of rejection of the null hypothesis that the true model is AR(1) by the overidentification test at nominal level 5%. Overall, the results suggest that the overidentification test has good size.

In order to assess the power of the previous test, we fit an AR(1) model to an MA(3) microstructure noise. Table 2 presents the results of the simulation. The first order autocorrelation of the MA(3) noise gives us a pseudo-true value for $\phi_m$. In all the scenarios, the noise variance is overestimated while the first order autocorrelation is underestimated. The model rejection rate is nearly 100%, which indicates that the overidentification test has power against MA(L) alternatives. Acting on these results, our preferred strategy for the empirical investigation will consist of first testing the null hypothesis that the noise is AR(1) and next, estimating an MA(L) noise if the AR(1) assumption is rejected.

Table 2: Estimation of a misspecified AR(1) noise model

True model is MA(3), $m = 390$, 1000 replications

| | no leverage | | | | with leverage | | | |
|---|---|---|---|---|---|---|---|---|
| | T=250 | | T=500 | | T=250 | | T=500 | |
| | $\omega_0$ $\left(\times 10^{-6}\right)$ | $\phi_m \equiv \frac{\omega_{m,1}}{\omega_0}$ | $\omega_0$ | $\phi_m$ | $\omega_0$ | $\phi_m$ | $\omega_0$ | $\phi_m$ |
| true | 2.25 | 0.472 | 2.25 | 0.472 | 2.25 | 0.472 | 2.25 | 0.472 |
| mean | 2.79 | 0.425 | 2.78 | 0.425 | 2.81 | 0.421 | 2.79 | 0.420 |
| median | 2.79 | 0.425 | 2.78 | 0.422 | 2.80 | 0.421 | 2.79 | 0.421 |
| std. dev. | 0.05 | 0.011 | 0.03 | 0.007 | 0.05 | 0.011 | 0.03 | 0.008 |
| mean std. dev. | 0.05 | 0.012 | 0.03 | 0.008 | 0.05 | 0.012 | 0.03 | 0.009 |
| rejection rate | 99.7% | | 100% | | 98.8% | | 100% | |

We now study the performance of the inference procedure outlined previously for an MA(L) noise. The first step consists of guessing an initial value $L_{\max}$ that is larger that the true dependence lag $L$. We use an heuristic based on the following empirical MSE:

$$\Delta(l) = \frac{1}{T}\sum_{t=1}^{T}\left(K_t^T - \overline{RV}_t^{(AC,m,l)}\right)^2,\ l = 1, ..., \lfloor 2H/3 \rfloor \tag{36}$$

where $\overline{RV}_t^{(AC,m,l+1)}$ is defined as in (27), $K_t^T = RV_t^{(AC,m,1)} + \frac{1}{T}\sum_{s=1}^{T}\sum_{h=2}^{H}\left(1 - \frac{h-1}{H}\right)\left(\gamma_{s,h} + \gamma_{s,-h}\right)$ and it is implicitly assumed that $H$ is large enough to ensure that $L \leq \lfloor 2H/3 \rfloor$. On the one hand,

$\overline{RV}_t^{(AC,m,l)}$ is obtained by truncating the formula of $\overline{RV}_t^{(AC,m,L+1)}$ to $l$ autocovariance terms and thus, it is thus unbiased for $IV_t$ when $l \geq L+1$. On the other hand, $K_t^T$ is a smoothed version of $\overline{RV}_t^{(AC,m,H)}$ and it is also unbiased for $IV_t$ if the bandwidth $H$ is selected sufficiently large. Hence the mean of $K_t^T - \overline{RV}_t^{(AC,m,l)}$ is decreasing in $l$ as $l$ increases to $L$ and it is equal to zero for $l \geq L+1$. Also, the variance of $K_t^T - \overline{RV}_t^{(AC,m,l)}$ is increasing in $l$. As a result, the curve of $\Delta(l)$ is L-shapped or convex. An initial estimate $\widetilde{L}$ of $L$ is given by the point where the curve $(l, \Delta(l))$ is bent the most or by the minimum of that curve. Figure 1 shows an L-shapped example with an MA(3) noise. Table 3 shows the simulation results for the estimation of $L$. The medians of $\widetilde{L}$ and $\widehat{L}$ coincide with the true value $L = 3$. The corresponding means are slightly biased downward, but this is repaired by rounding up the estimates to the next unit.

Figure 1: Plots of $\Delta(l)$ against $l$. An example with an MA(3) noise.

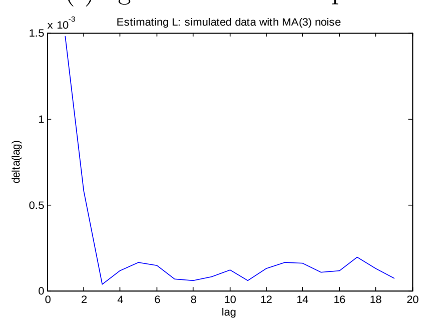



Table 3: Estimation of the dependence lag $L$

True model is MA(3), $m = 390$, $T = 250$, 1000 replications

| | | $\omega_0 = 2.25 \times 10^{-6}$ | | $\omega_0 = 2.5 \times 10^{-5}$ | |
|---|---|---|---|---|---|
| | | no leverage | with leverage | no leverage | with leverage |
| | min | 2.00 | 2.00 | 2.00 | 2.00 |
| $\widetilde{L}$ | mean | 2.97 | 2.91 | 2.99 | 2.99 |
| | median | 3.00 | 3.00 | 3.00 | 3.00 |
| | max | 3.00 | 3.00 | 3.00 | 3.00 |
| | min | 2.00 | 2.00 | 2.00 | 2.00 |
| $\widehat{L}$ | mean | 2.61 | 2.63 | 2.90 | 2.91 |
| | median | 3.00 | 3.00 | 3.00 | 3.00 |
| | max | 3.00 | 3.00 | 3.00 | 3.00 |

Below, we use $L_{\max} = \widetilde{L} + 3$ for the estimation of the correlogram of the noise. Table 4 presents the results. Note that the estimator of $\omega_0$ is expected to be biased upward because it reflect the size of the total noise contaminating the efficient price. Indeed, we have:

$$E\left(\widehat{\omega}_0\right) - \omega_0 = \left(1-\rho^2\right)\left(\frac{\beta_1^2}{m} + \frac{\beta_1\left(2\beta_0+1\right)}{\sqrt{m}} E\left[\sigma^*_{t,qk}\right] + \beta_0\left(\beta_0+1\right)E\left[\sigma^{*2}_{t,qk}\right]\right).$$

The results suggest that the autocovariances $\{\omega_l\}_{l=1}^4$ are estimated without bias. The mean standard deviation (mean std. dev.) is the average of the standard deviations implied by the analytical formula (30). Interestingly, the average of the standard deviations obtained by the analytical formula is close to the empirical standard deviation of the simulated estimates. The last column gives the rate of rejection of the null hypothesis that $\omega_h = 0$. It appears that a standard t-test for the null hypothesis $\omega_4 = 0$ has a good size at 5% nominal level. Also, the separate tests for the null hypotheses $\omega_h = 0$ have power against the alternatives $\omega_h \neq 0, h = 1, 2, 3$.

Table 4: Estimation of the Correlogram of the noise

$m = 390$, $T = 250$, 1000 replications

| | true $(\times 10^{-6})$ | mean $(\times 10^{-6})$ | std. dev. $(\times 10^{-6})$ | mean std. dev. $(\times 10^{-6})$ | Prob(t>1.96) (%) |
|---|---|---|---|---|---|
| | | | no leverage | | |
| $\widehat{\omega}_0$ | 2.250 | 2.573 | 0.050 | 0.051 | 100 |
| $\widehat{\omega}_{m,1}$ | 1.062 | 1.069 | 0.040 | 0.042 | 100 |
| $\widehat{\omega}_{m,2}$ | 0.392 | 0.399 | 0.031 | 0.032 | 100 |
| $\widehat{\omega}_{m,3}$ | 0.087 | 0.093 | 0.021 | 0.021 | 98.5 |
| $\widehat{\omega}_{m,4}$ | 0 | 0.004 | 0.011 | 0.011 | 6.9 |
| | | | with leverage | | |
| $\widehat{\omega}_0$ | 2.250 | 2.566 | 0.048 | 0.052 | 100 |
| $\widehat{\omega}_{m,1}$ | 1.062 | 1.063 | 0.039 | 0.042 | 100 |
| $\widehat{\omega}_{m,2}$ | 0.392 | 0.391 | 0.030 | 0.032 | 100 |
| $\widehat{\omega}_{m,3}$ | 0.087 | 0.090 | 0.020 | 0.022 | 99.5 |
| $\widehat{\omega}_{m,4}$ | 0 | 0.001 | 0.010 | 0.012 | 3.2 |

As a final step of this simulation study, we evaluate the performance of the adaptive realized kernels $K_t^{\varpi^*}$ by simulations. Under either type of noise, we set $K_t^{\varpi} = \sum_{i=1}^{4} \varpi_i \widehat{IV}_t^{(i)}$, with $\widehat{IV}_t^{(1)} = K_{t,15}^{BNHLS}$, $\widehat{IV}_t^{(2)} = K_{t,25}^{BNHLS}$, $\widehat{IV}_t^{(3)} = K_{t,35}^{BNHLS}$ and $\widehat{IV}_t^{(4)} = K_{t,45}^{BNHLS}$, where $K_{t,H}^{BNHLS}$ is the realized Bartlett kernels with bandwidth $H$. Let $\widehat{V}_t = \left(\widehat{IV}_t^{(1)}, \widehat{IV}_t^{(2)}, \widehat{IV}_t^{(3)}, \widehat{IV}_t^{(4)}\right)'$. Under AR(1) noise, the MSE matrix of $\widehat{V}_t$ is $\Omega = Var\left(\widehat{V}_t\right) + BB'$, where $B$ is the 4x1 vector of biases given by:

$$B = -2m\omega_0 \left(2 + \phi_m\right) \left( \frac{(\phi_m)^{15}}{15}, \ \frac{(\phi_m)^{25}}{25}, \ \frac{(\phi_m)^{35}}{35}, \ \frac{(\phi_m)^{45}}{45} \right)' \tag{37}$$

and $Var\left(\widehat{V}_t\right)$ is the covariance matrix of $\widehat{V}_t$. Note that the expression of the bias is deduced from Theorem 4. When the noise is MA(3), the bias vector is $B = (0,0,0,0)'$ and the MSE reduces to $\Omega = Var\left(\widehat{V}_t\right)$. In order to simplify the steps of the Monte Carlo simulation, we assume the ideal situation where $\omega_0$, $\phi_m$ and $L$ are known (in the empirical application, these parameters are replaced by their estimates). We estimate $\Omega$ by replacing $Var\left(\widehat{V}_t\right)$ by its sample counterpart:

$$\widehat{Var}\left(\widehat{V}_t\right) = \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{V}_t - \frac{1}{T}\sum_{l=1}^{T} V_l\right)\left(\widehat{V}_t - \frac{1}{T}\sum_{l=1}^{T} V_l\right)' \tag{38}$$

The MSE of each IV estimator reported in the tables is computed as:

$$MSE(\widehat{IV}_t^{(i)}) = \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{IV}_t^{(i)} - IV_t\right)^2, \ i = 1, ..., 4$$

where $IV_t$ is inferred from the simulated volatility path at one second frequency.

