2011s-29

Shrinkage Realized Kernels

Marine Carrasco, Rachidi Kotchoni

Série Scientifique
Scientific Series

Montréal
Février 2011

CIRANO
Allier savoir et décision

Centre interuniversitaire de recherche en analyse des organisations

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du Ministère du Développement économique et régional et de la Recherche, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère du Développement économique et régional et de la Recherche, and grants and research mandates obtained by its research teams.

Les partenaires du CIRANO

Partenaire majeur
Ministère du Développement économique, de l'Innovation et de l'Exportation

Partenaires corporatifs
Banque de développement du Canada
Banque du Canada
Banque Laurentienne du Canada
Banque Nationale du Canada
Banque Royale du Canada
Banque Scotia
Bell Canada
BMO Groupe financier
Caisse de dépôt et placement du Québec
Fédération des caisses Desjardins du Québec
Financière Sun Life, Québec
Gaz Métro
Hydro-Québec
Industrie Canada
Investissements PSP
Ministère des Finances du Québec
Power Corporation du Canada
Raymond Chabot Grant Thornton
Rio Tinto
State Street Global Advisors
Transat A.T.
Ville de Montréal

Partenaires universitaires
École Polytechnique de Montréal
HEC Montréal
McGill University
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web.

ISSN 1198-8177

Partenaire financier

Développement économique, Innovation et Exportation
Québec

Shrinkage Realized Kernels

Marine Carrasco [*], *Rachidi Kotchoni* [†]

Résumé / *Abstract*

Nous construisons un estimateur de volatilité intégrée qui se présente sous la forme d'une combinaison linéaire optimale d'autres estimateurs, dans le cadre d'un modèle semi-paramétrique de type moyenne mobile postulé pour le bruit de microstructure. L'ordre de ce processus moyen mobile est une fonction croissante de la fréquence des observations, ce qui implique que l'autocorrélation d'ordre 1 du bruit de microstructure tend vers l'unité lorsque la fréquence tend vers l'infini. Des estimateurs sont proposés pour les paramètres identifiables du modèle et leurs bonnes propriétés sont confirmées par simulation. Les résultats d'une application empirique basée sur des actifs du DJI suggèrent qu'en général, l'ordre du processus moyen mobile postulé pour le bruit de microstructure augmente moins vite que la racine carrée de la fréquence des observations.

Mots clés : Volatilité intégrée, méthode des moments, bruit de microstructure, estimateur à noyaux réalisés, combinaison linéaire optimale d'estimateurs.

A shrinkage estimator of the integrated volatility is derived within a semiparametric moving average microstructure noise model specified at the highest frequency. The order the moving average is allowed to increase with the sampling frequency, which implies that the first order autocorrelation of the noise converges to one as the sampling frequency goes to infinity. Estimators are derived for the identifiable parameters of the model and their good properties are confirmed in simulation. The results of an empirical application with stocks listed in the DJI suggest that the order of the moving average model postulated for the noise typically increases slower than the square root of the sampling frequency.

Keywords: *Integrated Volatility, method of moment, microstructure noise, realized kernel, shrinkage.*

[*] Université de Montreal - CIRANO - CIREQ. E-mail: marine.carrasco@umontreal.ca

[†] CRÉA, Université Laval. Correspondance: Pavillon Paul-Comtois, 2425, rue de l'Agriculture, Local#4424-F, Québec (QC) G1V 0A6. Phone: 418 656 2131 poste 3940. Fax: 418 656 7821.
E-mail: rachidi.kotchoni@fsaa.ulaval.ca.

1 Introduction

To estimate the monthly variance of a financial asset, 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 support their 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 (1994), 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 or jumps.

However, it is commonly admitted that recorded stock prices are contaminated with "market microstructure noise" (henceforth "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 be 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 very high frequency data to be a biased estimator of the object of interest. The sources of noise are discussed for example in Stoll (1989, 2000) or Hasbrouck (1993,1996). 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 this curse. One of them consists in choosing in an ad-hoc manner a moderate sampling frequency at which the impact of the noise is sufficiently mitigated[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 the day, Ait-Sahalia, Mykland and Zhang (2005) derive a maximum likelihood estimator of the IV that is robust to both IID noise and distributional mispecification. 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 (2006) to deal with dependent noise. Since then, other consistent estimators have become available among which the realized kernels of Barndorff-Nielsen, Hansen, Lunde and Shephard (2008a) and the pre-averaging estimator of Podolskij and Vetter (2006)[2]. An alternative line of research pursued by Corradi, Distaso and Swanson (2008) advocates the nonparametric estimation of the predictive density and confidence intervals for the IV rather than focusing on point estimates.

In a simulation study, Gatheral and Oomen (2007) find that consistent estimators often perform poorly at the sampling frequencies commonly encountered in practice. Our simulations of Section 6 show that this finding strongly depends on the size of the variance of the microstructure noise relative to the discretization error. In fact, the inconsistent estimator tends to perform better than the consistent one only when the variance of the microstructure noise is small. We also note that even when the variance of the inconsistent estimator is higher, it can still be optimally combined with the consistent estimator to obtain a new one that performs better than both. The weight of the linear combination is selected in order to minimize the variance and the resulting estimator is termed “shrinkage estimator”.

However, an unbiased estimator of the IV must be designed in accordance with the dependence properties of the noise. This leads us to propose a model for the microstructure noise that depart from the usual IID assumption. More precisely, we specify at the highest frequency a parsimonious relation between the microstructure noise on the one side, and the efficient return and the latent volatility process on the other side. We assume a general and flexible type of noise that includes an independent endogenous part ε_t^* and an L-dependent exogenous part ε_t, with the autocovariance structure of ε_t depending on the highest frequency m at which the data are recorded. It is assumed that the maximum lag at which the autocorrelation of ε_t dies out is increasing in m when measured in number of observations, while this lag goes to zero when measured in calendar time. The latter assumption has the implication that the first order autocorrelation of ε_t goes to one as m goes to infinity, contrary to what would imply an $AR(1)$ with constant autoregressive root. We provide an intuitive economic interpretation of this implication of our model.

We derive the properties of common realized measures under the new model. We find that the realized kernels of Barndorff-Nielsen and al. (2008a) is still delivering its best performance at the highest frequency, but its variance converges to a quantity of similar order of magnitude as the variance of the microstructure noise. While this quantity can be arbitrary small and negligible, it does not converge to zero. This suggests that a variance reduction technique can be useful if the noise displays the particular type of dependence assumed in our model. We propose to linearly combine the standard realized kernels of Barndorff-Nielsen and al. (2008a) with an alternative unbiased kernel estimator. The resulting estimator is termed “shrinkage realized kernels”, as it shares some feature with the Stein (1956) estimator and other model averaging techniques. Finally, a method-of-moment approach is proposed to estimate the correlogram of the exogenous noise. We illustrate by simulation the good performance of the various estimators proposed in the paper. An empirical application based on fifteen stocks listed in the Dow Jones Industrials shows evidences of correlation in the noise process and between the noise and the latent returns. If our model for the noise is true, the empirical results suggest that the memory parameter L grows slower than $\sqrt{m}$ in general.

The rest of the paper is organized as follows. The next section presents our assumptions on the frictionless price and our model for the microstructure noise. In section 2, we study the properties of three standard IV estimators in light our theoretical framework. In Section 3, we present and discuss the properties of a kernel type shrinkage estimator for the IV when the noise is L-dependent. Inference procedures about the noise parameters are presented in section 4. Sections 5 and 6 present respectively a simulation study and an empirical application based on twelve stocks listed in the Dow Jones Industrials. Section 7 concludes. The mathematical proofs are left in appendix.

2 The Framework

Firstly, we present a standard model for the efficient price that allows for leverage effect. Next, we argue that our analysis can be performed by ignoring the leverage effect and jumps with no loss of generality. Finally, we present our model for the microstructure noise.

2.1 A General Model for the Efficient Price

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

$$dp_s^* = \mu_s ds + \sigma_s dW_s; \quad p_0^* = 0, \tag{1}$$

where μ_s is the drift function, σ_s is the spot volatility and W_s is a standard Brownian motion. We assume that the volatility process $\{\sigma_s\}_{s\geq 0}$ is *càdlàg*, implying that all powers of the volatility process are locally integrable with respect to the Lebesgue Measure[3]. The drift function μ_s is smooth and adapted to the filtration generated by $\{W_u, \sigma_u, u < s\}$. In turn, the spot volatility obeys a stochastic differential equation of the following form:

$$d\sigma_s = f_s ds + g_s dB_s,$$

where f_s and g_s are adapted to the filtration generated by $\{W_u, B_u, u < s\}$ and f_s is smooth. We allow for leverage effect by assuming that:

$$E\left(dW_s, dB_s\right) = \rho ds.$$

Without loss of generality, we condition all our analysis on the volatility path but the conditioning is removed from the notations for simplicity. Unless otherwise mentioned, all expectations, variances and covariances are conditioned on $\{\sigma_s\}_{s\geq 0}$. Accordingly, all deterministic transformations of the volatility process are treated as constant objects. In particular, the integrated volatilities $IV_t = \int_{t-1}^{t} \sigma_s^2 ds$, $t = 1, 2, 3, ...T$ are fixed parameters that we aim to estimate. By definition, the microstructure noise equals $u_s = p_s - p_s^*$, that is, the difference between the observed log-price and the efficient log-price. Let r_t^* denote the latent log-return at period t, and r_t its observable counterpart. We consider a sampling scheme where the unit period is normalized to one in calendar time. Under the above conditions, the daily return is:

$$\begin{aligned} r_t &\equiv p_t - p_{t-1} = r_t^* + u_t - u_{t-1}, \text{ and} && (2)\\ r_t^* &= \int_{t-1}^{t} \mu_s ds + \int_{t-1}^{t} \sigma_s dW_s. && (3) \end{aligned}$$

Suppose that we have access to a large number m of intra-period returns $r_{t,1}, r_{t,2}, ..., r_{t,m}$, where $t = 1, ..., T$ are the period labels, m is the number of recorded prices in each period and $r_{t,j}$ is the j^{th} observed return during the period $[t-1,\ t]$. In the sequel, we use the expression "record frequency" to refer to the frequency m at which the data has been recorded. For simplicity, we assume that the high frequency observations are equidistant in calendar time. The j^{th} high frequency observed return within day t is given by:

$$r_{t,j} = r_{t,j}^* + u_{t,j} - u_{t,j-1},$$

where:

$$\begin{aligned} r_{t,j}^* &\equiv p_{t-1+j/m}^* - p_{t-1+(j-1)/m}^* = \int_{t-1+(j-1)/m}^{t-1+j/m} \sigma_s dW_s, \text{ and}\\ u_{t,j} &\equiv u_{t-1+j/m}. \end{aligned}$$

The noise-contaminated (observed) and true realized volatility (latent) computed at frequency m are:

$$RV_t^{(m)} = \sum_{j=1}^{m} r_{t,j}^2 \text{ and } RV_t^{*(m)} = \sum_{j=1}^{m} r_{t,j}^{*2}. \quad (4)$$

In the absence of leverage effect ($\rho = 0$), Barndorff-Nielsen and Shephard (2002) show that

$RV_t^{*(m)}$ converges to IV_t and derived the asymptotic distribution:

$$\frac{RV_t^{*(m)} - IV_t}{\sqrt{\frac{2}{3}\sum_{j=1}^{m} r_{t,j}^{*4}}} \to N(0,1), \tag{5}$$

as m goes to infinity. Meddahi (2002) studied the *finite frequency* behavior of the discretization error $RV_t^{*(m)} - IV_t$ with a focus on the specific case where the true model belongs to the Eigenfunction Stochastic Volatility family. Gonçalves and Meddahi (2009) proposed some bootstrap procedures as alternative inference tools to analyze the asymptotic behavior of realized measures. In both papers, no microstructure noise is assumed. In the presence of microstructure noise, $RV_t^{*(m)}$ is not feasible.