Table 5 shows the simulation results under IID noise, with and without leverage effect. We see that the MSE of all IV estimators are slightly smaller in the presence of leverage effect compared to when there is no leverage. Otherwise, the simulation results are qualitatively similar. When the noise variance is small ($\omega_0 = 25 \times 10^{-8}$), the estimator with smallest bandwidth ($K_{t,15}^{BNHLS}$) has the smallest MSE and it is assigned the largest weight in the design of the adaptive estimator $K_t^{\varpi^*}$. By contrast, when the noise variance is large ($\omega_0 = 2500 \times 10^{-8}$), $K_{t,35}^{BNHLS}$ is the most efficient estimator and it receives the largest or the second largest weight. In either case, the estimator with largest bandwidth ($K_{t,45}^{BNHLS}$) is not efficient because it does not optimally balance the discretization error against the microstructure noise. As expected, the adaptive realized kernel is more efficient than all other estimators taken individually.

Table 5. Assessing the Performance of the Adaptive Realized Kernel by Simulation under IID microstructure noise. $m = 390$, $T = 250$, 1000 Monte Carlo replications.

| | | MSE ($10^{-8}$) | | | Weights | | |
|---|---|---|---|---|---|---|---|
| | $\omega_0$ | $25 \times 10^{-8}$ | $225 \times 10^{-8}$ | $2500 \times 10^{-8}$ | 25 | 225 | 2500 |
| no leverage | $K_{t,15}^{BNHLS}$ | 0.1726 | 0.2044 | 1.8115 | 1.6628 | 1.1249 | 0.0689 |
| | $K_{t,25}^{BNHLS}$ | 0.2504 | 0.2713 | 1.2546 | −0.8205 | −0.0220 | 0.2413 |
| | $K_{t,35}^{BNHLS}$ | 0.3269 | 0.3463 | 1.1713 | 0.3098 | 0.0021 | 0.3265 |
| | $K_{t,45}^{BNHLS}$ | 0.4059 | 0.4263 | 1.2011 | −0.1521 | −0.1049 | 0.3632 |
| | $K_t^{\varpi^*}$ | 0.1568 | 0.1999 | 1.1006 | - | - | - |
| with leverage | $K_{t,15}^{BNHLS}$ | 0.1540 | 0.1858 | 1.7933 | 1.5813 | 1.1387 | 0.0878 |
| | $K_{t,25}^{BNHLS}$ | 0.2372 | 0.2593 | 1.2538 | −0.3598 | 0.1338 | 0.2753 |
| | $K_{t,35}^{BNHLS}$ | 0.3246 | 0.3455 | 1.1865 | −0.3278 | −0.2442 | 0.3293 |
| | $K_{t,45}^{BNHLS}$ | 0.4167 | 0.4384 | 1.2274 | 0.1063 | −0.0284 | 0.3076 |
| | $K_t^{\varpi^*}$ | 0.1383 | 0.1805 | 1.1081 | - | - | - |

Table 6 shows the simulation results under AR(1) microstructure noise and leverage effect. The upper part of the table presents the results for a noise with positive autoregressive root ($\phi_m = 0.9$) while the lower part of the table presents the results for a noise with negative autoregressive root ($\phi_m = -0.9$). The results are qualitatively the same under either type of AR noise. As in the IID noise scenario, the estimator with smallest bandwidth ($K_{t,15}^{BNHLS}$) has the smallest MSE and it is assigned the largest weight when the noise variance is small ($\omega_0 = 25 \times 10^{-8}$). Contrary to the IID noise case, the estimator with largest bandwidth ($K_{t,45}^{BNHLS}$) is the most efficient when the noise variance is large ($\omega_0 = 2500 \times 10^{-8}$). Intuitively, a serially correlated noise causes more harm to IV estimators compared to an IID noise with same variance. The adaptive realized kernel is more efficient than all the individual estimators in the small and large noise variance scenario. The results are nuanced when the noise variance is moderate ($\omega_0 = 225 \times 10^{-8}$).

Table 6. Assessing the Performance of the Adaptive Realized Kernel by Simulation under AR(1) microstructure noise and Leverage Effect. $m = 390$, $T = 250$, 1000 Monte Carlo replications.

| | | MSE ($10^{-8}$) | | | Weights | | |
|---|---|---|---|---|---|---|---|
| | $\omega_0$ | $25 \times 10^{-8}$ | $225 \times 10^{-8}$ | $2500 \times 10^{-8}$ | 25 | 225 | 2500 |
| $\phi_m = 0.9$ | $K_{t,15}^{BNHLS}$ | 0.1697 | 0.9767 | 92.1028 | 1.3314 | $-0.0397$ | $-0.0188$ |
| | $K_{t,25}^{BNHLS}$ | 0.2497 | 0.7023 | 47.4598 | $-0.1182$ | 1.0885 | 0.1082 |
| | $K_{t,35}^{BNHLS}$ | 0.3355 | 0.6356 | 28.2403 | $-0.2539$ | $-0.1691$ | $-0.3075$ |
| | $K_{t,45}^{BNHLS}$ | 0.4256 | 0.6428 | 18.7576 | 0.0407 | 0.1203 | 1.2181 |
| | $K_t^{\varpi^*}$ | 0.1589 | 0.6878 | 17.7389 | - | - | - |
| $\phi_m = -0.9$ | $K_{t,15}^{BNHLS}$ | 0.1782 | 1.8566 | 204.7480 | 1.5368 | $-0.0035$ | $-0.1446$ |
| | $K_{t,25}^{BNHLS}$ | 0.2436 | 0.7142 | 56.6989 | $-0.3094$ | 1.1856 | 0.1024 |
| | $K_{t,35}^{BNHLS}$ | 0.3263 | 0.5368 | 25.7243 | $-0.3184$ | 0.1353 | $-0.1343$ |
| | $K_{t,45}^{BNHLS}$ | 0.4169 | 0.5384 | 14.8193 | 0.0909 | $-0.3174$ | 1.1765 |
| | $K_t^{\varpi^*}$ | 0.1784 | 0.8086 | 6.9812 | - | - | - |

Table 7. Assessing the Performance of the Adaptive Realized Kernel by Simulation under MA(3) microstructure noise. $m = 390$, $T = 250$, 1000 Monte Carlo replications.

| | | MSE ($10^{-8}$) | | | Weights | | |
|---|---|---|---|---|---|---|---|
| | $\omega_0$ | $25 \times 10^{-8}$ | $225 \times 10^{-8}$ | $2500 \times 10^{-8}$ | 25 | 225 | 2500 |
| no leverage | $K_{t,15}^{BNHLS}$ | 0.1770 | 0.5595 | 45.7119 | 1.5982 | 0.9222 | $-0.0180$ |
| | $K_{t,25}^{BNHLS}$ | 0.2520 | 0.4110 | 18.5139 | $-0.7221$ | 0.1207 | 0.1475 |
| | $K_{t,35}^{BNHLS}$ | 0.3281 | 0.4276 | 10.7674 | 0.2706 | 0.0621 | 0.2997 |
| | $K_{t,45}^{BNHLS}$ | 0.4072 | 0.4832 | 7.5090 | $-0.1466$ | $-0.1050$ | 0.5708 |
| | $K_t^{\varpi^*}$ | 0.1644 | 0.5465 | 9.2932 | - | - | - |
| with leverage | $K_{t,15}^{BNHLS}$ | 0.1584 | 0.5404 | 45.7441 | 1.5288 | 0.9370 | $-0.0017$ |
| | $K_{t,25}^{BNHLS}$ | 0.2392 | 0.4026 | 18.5765 | $-0.2982$ | 0.2474 | 0.1654 |
| | $K_{t,35}^{BNHLS}$ | 0.3264 | 0.4306 | 10.8433 | $-0.3240$ | $-0.1137$ | 0.3155 |
| | $K_{t,45}^{BNHLS}$ | 0.4182 | 0.4960 | 7.5582 | 0.0934 | $-0.0707$ | 0.5208 |
| | $K_t^{\varpi^*}$ | 0.1450 | 0.5438 | 9.9502 | - | - | - |

Table 7 shows the simulation results for the MA(3) microstructure noise case. The upper part of the table presents the results for the scenario without leverage effect while the lower part of the table presents the results for the scenario with leverage effect. Qualitatively, the results are similar to what we have seen for the AR(1) noise scenario. Quantitatively, the MSEs of the IV estimators are larger than the MSE under IID noise but smaller than the MSE under AR(1) noise. This suggests that controlling for the noise variance, the more persistent the noise is, the larger the MSE of IV estimators are. This explains why larger bandwidths are needed when the dependence of the noise increases (cf. Theorem 4).

In summary, our empirical investigation strategy is successful in capturing the nature of the dependence of the microstructure noise and thus, it permits to design the adaptive realized kernel in accordance with the properties of the noise.

# 6 Empirical Application

For this application, we use data on twelve stocks listed in the Dow Jones Industrial (see the first column of Table 8). The prices are observed every one minute from January $1^{st}$, 2002 to December $31^{th}$, 2007 (1510 trading days). In a typical trading day, the market opens from 9:30 am to 4:00 pm and this results in $m = 390$ intradaily observations.[16] There are a few missing observations (less than 5 missing data per day) which we filled in using the previous tick method. Also, the time series of prices contain a few outlying observations that seem to be due to recording errors. To deal with such outliers in quote data, Barndorff-Nielsen and al. (2008b) suggest to delete entries for which the spread is more that 50 times the median spread on that day. Here, we proceed similarly by applying the following cleaning rule:

$$r_{t,j}^{NEW} = \begin{cases} r_{t,j}^{OLD} \text{ if } \left| r_{t,j}^{OLD} \right| \leq 50 \times \underline{r}^{OLD} \\ sign\left( r_{t,j}^{OLD} \right) \times 50 \times \underline{r}^{OLD} \text{ otherwise} \end{cases},$$

where $r_{t,j}^{OLD}$ is the initial data and $\underline{r}^{OLD}$ is the empirical median of $\left| r_{t,j}^{OLD} \right|$ across $t$ and $j$. As shown by Figure 2, this cleaning rule affects very few observations and it does not remove jumps from the data.