2.2 Simplifying the Model for the Efficient Price

Here we argue that the leverage effect (and jumps) may be ignored for the purpose of our analysis. In fact, the expression of the high frequency efficient return is given by:

$$r_{t,j}^* = \int_{t-1+(j-1)/m}^{t-1+j/m} \mu_s ds + \int_{t-1+(j-1)/m}^{t-1+j/m} \sigma_s dW_s.$$

By adding and streaking $\mu_{t-1+(j-1)/m}$ and $\sigma_{t-1+(j-1)/m}$ to the drift and volatility respectively, we obtain:

$$\begin{aligned} r_{t,j}^* \quad = \quad & \mu_{t-1+(j-1)/m} \int_{t-1+(j-1)/m}^{t-1+j/m} ds + \sigma_{t-1+(j-1)/m} \int_{t-1+(j-1)/m}^{t-1+j/m} dW_s \\ & + \int_{t-1+(j-1)/m}^{t-1+j/m} \left(\mu_s - \mu_{t-1+(j-1)/m}\right) ds + \int_{t-1+(j-1)/m}^{t-1+j/m} \left(\sigma_s - \sigma_{t-1+(j-1)/m}\right) dW_s. \end{aligned}$$

For the first term, we have:

$$\mu_{t-1+(j-1)/m} \int_{t-1+(j-1)/m}^{t-1+j/m} = \frac{\mu_{t-1+(j-1)/m}}{m} \tag{6}$$

The second term satisfies:

$$\sigma_{t-1+(j-1)/m} \int_{t-1+(j-1)/m}^{t-1+j/m} dW_s = \sigma_{t-1+(j-1)/m} \left(W_{t-1+j/m} - W_{t-1+(j-1)/m}\right) \tag{7}$$

For the third term, we use the fact that μ_s is smooth by assumption so that $\mu_s - \mu_{t-1+(j-1)/m} = O\left(\frac{1}{m}\right)$. We have:

$$\int_{t-1+(j-1)/m}^{t-1+j/m} \left(\mu_s - \mu_{t-1+(j-1)/m}\right) ds = O\left(m^{-2}\right), \tag{8}$$

For sufficiently large m, we have the following Euler-type approximation for $\sigma_s - \sigma_{t-1+(j-1)/m}$:

$$\begin{aligned} \sigma_s - \sigma_{t-1+(j-1)/m} \quad \simeq \quad & f_{t-1+(j-1)/m} \left(s - t + 1 - \frac{j-1}{m}\right) \\ & + g_{t-1+(j-1)/m} \left(B_s - B_{t-1+(j-1)/m}\right), \end{aligned}$$

for $s \in \left[t-1+\frac{j-1}{m}, t-1+\frac{j}{m}\right]$. Replacing this into the fourth term yields:

$$\begin{aligned}
&\int_{t-1+(j-1)/m}^{t-1+j/m} \left(\sigma_s - \sigma_{t-1+(j-1)/m}\right) dW_s \\
\simeq \quad & f_{t-1+(j-1)/m} \int_{t-1+(j-1)/m}^{t-1+j/m} \left(s - t + 1 - \frac{j-1}{m}\right) dW_s \\
& + g_{t-1+(j-1)/m} \int_{t-1+(j-1)/m}^{t-1+j/m} \left(B_s - B_{t-1+(j-1)/m}\right) dW_s.
\end{aligned}$$

Next, we use the following type of approximation:

$$\int_{t-1+(j-1)/m}^{t-1+j/m} \phi(s)\, dW_s \simeq \left[\phi(t-1+j/m) - \phi(t-1+(j-1)/m)\right] \left[W_{t-1+j/m} - W_{t-1+(j-1)/m}\right]$$

This leads to

$$\begin{aligned}
&\int_{t-1+(j-1)/m}^{t-1+j/m} \left(\sigma_s - \sigma_{t-1+(j-1)/m}\right) dW_s \\
\simeq \quad & \frac{f_{t-1+(j-1)/m}}{m} \left(W_{t-1+j/m} - W_{t-1+(j-1)/m}\right) \\
& + g_{t-1+(j-1)/m} \left(B_{t-1+j/m} - B_{t-1+(j-1)/m}\right) \left(W_{t-1+j/m} - W_{t-1+(j-1)/m}\right)
\end{aligned}$$

The leverage effect assumption implies:

$$\left(B_{t-1+j/m} - B_{t-1+(j-1)/m}\right) \left(W_{t-1+j/m} - W_{t-1+(j-1)/m}\right) \simeq \frac{\rho}{m}$$

so that finally, we obtain for the fourth term:

$$\begin{aligned}
&\int_{t-1+(j-1)/m}^{t-1+j/m} \left(\sigma_s - \sigma_{t-1+(j-1)/m}\right) dW_s \simeq \\
&\frac{f_{t-1+(j-1)/m}}{m} \left(W_{t-1+j/m} - W_{t-1+(j-1)/m}\right) + \frac{\rho g_{t-1+(j-1)/m}}{m}
\end{aligned} \tag{9}$$

The approximation of the high frequency return is obtained by taking the sum of (6), (7), (8) and (9. We see that the term with dominant variance in the high frequency return is given by (7). The dominant term above does not depend on the functions μ_s, f_s, g_s nor on the leverage effect parameter ρ. This shows that in the presence of leverage effect, the efficient log-price may be simply treated as if it were a semi-martingale with the additional drift term ρg_s. Hence without loss of generality and for sake of parsimony, we will assume in subsequent developments that:

$$\mu_s = \rho = 0,$$

or equivalently, that:

$$dp_s^* = \sigma_s dW_s; \quad p_0^* = 0,$$

where σ_s is independent of W_s.

It is maintained throughout the paper that there is no jump in the efficient price. However, our analysis of the microstructure noise remains valid if jumps that are uncorrelated with all other randomness are present in the model. In this case, the estimators we consider for the IV is now designed for the quadratic variation. Separating the IV from the contribution of the jumps in the quadratic variation would then be the relevant issue in practice.

2.3 A Semiparametric Model for the Microstructure Noise

Our approach to model the noise is based on the assumption that the time series properties of the microstructure noise may depend on the frequency at which the prices have been recorded. With this in mind, we specify a link between the noise $u_{t,j}$ and the latent return $r^*_{t,j}$ at the highest frequency and then deduce the properties of the realized volatility computed at lower frequencies. We assume that the noise process evolves in calendar time according to:

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

where $a_{t,j}$ is a deterministic but time varying coefficient and $\varepsilon_{t,j}$ is independent of the efficient high frequency return $r^*_{t,j}$. In the words of Hasbrouck (1993), $\varepsilon_{t,j}$ is the information uncorrelated or exogenous pricing error and $a_{t,j} r^*_{t,j}$ is the information correlated or endogenous pricing error. For sake of parsimony, our model assumes that time dependence in the noise process can only be due to its information uncorrelated part.