Figure 2: Impact of the cleaning rule on the data. Left: realized volatility for $r_{t,j}^{OLD}$. Right: realized volatility for $r_{t,j}^{NEW}$.

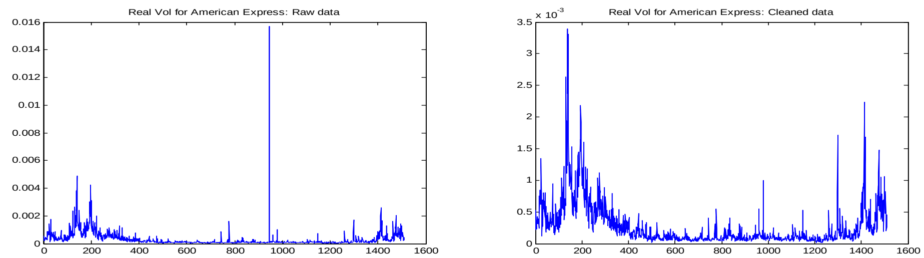



Figure 3 shows examples of volatility signature plots. Except for the General Motor index, the average RV decreases as one samples more and more sparsely. The shape of the graph for General Motors is not typical in the literature and it suggests that the bias of the RV is negative at the highest frequency.

Figure 3: Volatility signature plots for selected indices

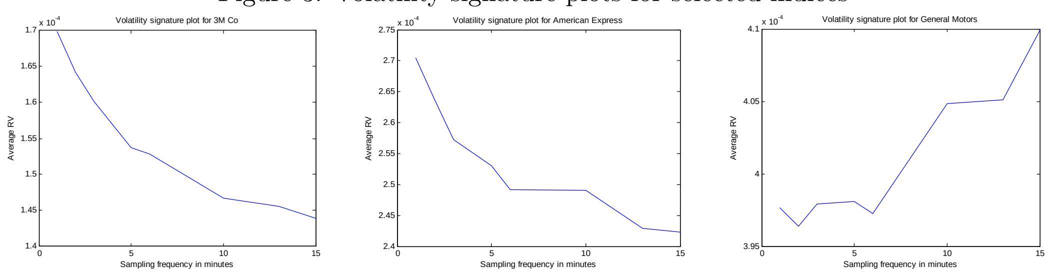



Table 8 shows the output of the GMM estimation of an AR(1) noise model. Of the twelve stocks considered, the AR(1) model is not rejected for six stocks. The overidentification test statistics for Intel Corp and Microsoft are only slightly above the rejection threshold. The autoregressive root $\phi_m$ is estimated to be positive in all cases and it is significantly different from zero in cases where the AR(1) noise model is not rejected. For the AIG stock, $\widehat{\phi}_m$ is very close to unity while it is degenerate (i.e. equal to one) for General Motors. This suggests that the noises contaminating AIG and General Motors obey more sophisticated unit root models.

Table 8: Estimating the AR(1) noise model by GMM based on 15 moments conditions

At level 5%, the AR(1) model is rejected if J-stat>22.362

| | $\widehat{\omega}_0$ | $\widehat{\phi}_m$ | J-stat | Rejection |
|---|---|---|---|---|
| 3M Co. | $1.2 \times 10^{-7}$ ($2.1 \times 10^{-8}$) | 0.675 (0.049) | 11.44 | No |
| Alcoa Inc. | $1.3 \times 10^{-6}$ ($1.1 \times 10^{-6}$) | 0.931 (0.035) | 31.92 | Yes |
| AIG | $3.8 \times 10^{-4}$ ($2.9 \times 10^{-2}$) | 0.998 (0.074) | 9.40 | No |
| American Express | $8.4 \times 10^{-8}$ ($2.1 \times 10^{-8}$) | 0.552 (0.110) | 14.7 | No |
| Dupont and Dupont | $3.3 \times 10^{-7}$ ($1.0 \times 10^{-7}$) | 0.846 (0.033) | 20.2 | No |
| Walt Disney | $4.4 \times 10^{-7}$ ($5.5 \times 10^{-8}$) | 0.727 (0.029) | 21.8 | No |
| General Electric | $3.8 \times 10^{-7}$ ($1.4 \times 10^{-7}$) | 0.868 (0.035) | 12.2 | No |
| General Motors | $1.7 \times 10^{-6}$ ($----$) | 1.000 ($--$) | 29.2 | Yes |
| IBM | $9.9 \times 10^{-8}$ ($2.1 \times 10^{-8}$) | 0.667 (0.065) | 31.1 | Yes |
| Intel Corp. | $4.4 \times 10^{-7}$ ($7.8 \times 10^{-8}$) | 0.766 (0.032) | 22.5 | Yes |
| Hewlett Packard | $7.9 \times 10^{-7}$ ($3.6 \times 10^{-8}$) | 0.881 (0.039) | 28.3 | Yes |
| Microsoft | $4.0 \times 10^{-7}$ ($7.7 \times 10^{-8}$) | 0.833 (0.025) | 23.5 | Yes |

We apply the MA(L) noise model to the stocks for which the AR(1) specification is rejected. Table 9 shows the estimates of the dependence lag of the noise. $\widetilde{L}$ is obtained by minimizing the $\Delta(l)$ criterion (cf. Equation (36) and Figure 4) while $\widehat{L}$ is deduced from the significance tests (32). The estimated dependence lags lies between 8 and 12 minutes.

Table 9: Estimated noise dependence lag

| | $\widetilde{L}$ | $\widehat{L}$ |
|---|---|---|
| Alcoa Inc (AA) | 12 | 12 |
| General Motors (GM) | 8 | 3 |
| IBM | 11 | 11 |
| Intel Corp. (INTC) | 10 | 9 |
| Hewlett-Packard (HPQ) | 9 | 8 |
| Microsoft (MSFT) | 10 | 10 |

Figure 4 shows the plots of $\Delta(l)$ against $L$ (left) and the estimated noise autocovariances (right). To the exception of General Motors, all estimated noise correlograms are positive. This explains the shape of the volatility signature plot of the General Motors index, and it supports that the estimate $\widehat{\phi}_m = 1$ found previously in Table 8 is spurious.

The final step of the empirical study concerns the estimation of the daily integrated volatility. For all assets, we set:

$$K_t^{\varpi} = \varpi_1 K_{t,15}^{BNHLS} + \varpi_2 K_{t,25}^{BNHLS} + \varpi_3 K_{t,35}^{BNHLS} + \varpi_4 K_{t,45}^{BNHLS},$$

and minimize the variance of $K_t^{\varpi}$ with respect to $\varpi = (\varpi_1, \varpi_2, \varpi_3, \varpi_4)$. We implement the adaptive realized kernel as explained in the previous section. The MSEs of all IV estimators are obtained by combining their bias and their variance (see Equations (37) and (38)). The minimum bandwidth $H = 15$ implies that $K_{t,H}^{BNHLS}$ is unbiased under the MA(L) noises identified in Table 9.

Table 10 shows the results. In eight cases out of twelve, the MSE of $K_{t,H}^{BNHLS}$ is minimized for either $H = 25$ or $H = 35$. The MSE in increasing in $H$ in three cases (3M Co, General Motors, IBM) and it is decreasing in one case (AIG). In the latter case for example, the initial estimators ($K_{t,H}^{BNHLS}$) have very similar variances and the differences seen in their MSEs are due to the squared bias term. In all other cases, the variance term dominates the squared bias term in the MSE. More often than not, the initial estimator with smallest MSE receives the largest positive weight when the noise is MA(L). As expected, the adaptive realized kernel $K_t^{\varpi^*}$ outperforms the most efficient of the initial estimators. Arguably, the design of $K_t^{\varpi^*}$ can be improved by combining other estimators based on different kernel functions (Parzen, Tuckey-Hanning, Quadratic spectral). The extra cost for such an improvement resides in the derivation of the biases of the initial estimators.

Figure 5 shows the estimated daily IV processes for all twelve stocks. Although many of the estimated weights in Table 10 are negative, we have found negative IV estimates for one stock only (the AIG index), and this happens for 5 days only out of 1510. An examination of the correlation matrix of the vector of the initial estimators for AIG shows that they are highly correlated. The minimum correlation is 0.9687 and it occurs between $K_{t,15}^{BNHLS}$ and $K_{t,45}^{BNHLS}$. In fact, the noise contaminating the AIG stock price is highly persistence ($\widehat{\phi}_m = 0.998$), and this causes the MSE matrix of the initial estimators to be nearly singular. For this particular stock, a bias corrected version of $K_{t,15}^{BNHLS}$ is more reliable than $K_t^{\varpi^*}$.

Figure 4: Estimation of MA(L) noise. Left: plot of $\Delta(l)$ against $l$. The minimum of $\Delta(l)$ is used as the first guess of $L$. Right: The correlogram of the noise (top) and the associated Student-t (bottom). The line crossing the student-t stats indicates the significance threshold, one-sided, 5%.

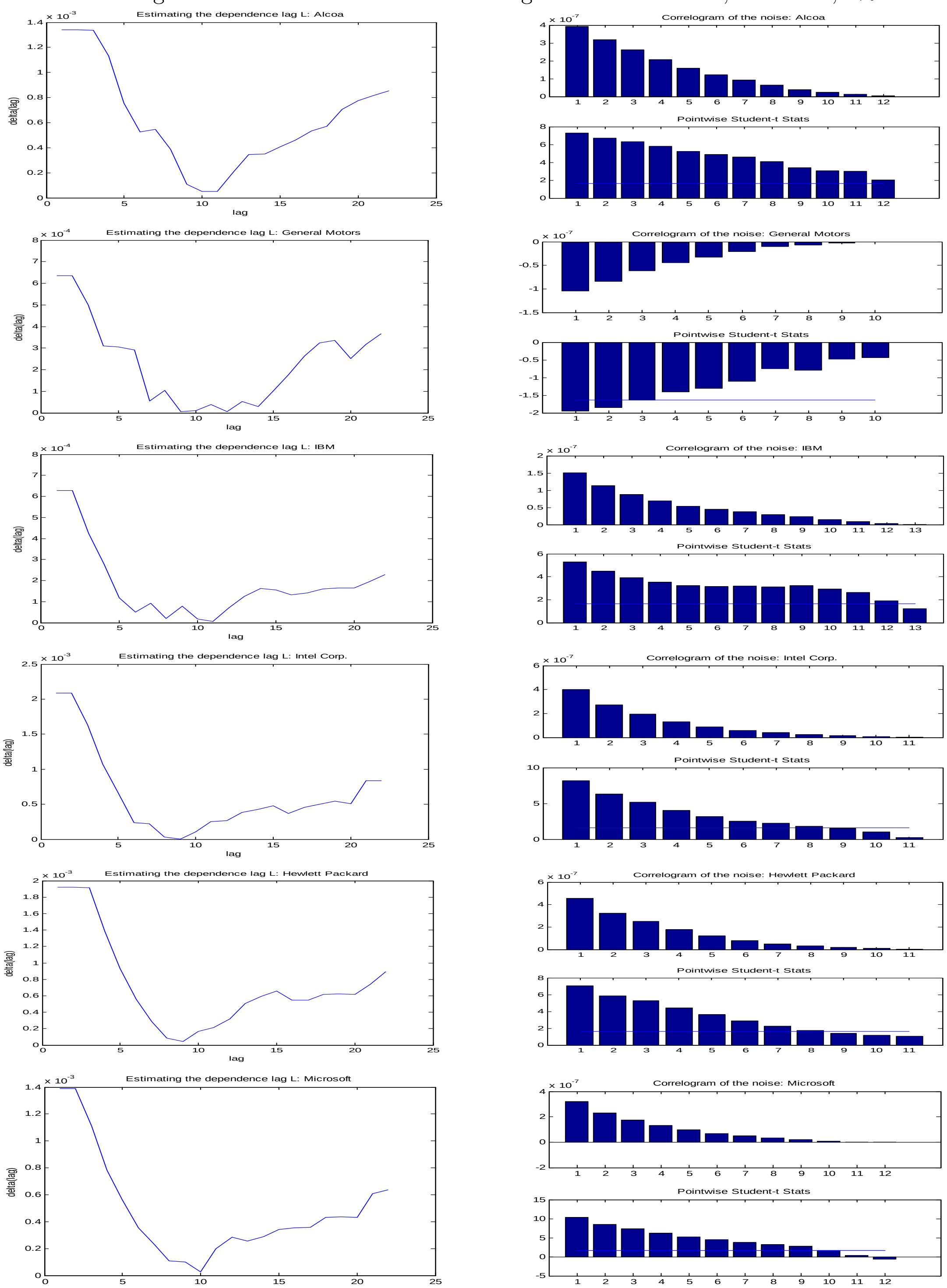
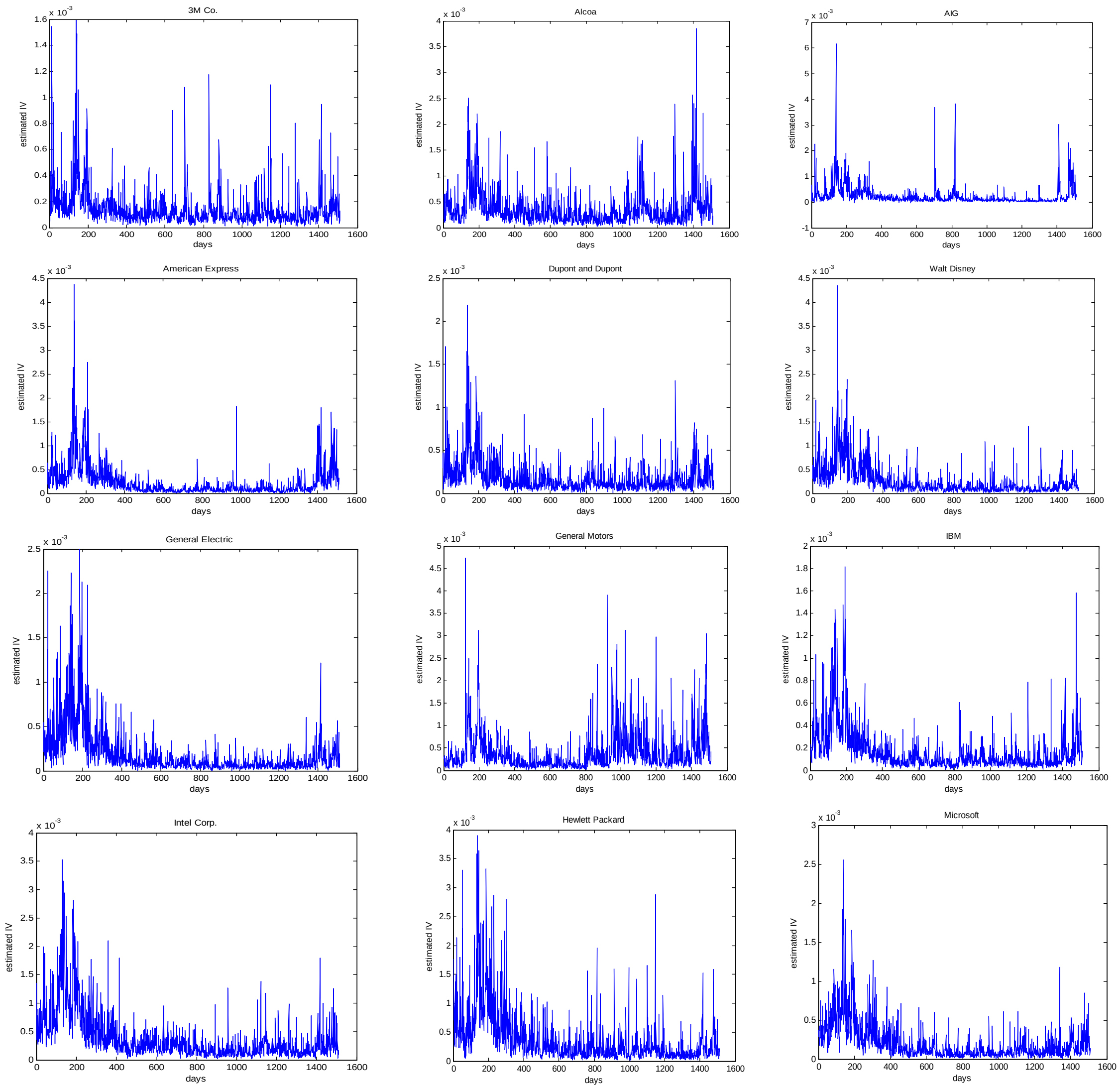


Table 10: Optimal weights for the adaptive realized kernel.

| | | $K_{t,15}^{BNHLS}$ | $K_{t,25}^{BNHLS}$ | $K_{t,35}^{BNHLS}$ | $K_{t,45}^{BNHLS}$ | $K_t^{\varpi^*}$ |
|---|---|---|---|---|---|---|
| | | AR(1) microstructure noise | | | | |
| 3M Co | weight | **2.9153** | −5.1125 | 5.4933 | −2.2960 | – |
| | MSE ($10^{-7}$) | **0.2486** | 0.2693 | 0.2840 | 0.3037 | 0.2324 |
| AIG | weight | −2.8496 | 10.5768 | −12.8812 | **6.1540** | – |
| | MSE ($10^{-3}$) | 3.2523 | 1.1273 | 0.5538 | **0.3226** | 0.0002 |
| American Express | weight | −0.2996 | **1.5967** | −0.4983 | 0.2012 | – |
| | MSE ($10^{-6}$) | 0.1260 | **0.1217** | 0.1234 | 0.1258 | 0.1215 |
| Dupont and Dupont | weight | −0.9431 | **1.8963** | 2.7955 | −2.7487 | – |
| | MSE ($10^{-7}$) | 0.4662 | **0.4364** | 0.4547 | 0.4936 | 0.3973 |
| Walt Disney | weight | −0.6308 | **3.0773** | −0.6188 | −0.8277 | – |
| | MSE ($10^{-6}$) | 0.1126 | **0.1100** | 0.1153 | 0.1213 | 0.1060 |
| General Electric | weight | −0.6004 | **−0.0484** | 4.7721 | −3.1233 | – |
| | MSE ($10^{-7}$) | 0.8050 | **0.7628** | 0.7658 | 0.8059 | 0.7172 |
| | | MA(L) microstructure noise | | | | |
| Alcoa Inc. | weight | −0.6769 | **3.3121** | −1.5601 | −0.0752 | – |
| | MSE ($10^{-6}$) | 0.1205 | **0.1180** | 0.1268 | 0.1354 | 0.1135 |
| General Motors | weight | **2.7928** | 0.0368 | −4.0519 | 2.2223 | – |
| | MSE ($10^{-6}$) | **0.2382** | 0.2838 | 0.3091 | 0.3173 | 0.1915 |
| IBM | weight | **0.7628** | 0.0022 | 1.4136 | −1.1786 | – |
| | MSE ($10^{-7}$) | **0.4057** | 0.4122 | 0.4286 | 0.4485 | 0.4019 |
| Intel Corp. | weight | −0.3987 | **1.7360** | 0.5748 | −0.9121 | – |
| | variance ($10^{-6}$) | 0.1869 | **0.1815** | 0.1867 | 0.1946 | 0.1794 |
| Hewlett Packard | weight | −0.4438 | **0.7396** | 2.4871 | −1.7829 | – |
| | MSE ($10^{-6}$) | 0.2310 | **0.2217** | 0.2245 | 0.2344 | 0.2174 |
| Microsoft | weight | 0.2129 | −0.8753 | **2.4685** | −0.8061 | – |
| | MSE ($10^{-7}$) | 0.6450 | 0.6241 | **0.6155** | 0.6254 | 0.6133 |

Figure 5: Estimated daily integrated volatility by adaptive realized kernels $K_t^{\varpi^*}$.




# 7 Conclusion

We design adaptive realized kernels to estimate the integrated volatility in a framework that combines, on the one hand, a Brownian stochastic volatility model with leverage effect for the frictionless price, and on the other hand, a semi-parametric model for the microstructure noise. The proposed noise model is tied to the frequency at which the price data are recorded and it specifies the noise as the sum of an endogenous term (correlated with the efficient returns) and an exogenous term (uncorrelated with the efficient returns). Our specification for the exogenous noise nests IID, $L$-dependent as well as AR(1) models. The simulation results show that the adaptive realized kernels achieve the optimal trade-off between the discretization error and the microstructure noise. Two inference procedures are proposed for the noise parameters. The first procedure is based on an overidentified generalized method of moments and it is designed for AR(1) types of noise. The second procedure is designed for MA(L) noises and it uses as many moment conditions as there are parameters to be estimated. The simulations show that the AR(1) inference procedure has power against MA(L) alternatives. Hence, our best investigation strategy in practice consists of first testing whether the noise is AR(1) and next, applying the MA(L) inference procedure if the AR(1) specification is rejected. We apply this strategy to twelve stocks listed in the Dow Jones Industrial and find that the AR(1) noise model cannot be rejected for six stocks. For the other stocks, we apply the MA(L) noise inference procedure and find estimates of $L$ that lie between 8 and 12 minutes.

**Appendix: Proofs**

The following Lemma will be used in the proof of Theorem 1.