The following assumptions are maintained throughout the paper.

Assumption E0. $a_{t,j} = \beta_0 + \frac{\beta_1}{\sqrt{m\sigma^{*2}_{t,j}}}$, where β_0 and β_1 are constants and

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

Assumption E1. For fixed m, we have:

E1(a) The process $\varepsilon_{t,j}$ is discrete time stationary with zero mean, and independent of $\{\sigma_s\}$ and $r^*_{t,j}$.

E1(b) $E\left|u_{t,j}u_{t,j-h}\right|^{4+\epsilon} < \infty$, for some $\epsilon > 0$, for all h.

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

Assumption E3. $\omega(0) \equiv \omega_{m,0} = \omega_0$ for all m, $\omega_{m,h} - \omega_{m,h+1} = \omega_0 O(m^{-\alpha})$ for some $\alpha > 0$, $h = 0, ..., L-1$.

Assumption E4. $L \propto m^\delta$ for some δ, and $0 \leq \delta \leq \min(\alpha, 2/3)$, where the notation "$\propto$" means "proportional to".

Assumption E5. For fixed m, $Var\left(n^{-1/2}m^{-1/2}\sum_{t=t'+1}^{t'+n}\sum_{j=1}^{m} r_{t,j}r_{t,j-h}\right) \to q_h$, uniformly in any t', as $n \to \infty$, where $r_{t,j} = r^*_{t,j} + u_{t,j} - u_{t,j-1}$ is the observed return.

Assumption E0 is aimed at introducing endogeneity in the microstructure noise process in such a way that both homoscedasticity ($\beta_0 = 0$) and heteroscedaticity ($\beta_1 = 0$) are allowed. This assumption implies that the variance of the endogenous part of the noise goes to zero at rate m since:

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

Assumption E1(a) is quite standard in the literature. Assumption E1(b) is stronger than needed to show the finiteness of the variance of the IV estimators. It is used in conjunction with E2 and E5 to derive an asymptotic theory for the estimators of the autocovariances of $\varepsilon_{t,j}$ in Section 4.

The semi-parametric nature of the microstructure noise model comes from Assumption E2 which only stipulates that $\varepsilon_{t,j}$ is L-dependent without specifying a parametric family for the distribution of $\varepsilon_{t,j}$. Hansen and Lunde (2006) construct a Haussman-type test to detect time dependence in the noise process. After applying their test to real data, they concluded that the noise process is time dependent, correlated with latent return, and possibly heteroscedastic. More recently, Ubukata and Oya (2009) proposed some procedures to test for dependence in the noise process with irregularly spaced and asynchronous bivariate data.

Assumption E3 implies that:

$$Cov\left(\varepsilon_{t,0}, \varepsilon_{t,j}\right) - \omega_0 = -\sum_{h=0}^{j-1}\left(\omega_{m,h} - \omega_{m,h+1}\right) = O(jm^{-\alpha}). \tag{11}$$

Hence for any fixed j, $Cov\left(\varepsilon_{t,0}, \varepsilon_{t,j}\right)$ converges to the constant variance ω_0 as m goes to infinity.

If $\alpha = 0$, then Assumption E4 implies that $\delta = 0$ so that $\varepsilon_{t,j}$ is an $MA(L)$ process with fixed L. More generally, L may grow with the record frequency m. In this case, if $j = \lfloor Lc \rfloor$ for some constant $c \in (0,1)$ (where $\lfloor x \rfloor$ denote the largest integer that is smaller than x), then we have:

$$Cov\left(\varepsilon_{t,0}, \varepsilon_{t,j}\right) - \omega_0 = O(m^{\delta-\alpha}). \tag{12}$$

We see that $\delta \leq \alpha$ as assumed in E4 is a necessary condition for $Cov\left(\varepsilon_{t,0}, \varepsilon_{t,j}\right)$ to be bounded. Also, the requirement that $\delta < 2/3$ is needed for the convergence of the realized kernels with Bartlett kernel. The lag L is longer for larger δ, but the time length $\frac{L}{m}$ after which the correlation dies out converges to zero as m goes to infinity.

Finally, Assumption E5 is analogue to the Assumption 2 of Ubukata and Oya (2009) and is needed for the central limit theorem of Politis, Romano and Wolf (1997) to obtain. This assumption is likely to be satisfied if the volatility increment process $\sigma_{t,j}^{*2}$ is stationary and mixing.

In summary, the proposed model has the implication that the first order autocorrelation of ε_t converges to one as m goes to infinity. This implication follows from the previous assumptions and should not be considered as an assumption in its own. Interestingly, this implication has an intuitive economic interpretation. In fact, transaction decisions are made by agents based on the information flow to which they have access to. For the econometrician, the information held by agents is latent but has observable consequences (including, but not limited to the bid-ask spread). As a deviation from the frictionless equilibrium price, the microstructure noise certainly incorporate the quality of the aggregate information process that drives the market. We may thus formally view the microstructure noise as being a function of this information process. If the information flow varies smoothly through time, we can reasonably expect two consecutive realizations of the noise to be correlated, and the closer the two realizations are in calendar time the higher the correlation is. Sudden and large variations of the information flow translate in jumps in the efficient price and are unlikely to go unnoticed. This interpretation implies that ε_t is generated endogenously by the aggregated trade flow even though independent of the efficient price process.

Imposing $\beta_0 = \beta_1 = \delta = 0$ in our model leads to $u_{t,j} = \varepsilon_{t,j}$ where $\varepsilon_{t,j}$ is a moving average model of fix order L for $u_{t,j}$. This case has been considered in Hansen and Lunde (2006). Further imposing $L = 0$ results in an uncorrelated noise, such as the IID noise considered by Ait-Sahalia, Mykland and Zhang (2005). One gets a version of Roll's model (1984) from our specification by setting $\beta_0 = \beta_1 = 0$ and $\varepsilon_{t,j} = \pm Q_{t,j}/2$, where $Q_{t,j}$ is the bid-ask spread. The model of Roll can thus be regarded as nested within our specification with the difference that $\varepsilon_{t,j}$ is now observable. Hasbrouck (1993) used the restriction $\beta_1 = 0$ with $\varepsilon_{t,j}$ IID to model the microstructure noise contaminating daily returns. This particular case results in an $MA(1)$ representation for $u_{t,j}$ which, as a function of the original parameters, is identifiable if one further imposes the restriction $\varepsilon_{t,j} = 0$ used in Beveridge and Nelson (1981) or the restriction $\beta_0 = 0$ used by Watson (1986). Ait-Sahalia, Mykland and Zhang (2006) considered an exogenous noise with general mixing properties. Kalnina and Linton (2008) advocated a microstructure noise model that features endogeneity and diurnal heteroscedaticity. These two models cannot be nested within our specification.

3 Properties of Three IV Estimators

We study successively the traditional realized variance, the kernel estimator of Hansen and Lunde (2006) and the realized kernels of Barndorff-Nielsen, Hansen, Lunde and Shephard (2008a). All three estimators admit the following decomposition:

$$\widehat{IV}_t = f_{r^*}\left(\left\{r^*_{t,j}\right\}_{j=1}^m\right) + f_{r^*,u}\left(\left\{r^*_{t,j}, u_{t,j}\right\}_{j=1}^m\right) + f_u\left(\left\{u_{t,j}\right\}_{j=1}^m\right), \tag{13}$$

where

$$E\left[f_{r^*}\left(\left\{r^*_{t,j}\right\}_{j=1}^m\right)\right] = IV_t, \tag{14}$$

$$E\left[f_{r^*,u}\left(\left\{r^*_{t,j}, u_{t,j}\right\}_{j=1}^m\right)\right] = 0, \tag{15}$$

and

$$f_{r^*,u}\left(\left\{r^*_{t,j}, 0\right\}_{j=1}^m\right) = f_u\left(\left\{0\right\}_{j=1}^m\right) = 0, \tag{16}$$

and the three terms in (13) are uncorrelated. This decomposition will be used in Section 4 to enhance our arguments in favor of a shrinkage estimator for the IV

3.1 The Realized Volatility

The realized volatility $RV_t^{(m)}$ sampled at the highest frequency satisfies (13) with:

$$\begin{aligned}
f_{r^*}\left(\left\{r^*_{t,j}\right\}_{j=1}^m\right) &= \sum_{j=1}^m r^{*2}_{t,j}, \\
f_{r^*,u}\left(\left\{r^*_{t,j}, u_{t,j}\right\}_{j=1}^m\right) &= 2\sum_{j=1}^m \left(u_{t,j} - u_{t,j-1}\right) r^*_{t,j} \text{ and} \\
f_u\left(\left\{u_{t,j}\right\}_{j=1}^m\right) &= \sum_{j=1}^m \left(u_{t,j} - u_{t,j-1}\right)^2.
\end{aligned}$$

Under IID noise, $RV_t^{(m)}$ is biased and inconsistent and its bias and variance are linearly increasing in m. See for example Zhang, Mykland and Ait-Sahalia (2005) and Hansen and Lunde (2006). Here the estimator of interest is the sparsely sampled realized variance given by:

$$RV_t^{(m_q)} = \sum_{k=1}^{m_q} \widetilde{r}^2_{t,k}, \tag{17}$$

where $\widetilde{r}_{t,k}$ is the sum of q consecutive returns, that is:

$$\widetilde{r}_{t,k} = \sum_{j=qk-q+1}^{qk} r_{t,j}, k = 1, ..., m_q = \frac{m}{q}, q \geq 1. \tag{18}$$

Hence if $r^*_{t,j}$ is a series of one minute returns for instance, then $\widetilde{r}_{t,k}$ would be a sequence of q minutes return. Figure 1 illustrates the corresponding subsampling scheme which is quite standard in this literature.

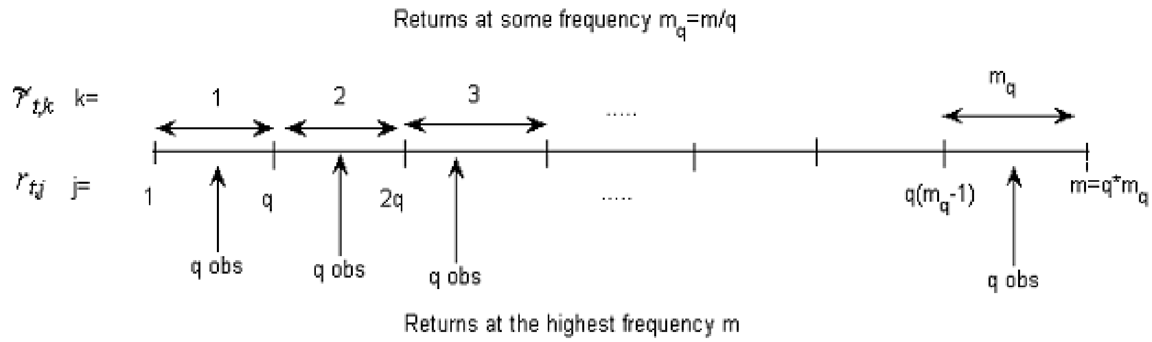

Figure 1: The subsampling scheme.

If the noise process is correctly described at the highest frequency by equation (10), then the expression of $\widetilde{r}_{t,k}$ is given by:

$$\begin{aligned}\widetilde{r}_{t,k} &= \left(1+\beta_0+\frac{\beta_1}{\sigma^*_{t,qk}}\right) r^*_{t,qk} + \sum_{j=qk-q+1}^{qk-1} r^*_{t,j} - \left(\beta_0+\frac{\beta_1}{\sigma^*_{t,qk-q}}\right) r^*_{t,qk-q} \\ &\quad + \left(\varepsilon_{t,qk}-\varepsilon_{t,qk-q}\right),\end{aligned} \tag{19}$$

for $k=1,...,m_q$ and for all t, with the convention that $\sum_{j=qk-q+1}^{qk-1} r^*_{t,j} = 0$ when $q=1$. The covariance between $\widetilde{r}_{t,k}$ and $\widetilde{r}_{t,k-1}$ is given by:

$$\begin{aligned} cov(\widetilde{r}_{t,k},\widetilde{r}_{t,k-1}) &= -\left(\beta_0+\frac{\beta_1}{\sigma^*_{t,qk-q}}\right)\left(1+\beta_0+\frac{\beta_1}{\sigma^*_{t,qk-q}}\right)\sigma^{*2}_{t,qk-q} \\ &\quad -\omega_0+2\omega_{m,q}-\omega_{m,2q}. \end{aligned} \tag{20}$$

The next theorem gives the bias and variance of $RV_t^{(m_q)}$. The expression of the bias will be useful for the estimation of the correlogram of the microstructure noise in Section 4.