**Lemma 5** *Assume that* $r_{t,j} = r^*_{(1),t,j} + (1 + a_{t,j})\, r^*_{(2),t,j} - a_{t,j-1} r^*_{(2),t,j-1} + (\varepsilon_{t,j} - \varepsilon_{t,j-1})$ *for some deterministic sequence* $\{a_{t,j}\}$, $j = 1, ..., m$. *Let* $\widetilde{r}_{t,k}$ *be the series of non-overlapping sums of* $q$ *consecutive observations of* $r_{t,j}$, *that is,* $\widetilde{r}_{t,k} = \widetilde{r}_{(1),t,k} + \widetilde{r}_{(2),t,k}$ *with* $\widetilde{r}_{(1),t,k} = \sum_{j=qk-q+1}^{qk} r^*_{(1),t,j}$ *and:*

$$\begin{aligned}\widetilde{r}_{(2),t,k} &= \sum_{j=qk-q+1}^{qk} r^*_{(1),t,j} + (1 + a_{t,qk})\, r^*_{(2),t,qk} + \sum_{j=qk-q+1}^{qk-1} r^*_{(2),t,j} \\ &\quad - a_{t,qk-q} r^*_{(2),t,qk-q} + (\varepsilon_{t,qk} - \varepsilon_{t,qk-q})\end{aligned}$$

*for* $k = 1, ..., m_q$ *and some positive integer* $q \geq 1$ *such that* $m_q = \lfloor m/q \rfloor$. *Then we have:*

$$E\left[RV_t^{(m_q)}\right] = IV_t + 2\sum_{k=1}^{m_q}\left(a_{t,qk} + a^2_{t,qk}\right)\sigma^{*2}_{(2),t,qk} + a^2_{t,0}\sigma^{*2}_{(2),t,0} - a^2_{t,qm_q}\sigma^{*2}_{(2),t,qm_q} + 2m_q\left(\omega_0 - \omega_{m,q}\right),$$
$$Var\left[RV_t^{(m_q)}\right] = O(m_q).$$

**Proof of Lemma 5:** We have: $RV^{(m_q)} = \sum_{k=1}^{m_q} \widetilde{r}^2_{(1),t,k} + 2\sum_{k=1}^{m_q} \widetilde{r}_{(1),t,k}\widetilde{r}_{(2),t,k} + \sum_{k=1}^{m_q} \widetilde{r}^2_{(2),t,k}$, with:

$$\sum_{k=1}^{m_q} \widetilde{r}^2_{(2),t,k} = (1) + (2) + (3) + (4) + (5) + (6) + (7) + (8) + (9)$$

where

$(1) = \sum_{k=1}^{m_q} \left[ (1+a_{t,qk})^2 + a_{t,qk}^2 \right] r_{(2),t,qk}^{*2} + a_{t,0}^2 r_{(2),t,0}^{*2} - a_{t,qm_q}^2 r_{(2),t,qm_q}^{*2}$.

$(2) = \sum_{k=1}^{m_q} \left( \sum_{j=qk-q+1}^{qk-1} r_{(2),t,j}^{*} \right)^2$.

$(3) = \sum_{k=1}^{m_q} (\varepsilon_{t,qk} - \varepsilon_{t,qk-q})^2$.

$(4) = 2\sum_{k=1}^{m_q} \sum_{j=qk-q+1}^{qk-1} (1+a_{t,qk}) r_{(2),t,j}^{*} r_{(2),t,qk}^{*}$.

$(5) = 2\sum_{k=1}^{m_q} (1+a_{t,qk}) a_{t,qk-q} r_{(2),t,qk-q}^{*} r_{(2),t,qk}^{*}$.

$(6) = 2\sum_{k=1}^{m_q} (1+a_{t,qk}) (\varepsilon_{t,qk} - \varepsilon_{t,qk-q}) r_{(2),t,qk}^{*}$.

$(7) = -2\sum_{k=1}^{m_q} \sum_{j=qk-q+1}^{qk-1} a_{t,qk-q} r_{(2),t,j}^{*} r_{(2),t,qk-q}^{*}$.

$(8) = 2\sum_{k=1}^{m_q} \sum_{j=qk-q+1}^{qk-1} (\varepsilon_{t,qk} - \varepsilon_{t,qk-q}) r_{(2),t,j}^{*}$.

$(9) = -2\sum_{k=1}^{m_q} a_{t,qk-q} (\varepsilon_{t,qk} - \varepsilon_{t,qk-q}) r_{(2),t,qk-q}^{*}$.

Only squared terms have nonzero expectation:

$$\begin{aligned} E\left[RV^{(m_q)}\right] &= \sum_{k=1}^{m_q} \sum_{j=qk-q+1}^{qk} \sigma_{(1),t,j}^{*2} + \sum_{k=1}^{m_q} \sum_{j=qk-q+1}^{qk-1} \sigma_{(2),t,j}^{*2} + \sum_{k=1}^{m_q} \left[ (1+a_{t,qk})^2 + a_{t,qk}^2 \right] \sigma_{(2),t,qk}^{*2} \\ &\quad + m_q E\left[ (\varepsilon_{t,qk} - \varepsilon_{t,qk-q})^2 \right] + a_{t,0}^2 \sigma_{(2),t,0}^{*2} - a_{t,qm_q}^2 \sigma_{(2),t,qm_q}^{*2} \\ &= IV_t + 2\sum_{k=1}^{m_q} \left( a_{t,qk} + a_{t,qk}^2 \right) \sigma_{(2),t,qk}^{*2} + 2m_q (\omega_0 - \omega_{m,q}) + a_{t,0}^2 \sigma_{(2),t,0}^{*2} - a_{t,qm_q}^2 \sigma_{(2),t,qm_q}^{*2}. \end{aligned}$$

where $\omega_{m,q} = E[\varepsilon_{t,j}\varepsilon_{t,j-q}]$ is independent of $t$ and $j$. Also, all the terms involved in the expression of $\sum_{k=1}^{m_q} \widetilde{r}_{(2),t,k}^2$ are uncorrelated. Thus:

$$\begin{aligned} Var\left[ \sum_{k=1}^{m_q} \widetilde{r}_{(2),t,k}^2 \right] &= Var((1)) + Var((2)) + Var((3)) + Var((4)) \\ &\quad + Var((5)) + Var((6)) + Var((7)) + Var((8)) + Var((9)), \end{aligned}$$

where

$$\begin{aligned} Var((1)) = 2\sum_{k=1}^{m_q} \left[ (1+a_{t,qk})^2 + a_{t,qk}^2 \right]^2 \sigma_{(2),t,qk}^{*4} + 2a_{t,0}^4 \sigma_{(2),t,0}^{*4} \\ -2a_{t,qm_q}^4 \sigma_{(2),t,qm_q}^{*4} - 4a_{t,qm_q}^2 \left(1+a_{t,qm_q}\right)^2 \sigma_{(2),t,qm_q}^{*4}. \end{aligned}$$

$Var((2)) = 2\sum_{k=1}^{m_q} \left( \sum_{l=qk-q+1}^{qk-1} \sum_{j=qk-q+1}^{qk-1} \sigma_{(2),t,j}^{*2} \sigma_{(2),t,l}^{*2} \right)$.

$Var((4)) = 4\sum_{k=1}^{m_q} \sum_{j=qk-q+1}^{qk-1} (1+a_{t,qk})^2 \sigma_{(2),t,j}^{*2} \sigma_{(2),t,qk}^{*2}$.

$Var((5)) = 4\sum_{k=1}^{m_q} (1+a_{t,qk})^2 a_{t,qk-q}^2 \sigma_{(2),t,qk-q}^{*2} \sigma_{(2),t,qk}^{*2}$.

$$\begin{aligned} Var((6)) &= 4\sum_{k=1}^{m_q} (1+a_{t,qk})^2 Var(\varepsilon_{t,qk} - \varepsilon_{t,qk-q}) Var\left( r_{(2),t,qk}^{*} \right) \\ &= 8(\omega_0 - \omega_{m,q}) \sum_{k=1}^{m} (1+a_{t,qk})^2 \sigma_{(2),t,qk}^{*2}. \end{aligned}$$

$Var((7)) = 4\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1} a_{t,qk-q}^2 \sigma_{(2),t,j}^{*2}\sigma_{(2),t,qk-q}^{*2}$.

$Var((8)) = 8\left(\omega_0 - \omega_{m,q}\right)\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{(2),t,j}^{*2}$.

$Var((9)) = 8\left(\omega_0 - \omega_{m,q}\right)\sum_{k=1}^{m_q} a_{t,qk-q}^2\sigma_{(2),t,qk-q}^{*2}$.

Hence:

$$
\begin{aligned}
Var\left[\sum_{k=1}^{m_q}\widetilde{r}_{(2),t,k}^2\right] = {} & 2\sum_{k=1}^{m_q}\left[\left(1+a_{t,qk}\right)^2 + a_{t,qk}^2\right]^2 \sigma_{(2),t,qk}^{*4} \\
& +2\sum_{k=1}^{m_q}\left(\sum_{l=qk-q+1}^{qk-1}\sum_{j=qk-q+1}^{qk-1}\sigma_{(2),t,j}^{*2}\sigma_{(2),t,l}^{*2}\right) \\
& +Var\left[\sum_{k=1}^{m_q}\left(\varepsilon_{t,qk}-\varepsilon_{t,qk-q}\right)^2\right] + 4\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\left(1+a_{t,qk}\right)^2\sigma_{(2),t,j}^{*2}\sigma_{(2),t,qk}^{*2} \\
& +4\sum_{k=1}^{m_q}\left(1+a_{t,qk}\right)^2 a_{t,qk-q}^2\sigma_{(2),t,qk-q}^{*2}\sigma_{(2),t,qk}^{*2} + 4\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1} a_{t,qk-q}^2\sigma_{(2),t,j}^{*2}\sigma_{(2),t,qk}^{*2} \\
& +8\left(\omega_0-\omega_{m,q}\right)\sum_{k=1}^{m_q}\left(1+a_{t,qk}\right)^2\sigma_{(2),t,qk}^{*2} + 8\left(\omega_0-\omega_{m,q}\right)\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{(2),t,j}^{*2} \\
& +8\left(\omega_0-\omega_{m,q}\right)\sum_{k=1}^{m_q} a_{t,qk-q}^2\sigma_{(2),t,qk-q}^{*2} + 2a_{t,0}^4\sigma_{(2),t,0}^{*4} - 2a_{t,qm_q}^4\sigma_{(2),t,qm_q}^{*4} \\
& -4a_{t,qm_q}^2\left(1+a_{t,qm_q}\right)^2\sigma_{(2),t,qm_q}^{*4}.
\end{aligned}
$$

The presence of the term $Var\left[\sum_{k=1}^{m_q}\left(\varepsilon_{t,qk}-\varepsilon_{t,qk-q}\right)^2\right]$ in the expression of the variance of $\sum_{k=1}^{m_q}\widetilde{r}_{(2),t,k}^2$ shows that $Var\left[RV^{(m_q)}\right] = O(m_q)$■

The following Lemma will be used in the proof of Theorem 3.