Theorem 1 *Assume that the noise process evolves according to equation (10), and let* $RV_t^{(m_q)} = \sum_{k=1}^{m_q} \widetilde{r}^2_{t,k}$ *with* $m_q = \frac{m}{q}$, $q \geq 1$ *and* m *the record frequency. Then we have:*

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

$$
\begin{aligned}
&Var\left[RV_t^{(m_q)}\right] = m_q\kappa + \tfrac{16\beta_1^2}{q}\left(\omega_0-\omega_{m,q}\right) + \tfrac{12\beta_1^4}{qm}\\
&\quad +8\left[\tfrac{(3+5\beta_0)\beta_1^3}{m\sqrt{m}} + \tfrac{2(1+\beta_0)\beta_1}{\sqrt{m}} + \tfrac{2\beta_0\beta_1}{\sqrt{m}}\left(\omega_0-\omega_{m,q}\right) + \tfrac{\beta_0\beta_1^3}{m\sqrt{m}}\right]\textstyle\sum_{k=1}^{m_q}\sigma_{t,qk}^{*}\\
&\quad +4\left(1+2\beta_0+2\beta_0^2\right)\left[\tfrac{7\left(1+2\beta_0+2\beta_0^2\right)\beta_1^2}{m} + 2\left(\omega_0-\omega_{m,q}\right)\right]\textstyle\sum_{k=1}^{m_q}\sigma_{t,qk}^{*2}\\
&\quad +8\tfrac{\left(1+4\beta_0+6\beta_0^2+4\beta_0^3\right)\beta_1}{\sqrt{m}}\textstyle\sum_{k=1}^{m_q}\sigma_{t,qk}^{3} + 2\sum_{k=1}^{m_q}\left(\sum_{j=qk-q+1}^{qk}\sigma_{t,j}^{*2}\right)^2\\
&\quad +\tfrac{16\beta_0(1+\beta_0)\beta_1^2}{m}\textstyle\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*}\sigma_{t,qk}^{*} + \tfrac{8(1+\beta_0)\beta_1}{\sqrt{m}}\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk}^{*}\\
&\quad +\tfrac{8\beta_0^2(1+\beta_0)\beta_1}{\sqrt{m}}\textstyle\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*2}\sigma_{t,qk}^{*} + \tfrac{8\beta_0(1+\beta_0)^2\beta_1}{\sqrt{m}}\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*}\sigma_{t,qk}^{*2}\\
&\quad +\tfrac{8\beta_0\beta_1}{\sqrt{m}}\textstyle\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk-q}^{*} + 2\left(4\beta_0+8\beta_0^2+8\beta_0^3+4\beta_0^4\right)\sum_{k=1}^{m_q}\sigma_{t,qk}^{4}\\
&\quad +4\left(2\beta_0+\beta_0^2\right)\textstyle\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk}^{*2} + 4\beta_0^2\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk-q}^{*2}\\
&\quad +4\beta_0^2\left(1+\beta_0\right)^2\textstyle\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*2}\sigma_{t,qk}^{*2} + 8\left(\omega_0-\omega_{m,q}\right)\left(\beta_0^2+\tfrac{2\beta_0\beta_1}{\sqrt{m}}\right) + O(m^{-1}),\\
&\textit{where } \kappa = \tfrac{1}{m_q}Var\left[\textstyle\sum_{k=1}^{m_q}\left(\varepsilon_{t,kq}-\varepsilon_{t,kq-q}\right)^2\right].
\end{aligned}
$$

Assumption E1(b) ensures that κ is finite, and computing explicitly its exact expression is not of direct interest in our analysis. Note that the dominant terms of the bias and of the variance of $RV^{(m_q)}$ are $O(m_q)$. In the case where $\varepsilon_{t,j}$ is IID, replacing $\beta_0 = \beta_1 = 0$ in the above expressions yields the result of Lemma 4 of Hansen and Lunde (2006) up to some changes in notations:

$$
\begin{aligned}
E\left[RV_t^{(m_q)}\right] &= IV_t + 2m_q\omega_0, \text{ and} \\
Var\left[RV_t^{(m_q)}\right] &= m_q\kappa + 8\omega_0 IV_t + 2\sum_{k=1}^{m_q}\left(\sum_{j=qk-q+1}^{qk}\sigma_{t,j}^{*2}\right)^2,
\end{aligned} \tag{21}
$$

where $m_q\kappa = 4m_qE\left[\varepsilon_{t,j}^4\right] + 2\left(\omega_0^2 - E\left[\varepsilon_{t,j}^4\right]\right)$ when $\varepsilon_{t,j}$ is IID.

We see that the volatility signature plot may not be able to reveal the presence of the noise in the data if $\varepsilon_{t,j} = 0$, since in this case the bias is equal to:

$$
\frac{2\beta_1^2}{q} + 2\beta_1\left(2\beta_0+1\right)\frac{1}{\sqrt{m}}\sum_{k=1}^{m_q}\sigma_{t,qk}^{*} + 2q\beta_0\left(\beta_0+1\right)\sum_{k=1}^{m_q}\sigma_{t,qk}^{*2} = O(1) \text{ for all } m_q.
$$

Moreover, this bias can be negative at some sampling frequencies provided that $\beta_1 < 0$ or $\beta_0 < 0$. Finally, note that the total bias of the RV sampled at the highest frequency may diverge at a lower rate than m, since:

$$
2m\left(\omega_0 - \omega_{m,1}\right) = O(m^{1-\alpha}).
$$

The bias of the realized provides one of the moment conditions that will be used in Section 5 to estimate the correlogram of the microstructure noise. In the next section, we pursue with the examination of the implication of the microstructure noise model for two kernel-based estimators. This preliminary exercise if a useful step in the process of designing a good shrinkage estimator for the IV.