**Lemma 6** *Under the assumptions of Theorem 3, we have:*

$$
\begin{aligned}
& E\left[RV_t^{(AC,m,1)}\right] = IV_t + \left(2a_{t,m}+a_{t,m}^2\right)\sigma_{t,m}^{*2} - \left(2a_{t,0}+a_{t,0}^2\right)\sigma_{t,0}^{*2} \\
& Var\left[RV_t^{(AC,m,1)}\right] = O(m).
\end{aligned}
$$

**Proof of Lemma 6:** Let $r_{t,j} = r_{(1),t,j}^* + \widetilde{r}_{(2),t,j}$, where $\widetilde{r}_{(2),t,j} = \left(1+a_{t,j}\right)r_{(2),t,j}^* - a_{t,j-1}r_{(2),t,j-1}^* + \left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)$.

We first note that:

$$
\begin{aligned}
RV_t^{(AC,m,1)} &= RV_t^{(AC,m,1)}\left(r_{(1)}^*\right) + 2\sum_{j=1}^{m} r_{(1),t,j}^*\widetilde{r}_{(2),t,j} + 2\sum_{j=1}^{m} r_{(1),t,j}^*\widetilde{r}_{(2),t,j-1} \\
&\quad +2\sum_{j=1}^{m}\widetilde{r}_{(2),t,j}r_{(1),t,j-1}^* + RV_t^{(AC,m,1)}\left(\widetilde{r}_{(2)}\right)
\end{aligned}
$$

where $RV_t^{(AC,m,1)}\left(r_{(1)}^*\right) = \sum_{j=1}^{m} r_{(1),t,j}^{*2} + 2\sum_{j=1}^{m} r_{(1),t,j}^* r_{(2),t,j-1}^*$ and $RV_t^{(AC,m,1)}\left(\widetilde{r}_{(2)}\right) = \sum_{j=1}^{m}\widetilde{r}_{(2),t,j}^2 + 2\sum_{j=1}^{m}\widetilde{r}_{(2),t,j}\widetilde{r}_{(2),t,j-1}$.

$$
RV_t^{(AC,m,1)}\left(\widetilde{r}_{(2)}\right) = (I)+(II)+(III)+(IV)+(V)+(VI)+(VII)+(VIII)+(IX),
$$

where

$(I)=\sum_{j=1}^{m}r_{(2),t,j}^{*2}+\left(2a_{t,m}+a_{t,m}^{2}\right)r_{(2),t,m}^{*2}-\left(2a_{t,0}+a_{t,0}^{2}\right)r_{(2),t,0}^{*2}$.

$(II)=2\sum_{j=1}^{m}\left(1+a_{t,j}+a_{t,j}a_{t,j-1}\right)r_{(2),t,j}^{*}r_{(2),t,j-1}^{*}+2a_{t,-1}a_{t,0}r_{(2),t,-1}^{*}r_{(2),t,0}^{*}-2a_{t,m-1}a_{t,m}r_{(2),t,m-1}^{*}r_{(2),t,m}^{*}$.

$(III)=-2\sum_{j=1}^{m}\left(1+a_{t,j}\right)a_{t,j-2}r_{(2),t,j}^{*}r_{(2),t,j-2}^{*}$.

$(IV)=2\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)r_{(2),t,j}^{*}-2a_{t,0}\left(\varepsilon_{t,0}-\varepsilon_{t,-1}\right)r_{(2),t,0}^{*}+2a_{t,m}\left(\varepsilon_{t,m}-\varepsilon_{t,m-1}\right)r_{(2),t,m}^{*}$.

$(V)=2\sum_{j=1}^{m}\left(1+a_{t,j}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)r_{(2),t,j}^{*}$.

$(VI)=2\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)r_{(2),t,j-1}^{*}$.

$(VII)=-2\sum_{j=1}^{m}a_{t,j-2}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)r_{(2),t,j-2}^{*}$.

$(VIII)=2\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)$.

$(IX)=\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)^{2}$.

Because only squared terms will have nonzero expectation, we have:

$$\begin{aligned}E\left[RV_t^{(AC,m,1)}\right] &= E\left[RV_t^{(AC,m,1)}\left(r_{(1)}^{*}\right)\right]+E\left[(I)\right]\\ &= IV_t+\left(1-\rho^{2}\right)\left[\left(2a_{t,m}+a_{t,m}^{2}\right)\sigma_{t,m}^{*2}-\left(2a_{t,0}+a_{t,0}^{2}\right)\sigma_{t,0}^{*2}\right].\end{aligned}$$

The calculation of that variance of $RV_t^{(AC,m,1)}\left(\widetilde{r}_{(2)}\right)$ is simplified by noting that only the terms $(IV)$ to $(IX)$ are possibly correlated. We have:

$$\begin{aligned}Var((I)) &= 2\sum_{j=1}^{m}\sigma_{(2),t,j}^{*4}+2\left(2a_{t,0}+a_{t,0}^{2}\right)^{2}\sigma_{(2),t,0}^{*4}+2\left(2a_{t,m}+a_{t,m}^{2}\right)^{2}\sigma_{(2),t,m}^{*4}\\ &\quad +2\left(2a_{t,m}+a_{t,m}^{2}\right)\sigma_{(2),t,m}^{*4}\end{aligned}$$

$$\begin{aligned}Var((II)) &= 4\sum_{j=1}^{m}\left(1+a_{t,j}+a_{t,j}a_{t,j-1}\right)^{2}\sigma_{(2),t,j}^{*2}\sigma_{(2),t,j-1}^{*2}+4a_{t,-1}^{2}a_{t,0}^{2}\sigma_{(2),t,-1}^{*2}\sigma_{(2),t,0}^{*2}\\ &\quad +4a_{t,m-1}^{2}a_{t,m}^{2}\sigma_{(2),t,m-1}^{*2}\sigma_{(2),t,m}^{*2}\\ &\quad -8a_{t,m-1}a_{t,m}\left(1+a_{t,m}+a_{t,m}a_{t,m-1}\right)\sigma_{(2),t,m-1}^{*2}\sigma_{(2),t,m}^{*2}\end{aligned}$$

$$Var((III))=4\sum_{j=1}^{m}\left(1+a_{t,j}\right)^{2}a_{t,j-2}^{2}\sigma_{(2),t,j}^{*2}\sigma_{(2),t,j-2}^{*2}$$

$$Var((IV))=8\omega_0\sum_{j=1}^{m}\sigma_{(2),t,j}^{*2}+8\omega_0\left(a_{t,0}^{2}\sigma_{(2),t,0}^{*2}+a_{t,m}^{2}\sigma_{(2),t,m}^{*2}\right)+16\omega_0 a_{t,m}\sigma_{(2),t,m}^{*2}$$

$$\begin{aligned}2Cov\left((IV),(V)\right) &= 8\sum_{j=1}^{m}\left(1+a_{t,j}\right)E\left[\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)\right]E\left(r_{t,j}^{*2}\right)\\ &= -8\omega_0\sum_{j=1}^{m}\sigma_{(2),t,j}^{*2}-8\omega_0\sum_{j=1}^{m}a_{t,j}\sigma_{(2),t,j}^{*2}\end{aligned}$$

$$\begin{aligned}2Cov\left((IV),(VI)\right) &= 8\sum_{j=1}^{m-1}E\left[\left(\varepsilon_{t,j+1}-\varepsilon_{t,j}\right)\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\right]E\left(r_{t,j}^{*2}\right)\\ &= -8\omega_0\sum_{j=1}^{m}\sigma_{(2),t,j}^{*2}+8\omega_0\sigma_{(2),t,m-1}^{*2}\end{aligned}$$

$$2Cov\left((IV),(VII)\right)=-8\sum_{j=1}^{m-2}a_{t,j}E\left[\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j+2}-\varepsilon_{t,j+1}\right)\right]E\left(r_{t,j}^{*2}\right)=0$$

$$2Cov\left((IV),(VIII)\right)=2Cov\left((IV),(IX)\right)=0$$

$Var((V)) = 8\omega_0 \sum_{j=1}^{m} (1+a_{t,j})^2 \sigma^{*2}_{(2),t,j}$

$2Cov((V),(VI)) = 2Cov((V),(VII)) =$

$$2Cov((V),(VIII)) = 2Cov((V),(IX)) = 0$$

$Var((VI)) = 8\omega_0 \sum_{j=1}^{m} \sigma^{*2}_{(2),t,j} - 8\omega_0 \sigma^{*2}_{(2),t,0}$

$2Cov((VI),(VII)) = -8 \sum_{j=1}^{m-2} a_{t,j} E[(\varepsilon_{t,j+1}-\varepsilon_{t,j})(\varepsilon_{t,j+2}-\varepsilon_{t,j+1})] E\left(r^{*2}_{t,j}\right)$

$$= 8\omega_0 \sum_{j=1}^{m} a_{t,j}\sigma^{*2}_{(2),t,j} - 8\omega_0 \left(a_{t,m-1}\sigma^{*2}_{(2),t,m-1} + a_{t,m}\sigma^{*2}_{(2),t,m}\right)$$

$2Cov((VI),(VIII)) = 2Cov((VI),(IX)) = 0$

$Var((VII)) = 8\omega_0 \sum_{j=1}^{m} a^2_{t,j}\sigma^{*2}_{(2),t,j} + 8\omega_0 \left(a^2_{t,-1}\sigma^{*2}_{(2),t,-1} + a^2_{t,0}\sigma^{*2}_{(2),t,0} - a^2_{t,m-1}\sigma^{*2}_{(2),t,m-1} - a^2_{t,m}\sigma^{*2}_{(2),t,m}\right)$

$Cov((VII),(VIII)) = Cov((VII),(IX)) = 0$

$Var((VIII)) = Var\left[2\sum_{j=1}^{m} (\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-1}-\varepsilon_{t,j-2})\right]$

$$= 4Var\left[2\sum_{j=1}^{m} \varepsilon_{t,j}\varepsilon_{t,j-1} + \varepsilon_{t,0}\varepsilon_{t,-1} + \varepsilon_{t,m}\varepsilon_{t,m-1} + \sum_{j=1}^{m} +\varepsilon_{t,j}\varepsilon_{t,j-2} + \sum_{j=1}^{m} \varepsilon^2_{t,j-1}\right]$$

$$= 4mE[\varepsilon^4_{t,j}] + 16m\omega_0^2 - 8\omega_0^2$$

$2Cov((VIII),(IX)) = 4Cov\left[\sum_{j=1}^{m} (\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}),\ \sum_{k=1}^{m} (\varepsilon_{t,k}-\varepsilon_{t,k-1})^2\right]$

$$= -(8m-4)\left(E[\varepsilon^4_{t,j}] + \omega_0^2\right)$$

since we have:

$E[(\varepsilon_{t,j+k}-\varepsilon_{t,j+k-1})^2 (\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-1}-\varepsilon_{t,j-2})] = -2\omega_0^2 \ \forall\ k \geq 1$

$E[(\varepsilon_{t,j}-\varepsilon_{t,j-1})^3 (\varepsilon_{t,j-1}-\varepsilon_{t,j-2})] = -E[\varepsilon^4_{t,j}] - 3\omega_0^2$ (for $k = j$)

$E[(\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-1}-\varepsilon_{t,j-2})^3] = -E[\varepsilon^4_{t,j}] - 3\omega_0^2$ (for $k = j-1$)

$E[(\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-1}-\varepsilon_{t,j-2})(\varepsilon_{t,j-k-1}-\varepsilon_{t,j-k-2})^2] = -2\omega_0^2 \ \forall\ k \geq 1$

$\Rightarrow E\left[\left(\sum_{j=1}^{m} (\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-1}-\varepsilon_{t,j-2})\right)\left(\sum_{j=1}^{m} (\varepsilon_{t,j}-\varepsilon_{t,j-1})^2\right)\right]$

$$= (-2m+1)E[\varepsilon^4_{t,j}] + (-2m^2-2m+1)\omega_0^2$$

Also: $E\left(\sum_{j=1}^{m} (\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-1}-\varepsilon_{t,j-2})\right) = -m\omega_0$

and $E\left(\sum_{j=1}^{m} (\varepsilon_{t,j}-\varepsilon_{t,j-1})^2\right) = 2m\omega_0$

Thus $Cov((VIII),(IX)) = (-2m+1)E[\varepsilon^4_{t,j}] + (-2m^2-2m+1)\omega_0^2 + 2m^2\omega_0^2$

$$= -(2m-1)\left(E[\varepsilon^4_{t,j}] + \omega_0^2\right)$$

$Var((IX)) = 4mE\left[\varepsilon^4_{t,j}\right] + 2\left(\omega_0^2 - E\left[\varepsilon^4_{t,j}\right]\right)$

The sum of all these terms gives:

$$\begin{aligned} Var\left[RV_t^{(AC,m,1)}\left(\widetilde{r}_{(2)}\right)\right] &= 2\sum_{j=1}^{m} \sigma^{*4}_{(2),t,j} + 4\sum_{j=1}^{m} (1+a_{t,j}+a_{t,j}a_{t,j-1})^2 \sigma^{*2}_{(2),t,j}\sigma^{*2}_{(2),t,j-1} \\ &+4\sum_{j=1}^{m} (1+a_{t,j})^2 a^2_{t,j-2}\sigma^{*2}_{(2),t,j}\sigma^{*2}_{(2),t,j-2} + 8\omega_0 \sum_{j=1}^{m} (1+a_{t,j})^2 \sigma^{*2}_{(2),t,j} \\ &+8\omega_0 \sum_{j=1}^{m} a^2_{t,j}\sigma^{*2}_{(2),t,j} + 8m\omega_0^2 + 2\left(E\left[\varepsilon^4_{t,j}\right] - \omega_0^2\right) + 2\left(2a_{t,0}+a^2_{t,0}\right)^2 \sigma^{*4}_{(2),t,0} \end{aligned}$$

$$
\begin{aligned}
&+2\left(2a_{t,m}+a_{t,m}^2\right)^2\sigma_{(2),t,m}^{*4}+2\left(2a_{t,m}+a_{t,m}^2\right)\sigma_{(2),t,m}^{*4}+4a_{t,-1}^2a_{t,0}^2\sigma_{(2),t,-1}^{*2}\sigma_{(2),t,0}^{*2}\\
&-8a_{t,m-1}a_{t,m}\left(1+a_{t,m}+a_{t,m}a_{t,m-1}\right)\sigma_{(2),t,m-1}^{*2}\sigma_{(2),t,m}^{*2}\\
&+4a_{t,m-1}^2a_{t,m}^2\sigma_{(2),t,m-1}^{*2}\sigma_{(2),t,m}^{*2}+8\omega_0\left(\sigma_{(2),t,m-1}^{*2}-\sigma_{(2),t,0}^{*2}\right)\\
&+8\omega_0\left(a_{t,-1}^2\sigma_{(2),t,-1}^{*2}+2a_{t,0}^2\sigma_{(2),t,0}^{*2}+a_{t,m}\sigma_{(2),t,m}^{*2}\right)\\
&-8\omega_0\left(a_{t,m-1}\sigma_{(2),t,m-1}^{*2}+a_{t,m-1}^2\sigma_{(2),t,m-1}^{*2}\right).
\end{aligned}
$$

The presence of the term $8m\omega_0^2$ in the expression of this variance shows that $Var\left[RV_t^{(AC,m,1)}\right]=O(m)$■

**Proof of Theorem 1:** Substituting for $a_{t,j}=\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,j}^*}$ in Lemma 5, we get the expectation:

$$
\begin{aligned}
E\left[RV_t^{(m_q)}\right] &= IV_t+\left(1-\rho^2\right)\left[\frac{2\beta_1^2}{q}+\frac{2\left(2\beta_0+1\right)\beta_1}{\sqrt{m}}\sum_{k=1}^{m_q}\sigma_{t,qk}^*+2\beta_0\left(\beta_0+1\right)\sum_{k=1}^{m_q}\sigma_{t,qk}^{*2}\right]\\
&\quad +2m_q\left(\omega_0-\omega_{m,q}\right)+\left(1-\rho^2\right)\left[\beta_0^2\left(\sigma_{t,0}^{*2}-\sigma_{t,m}^{*2}\right)+\frac{2\beta_0\beta_1}{\sqrt{m}}\left(\sigma_{t,0}^*-\sigma_{t,m}^*\right)\right].
\end{aligned}
$$

We do not calculate the exact variance of $RV_t^{(m_q)}$ because all we need to know is that it is $O(m)$, as shown in Lemma 5■

**Proof of Theorem 2:** Substituting for $a_{t,j}=\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,j}}$ in Lemma 6, yield:

$$
E\left[RV_t^{(AC,m,1)}\right]=IV_t+\left(1-\rho^2\right)\left[\left(\beta_0^2+2\beta_0\right)\left(\sigma_{t,m}^{*2}-\sigma_{t,0}^{*2}\right)-\frac{2\beta_1\left(1+\beta_0\right)}{\sqrt{m}}\left(\sigma_{t,m}^*-\sigma_{t,0}^*\right)\right].
$$

We do not calculate the exact variance of $RV_t^{(AC,m,1)}$ because all we need to know is that it is $O(m)$, as shown in Lemma 6■

**Lemma 7** *Assume that $\beta_0=\beta_1=\rho=0$ and $k(x)=1-x$ (the Bartlett kernel). Under Assumptions E1 and E2, we have:*

$$
\begin{aligned}
K_t\left(r^*\right)-IV_t &= O_p(H^{1/2}m^{-1/2}),\\
K_t\left(\Delta u\right) &= -\varepsilon_{t,0}^2+\varepsilon_{t,m}^2-\frac{4}{H}\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-H}-\frac{2}{H}\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-H-1}\\
&\quad -\frac{2}{H}\sum_{h=2}^{H-1}\left(\varepsilon_{t,0}\varepsilon_{t,-h}-\varepsilon_{t,m}\varepsilon_{t,m-h}\right)+\frac{2}{H}\left(\varepsilon_{t,0}\varepsilon_{t,-H}-\varepsilon_{t,m}\varepsilon_{t,m-H}\right),
\end{aligned}
$$

*as $m\to\infty$ and $H=Dm^\gamma$ for $\gamma\in(0,1)$.*

**Proof of Lemma 7:** The result for $K_t(r^*)$ follows from Theorem 1 of Barndorff-Nielsen and al (2008a). We now examine the term $K_t(\Delta\varepsilon)$. We have:

$$\begin{aligned} K_t(\Delta\varepsilon) &= \sum_{j=1}^{m}(\varepsilon_{t,j}-\varepsilon_{t,j-1})^2 + 2\sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\sum_{j=1}^{m}(\varepsilon_{s,j}-\varepsilon_{s,j-1})(\varepsilon_{s,j-h}-\varepsilon_{s,j-h-1}), \\ &= V_t^{(AC,m,1)} + 2\sum_{h=2}^{H} k\left(\frac{h-1}{H}\right)\sum_{j=1}^{m}(\varepsilon_{s,j}-\varepsilon_{s,j-1})(\varepsilon_{s,j-h}-\varepsilon_{s,j-h-1}), \end{aligned}$$

with $V_t^{(AC,m,1)} = 2\sum_{j=1}^{m}\varepsilon_{t,j}(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}) - \varepsilon_{t,0}^2 + \varepsilon_{t,m}^2 + 2(\varepsilon_{t,0}\varepsilon_{t,-1}-\varepsilon_{t,m}\varepsilon_{t,m-1})$, and for $h \geq 2$:

$$\begin{aligned} \sum_{j=1}^{m}(\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-h}-\varepsilon_{t,j-h-1}) &= -\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-h+1} + 2\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-h} - \sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-h-1} \\ &\quad -(\varepsilon_{t,0}\varepsilon_{t,-h+1}-\varepsilon_{t,m}\varepsilon_{t,m-h+1}) + (\varepsilon_{t,0}\varepsilon_{t,-h}-\varepsilon_{t,m}\varepsilon_{t,m-h}) \end{aligned}$$

Summing over $H$ yields:

$$\begin{aligned} & 2\sum_{h=2}^{H} k\left(\frac{h-1}{H}\right)\sum_{j=1}^{m}(\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-h}-\varepsilon_{t,j-h-1}) \\ =\ & -2\sum_{j=1}^{m}\varepsilon_{t,j}(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}) - \frac{4}{H}\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-H} - \frac{2}{H}\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-H-1} \\ & -\frac{2}{H}\sum_{h=2}^{H-1}(\varepsilon_{t,0}\varepsilon_{t,-h}-\varepsilon_{t,m}\varepsilon_{t,m-h}) - 2(\varepsilon_{t,0}\varepsilon_{t,-1}-\varepsilon_{t,m}\varepsilon_{t,m-1}) + \frac{2}{H}(\varepsilon_{t,0}\varepsilon_{t,-H}-\varepsilon_{t,m}\varepsilon_{t,m-H}) \end{aligned}$$

Finally, we have:

$$\begin{aligned} K_t(\Delta\varepsilon) &= -\varepsilon_{t,0}^2 + \varepsilon_{t,m}^2 - \frac{4}{H}\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-H} - \frac{2}{H}\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-H-1} \\ &\quad -\frac{2}{H}\sum_{h=2}^{H-1}(\varepsilon_{t,0}\varepsilon_{t,-h}-\varepsilon_{t,m}\varepsilon_{t,m-h}) + \frac{2}{H}(\varepsilon_{t,0}\varepsilon_{t,-H}-\varepsilon_{t,m}\varepsilon_{t,m-H}) \end{aligned}$$

■

**Proof of Theorem 3:** When $\varepsilon_{t,j}$ is IID, all autocovariances of order $h \geq 1$ are equal to zero and Lemma 7 implies: $K_t(r^*) - IV_t = O_p(H^{1/2}m^{-1/2})$, $K_t(r^*, \Delta u) = O_p(H^{-1/2})$ and $K_t(\Delta u) =$

$-\varepsilon_{t,0}^2 + \varepsilon_{t,m}^2 + O_p(H^{-1}m^{1/2})$. By setting $H$ proportional to $m^{2/3}$, we obtain:

$$\begin{aligned}
K_{t,Lead}^{BNHLS} &= K_t\left(r^*\right) + K_t\left(r^*, \Delta u\right) + K_t\left(\Delta u, r^*\right) + K_t\left(\Delta u\right) \\
&= O_p(m^{-1/6}) + O_p(m^{-1/3}) + O_p(m^{-1/3}) - \varepsilon_{t,0}^2 + \varepsilon_{t,m}^2 + O_p(m^{-1/6}) \\
&= -\varepsilon_{t,0}^2 + \varepsilon_{t,m}^2 + O_p(m^{-1/6})
\end{aligned}$$

■

**Proof of Theorem 4:** Under AR(1) noise with autoregressive root $\phi_m$, we see from Lemma 7 that the bias of $K_{t,Lead}^{BNHLS}$ induced by $K_t\left(\Delta u\right)$ is equal to

$$E\left(K_t\left(\Delta u\right)\right) = -mH^{-1}\left(4\omega_{m,H} + 2\omega_{m,H+1}\right) = -2mH^{-1}\left(\phi_m\right)^H\left(2+\phi_m\right)\omega_0.$$

Let $H = Cm^\gamma$ for strictly positive constants $C$ and $\gamma$. Then, the absolute bias is:

$$\left|E\left(K_t\left(\Delta u\right)\right)\right| = 2\left(2+\phi_m\right)\omega_0 C\exp\left(\left(1-\gamma\right)\log m + Cm^\gamma \ln\left|\phi_m\right|\right),$$

which converges to zero fast if $\phi_m \in (-1, 1)$. If $\phi_m = 1 - Dm^{-\alpha}$, then:

$$\left|E\left(K_t\left(\Delta u\right)\right)\right| \simeq 2\left(2+\phi_m\right)\omega_0 C\exp\left(\left(1-\gamma\right)\log m - CDm^{\gamma-\alpha}\right),$$

with the convention that $\ln\left|\phi_m\right| = D$ if $\alpha = 0$. This yields the first result.

For the second result, $E\left(K_t\left(\Delta u\right)\right) = -mH^{-1}\left(4\omega_{m,H} + 2\omega_{m,H+1}\right)$ while $\omega_{m,H} = 0$ for $H \geq L+1$ under an MA(L) noise■

**Acknowledgement:**

We are grateful to the editor Torben Andersen, an associate editor and two anonymous referees for their valuable suggestions. We also thank Valentina Corradi, Prosper Dovonon, Rene Garcia, Silvia Goncalves, Ilze Kalnina, and Nour Meddahi for helpful comments. An earlier draft of this paper has circulated under the title "*Assessing the Nature of Pricing Inefficiencies via Realized Measures*". Marine Carrasco gratefully acknowledges partial funding from SSHRC.

## References


[1] 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. *The Review of Financial Studies* 18, No 2.

[2] Ait-Sahalia, Y., P.A. Mykland and L. Zhang. 2011. Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160, 160-175.

[3] Andersen, T. G., and T. Bollerslev. 1998. Answering the skeptics: Yes, standard volatility model do provide accurate forecast. *International Economics Review* 39, No. 4.

[4] Andersen, T. G., T. Bollerslev, F.X. Diebold, and P. Labys. 2003. Modelling and forecasting realized volatility. *Econometrica* 71, 579-625.

[5] Andreou, E., and E. Ghysels. 2002. Rolling-Sample Volatility Estimators: Some new theoretical, simulation and empirical Results. *Journal of Business and Economic Statistics*, Forthcoming.

[6] Andrews, D.W.K., and J.C. Monahan. 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60, 953-966.

[7] Bandi, F. M., and J. R. Russell. 2008. Microstructure noise, realized variance and optimal sampling. *Review of Economic Studies* 75, 339-369.

[8] Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard. 2008a. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* 76, 1481-1536.

[9] Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard. 2008b. Realised kernels in practice: trades and quotes. *Econometrics Journal 04*, 1-32.

[10] Barndorff-Nielsen, O. E., and N. Shephard. 2002. Estimating quadratic variation using realized variance. *Journal of Applied Econometrics* 17, 457-477.

[11] Bollen, B. and B. Inder. 2002. Estimating daily volatility in financial markets utilizing intraday data. *Journal of Empirical Finance* 9, 551-562.

[12] Devroye, L. 1986. *Non-Uniform Random Variate Generation.* Spinger-Verlag.

[13] French, K. R., G. W. Schwert, and R. F. Stambaugh. 1987. Expected stock returns and volatility. *Journal of Financial Economics 19, 3-29.*

[14] Gatheral, J., and R. C. A. Oomen. 2007. Zero-intelligence realized variance estimation. *Finance Stochastics* 14, 249–283.

[15] Ghysels, E., P. Mykland, and E. Renault. 2007. In-sample asymptotics and across-sample efficiency gains for high frequency data statistics. *Working paper*, University of North Carolina - Chapel hill.

[16] Gloter, A., and J. Jacod. 2001. Diffusions with Measurement Errors I, Local Asymptotic Normality. ESAIM: Probability and Statistics 5, p225-242.

[17] Hansen, B. E. 2007. Least squares model averaging. *Econometrica* 75, 1175-1189.

[18] Hansen, P.R., and A. Lunde. 2006. Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24, No. 2.

[19] Hasbrouck, J. 1993. Assessing the quality of a security market: a new approach to transaction cost measurement. *The Review of Financial Studies* 6, No. 1, 191-212.

[20] Heston, S. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 1993, Vol. 6, number 2, 327-343.

[21] Hill, B.M. 1975. A simple general approach to inference about the tail of a distribution. *Annals of Statistics.* 13, 331-341.

[22] Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter. 2009. Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and their Applications,* 119 2249–2276.

[23] Jacod, J., and P. Protter. 1998. Asymptotic error distribution for the Euler method for stochastic Differential Equations. *Annals of Probability* 26, 267-307.

[24] Merton, R. C. 1980. On estimating the expected return on the market: an exploratory investigation. *Journal of Financial Economics* 8, 323-361.

[25] Newey, W., and K. West. 1987. A simple positive definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.

[26] Podolskij, M., and M. Vetter. 2009. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. Bernoulli 15(3), 634-658.

[27] Politis, D. N., J. P. Romano, and M. Wolf. 1997. Subsampling for heteroskedastic time series. Journal of Econometrics 81, 281-317.

[28] Stoll, H. R. 1989. Infering the components of the bid-ask spread: theory and empirical evidence. *Journal of Finance* 44, 115-134.

[29] Ubukata, O., and K. Oya. 2009. Estimation and testing for dependence in market microstructure noise. *Journal of Financial Econometrics*, 2008, 1-46. DOI: 10.1093/jjfinec/nbn021.

[30] Zhang, L, P.A. Mykland, and Y. Ait-Sahalia. 2005. A tale of two time scale: determining integrated volatility with noisy high-frequency data. *Journal of The American Statistical Association* 100, No 472.

[31] Zhou, B. 1996. High frequency data and volatility in foreign-exchange rates. *Journal of Business and Economics Statistics* 14, 45-52.

# Notes

[1]See Andersen, Bollerslev, Diebold and Labys (2003); Bandi and Russell (2008).

[2]See also Jacod, Li, Mykland, Podolskij and Vetter (2009).

[3]BNHLS (2008a, Section 5.6) show that the contribution of jumps may not disturb the asymptotic orders.

[4]A popular model often postulated for the spot variance is the square-root diffusion $d\sigma_s^2 = \kappa\left(\theta - \sigma_s^2\right) ds + \delta\sqrt{\sigma_s^2} dB_s$. Under this model, the spot volatility follows the diffusion $d\sigma_s = f\left(\sigma_s\right) ds + g\left(\sigma_s\right) dB_s$, where $f\left(\sigma_s\right) = \frac{1}{2\sigma_s}\left[\kappa\theta - \frac{\delta^2}{4} - \kappa\sigma_s^2\right]$ and $g\left(\sigma_s\right) = \frac{\delta}{2}$. In this case, the function $p_{(1)}^*\left(\sigma\right)$ solves $\frac{\partial p_{(1)}^*(\sigma)}{\partial\sigma} = \frac{\rho\sigma}{\frac{\delta}{2}}$, which yields $p_{(1)}^*\left(\sigma_s\right) = \left(\frac{\rho}{\delta}\right)\sigma_s^2$.

[5]This situation may apply either to a single asset with a market activity that varies across days (time series dimension) or to several assets with different liquidity level (cross-sectional dimension).

[6]For instance, $\phi_m \simeq Dm^{-\alpha}$ or $\phi_m \simeq 1 - Dm^{-\alpha}$ as $m \to \infty$, $D > 0$.

[7]For instance, if $r_{t,j}^*$ is a series of one minute returns, then $\widetilde{r}_{t,k}$ would be a $q$ minutes return.

[8]The dependence lag $L$ must be estimated before $RV_t^{(AC,m,L+1)}$ can be feasible. Likewise, an estimate of $\phi_m$ must be available before $RV_t^{(AC,m,\infty)}$ can be implemented.

[9]The noise parameters $\beta_0$, $\beta_1$, $\rho$, $\alpha$ and $\delta$ are not empirically relevant for the realized kernels and thus, they are not estimated.

[10]Note that $O(m^{-1})$ end effects are neglected so that Equation (20) is an *asymptotic* moment condition.

[11]Note that the noise parameters are not necessarily constant across $m$ in the current framework. Hence, it is important to read the central limit result (31) for a fixed $m$.

[12]A heuristic for chosing $L_{\max}$ is proposed in the next section.

[13]We assume that the market opens from 9:30 am to 4:00 pm, which implies 4680 discretization points within each day.

[14]Note that $\phi_m = 0$ features an IID exogenous noise.

[15]One should expect the sign of the bias of $\widehat{\omega}_0$ to depend on the sign of the correlation between the endogenous noise and the spot volatility. It may also depends on the sample size because the volatility path is fixed throughout the Monte Carlo replications.

[16]The data we use in this paper have been purchased from a private provider who has ensured its accuracy by comparison with three other independent financial data providers.