3.2 Hansen and Lunde (2006)

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

$$RV_t^{(AC,m,L+1)} = \sum_{j=1}^{m} r_{t,j}^2 + \sum_{h=1}^{L+1} \sum_{j=1}^{m} r_{t,j} \left(r_{t,j+h} + r_{t,j-h} \right), \tag{22}$$

where L is the memory of the noise as defined in E2. Note that 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):

$$RV_t^{(AC,m,1)} = \sum_{j=1}^{m} r_{t,j}^2 + 2 \sum_{j=1}^{m} r_{t,j} r_{t,j-1} + \underbrace{\left(r_{t,m+1} r_{t,m} - r_{t,1} r_{t,0} \right)}_{\text{end effects}}. \tag{23}$$

Note that $RV_t^{(AC,m,1)}$ satisfies (13) with:

$$\begin{aligned}
f_{r^*}\left(\left\{ r_{t,j}^* \right\}_{j=1}^{m} \right) &= \sum_{j=1}^{m} r_{t,j}^{*2} + \sum_{j=1}^{m} r_{t,j}^* \left(r_{t,j+1}^* + r_{t,j-1}^* \right), \\
f_{r^*,u}\left(\left\{ r_{t,j}^*, u_{t,j} \right\}_{j=1}^{m} \right) &= 2 \sum_{j=1}^{m} \Delta u_{t,j} r_{t,j}^* + \sum_{j=1}^{m} \Delta u_{t,j} \left(r_{t,j+1}^* + r_{t,j-1}^* \right) \\
&\quad + \sum_{j=1}^{m} r_{t,j}^* \left(\Delta u_{t,j+1} + \Delta u_{t,j-1} \right) \text{ and} \\
f_u\left(\left\{ u_{t,j} \right\}_{j=1}^{m} \right) &= \sum_{j=1}^{m} \Delta u_{t,j}^2 + \sum_{j=1}^{m} \Delta u_{t,j} \left(\Delta u_{t,j+1} + \Delta u_{t,j-1} \right),
\end{aligned}$$

where $\Delta u_{t,j} = u_{t,j} - u_{t,j-1}$.

Under IID noise, it is shown in Hansen and Lunde (2006) that $RV_t^{(AC,m,1)}$ is unbiased for IV while its variance is linearly increasing in m. Bandi and Russell (2006) and Hansen and Lunde (2006) derived optimal sampling frequencies for $RV_t^{(m)}$ and $RV_t^{(AC,m,1)}$ based on a signal-to-noise ratio maximization. The variance of $RV_t^{(AC,m,L+1)}$ is hard to derive in the general case. However, assuming that $\varepsilon_{t,j}$ is IID and neglecting the end effects in (23) leads to the following result for $RV_t^{(AC,m,1)}$.

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

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

$$\begin{aligned}
Var\left[RV_t^{(AC,m,1)} \right] &= 8m\omega_0^2 + 2\textstyle\sum_{j=1}^{m} \sigma_{t,j}^{*4} + 2\left(E\left[\varepsilon_{t,j}^4 \right] - \omega_0^2 \right) \\
&+ \tfrac{\beta_1^4 + 6\beta_1^2\omega_0}{m} + \tfrac{8\beta_1^4}{m^2} + \tfrac{8\beta_1}{\sqrt{m}} \left[\tfrac{(\beta_0+1)^2\beta_1^2}{m} + \tfrac{\beta_1}{\sqrt{m}} + 2\omega_0 \left(1 + 2\beta_0 \right) \right] \textstyle\sum_{j=1}^{m} \sigma_{t,j}^* \\
&+ 8\left[\tfrac{\beta_0^2\beta_1^2}{m} + \left(\tfrac{\beta_1^2}{m} + \omega_0 \right) \left(1 + \beta_0 \right)^2 + 2\omega_0\beta_0^2 \right] \textstyle\sum_{j=1}^{m} \sigma_{t,j}^{*2} \\
&+ \tfrac{8\beta_1^2}{m} \left(1 + 2\beta_0 + 2\beta_0^2 \right) \textstyle\sum_{j=1}^{m} \sigma_{t,j}^* \sigma_{t,j-1}^* + \tfrac{16\beta_0\beta_1^2}{m} \left(1 + \beta_0 \right) \textstyle\sum_{j=1}^{m} \sigma_{t,j}^* \sigma_{t,j-2}^* \\
&+ \tfrac{8\beta_0\beta_1}{\sqrt{m}} \left(1 + \beta_0 + \beta_0^3 \right) \textstyle\sum_{j=1}^{m} \sigma_{t,j}^{*2} \sigma_{t,j-1}^* + \tfrac{8\beta_1}{\sqrt{m}} \left(1 + 2\beta_0 + 2\beta_0^2 + \beta_0^3 \right) \textstyle\sum_{j=1}^{m} \sigma_{t,j-1}^{*2} \sigma_{t,j}^* \\
&+ \tfrac{8\beta_1(1+\beta_0)\beta_0^2}{\sqrt{m}} \textstyle\sum_{j=1}^{m} \sigma_{t,j-2}^{*2} \sigma_{t,j}^* + \tfrac{8\beta_1\beta_0(1+\beta_0)^2}{\sqrt{m}} \textstyle\sum_{j=1}^{m} \sigma_{t,j}^{*2} \sigma_{t,j-2}^* \\
&+ 4\left(1 + 2\beta_0 + 3\beta_0^2 + 2\beta_0^3 + \beta_0^4 \right) \textstyle\sum_{j=1}^{m} \sigma_{t,j-1}^{*2} \sigma_{t,j}^{*2}
\end{aligned}$$

$+4\beta_0^2 (1+\beta_0)^2 \sum_{j=1}^{m} \sigma_{t,j}^{*2}\sigma_{t,j-2}^{*2} + \beta_0 O\left(m^{-1/2}\right).$

Replacing $\beta_0 = \beta_1 = 0$ in this theorem yields a known result (Lemma 5 of Hansen and Lunde, 2006):

$$\begin{aligned} E\left[RV_t^{(AC,m,1)}\right] &= IV_t, \\ Var\left[RV_t^{(AC,m,1)}\right] &\simeq 8m\omega_0^2 + 8\omega_0 IV_t - 6\omega_{m,0}^2 + 2\sum_{j=1}^{m}\sigma_{t,j}^{*4} + 4\sum_{j=1}^{m}\sigma_{t,j}^{*2}\sigma_{t,j-1}^{*2}. \end{aligned}$$

When the exogenous noise is absent ($\varepsilon_{t,j} = 0$) and $\beta_0 \neq 0$ or $\beta_1 \neq 0$, the estimator $RV_t^{(AC,m,1)}$ is slightly biased and the bias vanishes at rate $O\left(m^{-1}\right)$.

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

By examining each of the individual terms in the expression of the variance of $RV_t^{(AC,m,1)}$, it is seen that $RV_t^{(AC,m,1)}$ converges to IV_t at rate $\sqrt{m}$ when $\varepsilon_{t,j} = 0$. In summary, Theorem 2 permits to see that the presence of the endogenous noise alone does not jeopardize the consistency of $RV_t^{(AC,m,1)}$. This allows us to study the properties of the next estimator under the exogenous noise only.

3.3 Barndorff-Nielsen, Hansen, Lunde and Shephard (2008a)

We consider the realized Kernel of Barndorff-Nielsen, Hansen, Lunde and Shephard (2008a) given by:

$$K_t^{BNHLS} = \gamma_{t,0}(r) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right), \tag{24}$$

for a kernel function $k\left(\frac{h-1}{H}\right)$ such that $k(0) = 1$ and $k(1) = 0$. where:

$$\gamma_{t,h}(x) = \sum_{j=1}^{m} x_{t,j}x_{t,j-h}, \tag{25}$$

for all variable x and h. If we further define:

$$\begin{aligned} \gamma_{t,h}(x,y) &= \sum_{j=1}^{m} x_{t,j}y_{t,j-h}, \\ K_t(x) &= \gamma_{t,0}(x) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\left[\gamma_{t,h}(x) + \gamma_{t,-h}(x)\right] \text{ and} \\ K_t(x,y) &= \gamma_{t,0}(x,y) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\left[\gamma_{t,h}(x,y) + \gamma_{t,-h}(x,y)\right], \end{aligned}$$

then K_t^{BNHLS} satisfies (13) with:

$$\begin{aligned} f_{r^*}\left(\left\{r_{t,j}^*\right\}_{j=1}^m\right) &= K_t\left(r^*\right), \\ f_{r^*,u}\left(\left\{r_{t,j}^*, u_{t,j}\right\}_{j=1}^m\right) &= K_t\left(r^*, \Delta u\right) + K_t\left(\Delta u, r^*\right) \text{ and} \\ f_u\left(\left\{u_{t,j}\right\}_{j=1}^m\right) &= K_t\left(\Delta u\right), \end{aligned}$$

where $\Delta u_{t,j} = u_{t,j} - u_{t,j-1}$.

Barndorff-Nielsen and al. (2008a) show that K_t^{BNHLS} is consistent for IV_t in the presence of microstructure noise under various choice of kernel function. For example, setting $k(x) = 1 - x$ (the Bartlett kernel) and $H \propto m^{2/3}$ leads to $K_t^{BNHLS} - IV_t = O_p\left(m^{-1/6}\right)$ under IID noise. Furthermore, this estimator is robust to special forms of endogeneity and serial correlation in the microstructure noise process[4]. As we can see, the expression of K_t^{BNHLS} is reminiscent of the long run variance estimators of Newey and West (1987) and Andrews and Monahan (1992). For practical purpose, we shall rewrite this as:

$$\begin{aligned} K_t^{BNHLS} &= \frac{1}{2}\left(K_{t,Lead}^{BNHLS} + K_{t,Lag}^{BNHLS}\right) \\ &= K_{t,Lead}^{BNHLS} + \frac{1}{2}\left(K_{t,Lag}^{BNHLS} - K_{t,Lead}^{BNHLS}\right), \end{aligned}$$

where

$$\begin{aligned} K_{t,Lead}^{BNHLS} &= \gamma_{t,0}(r) + 2\sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\gamma_{s,h}(r), \text{ and} \\ K_{t,Lag}^{BNHLS} &= \gamma_{t,0}(r) + 2\sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\gamma_{s,-h}(r). \end{aligned}$$

In studying the asymptotic properties of K_t^{BNHLS}, the remainder $\frac{1}{2}\left(K_{t,Lag}^{BNHLS} - K_{t,Lead}^{BNHLS}\right)$ is difficult to handle. However, $K_{t,Lead}^{BNHLS}$ and $K_{t,Lag}^{BNHLS}$ have the same expectation and asymptotic variances. This implies that:

$$\begin{aligned} E\left[K_{t,Lag}^{BNHLS} - K_{t,Lead}^{BNHLS}\right] &= 0, \text{ and} \\ Var\left(K_t^{BNHLS}\right) &\leq Var\left(K_{t,Lead}^{BNHLS}\right). \end{aligned}$$

For simplicity, we shall thus ignore the remainder $\frac{1}{2}\left(K_{t,Lag}^{BNHLS} - K_{t,Lead}^{BNHLS}\right)$ in subsequent analysis. By letting $K_t(x)$ and $K_t(x,y)$ represent their "Lead" versions, we are able to write:

$$K_t^{BNHLS} = K_t\left(r^*\right) + K_t\left(r^*, \Delta u\right) + K_t\left(\Delta u, r^*\right) + K_t\left(\Delta u\right),$$

where

$$\begin{aligned} r_{t,j} &= r_{t,j}^* + \Delta u_{t,j} \text{ and} \\ \Delta u_{t,j} &= \left(\beta_0 + \frac{\beta_1}{\sqrt{m}\sigma_{t,j}^*}\right) r_{t,j}^* - \left(\beta_0 + \frac{\beta_1}{\sqrt{m}\sigma_{t,j-1}^*}\right) r_{t,j-1}^* + \left(\varepsilon_{t,j} - \varepsilon_{t,j-1}\right). \end{aligned}$$

Interestingly, K_t^{BNHLS} has the following representation:

$$K_t^{BNHLS} = RV_t^{(AC,m,1)} + \sum_{h=2}^{H} k\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right),$$

where $\sum_{h=2}^{H} k\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right)$ is unbiased and consistent for zero when $\varepsilon_{t,j} = 0$. In fact, the observed log-return $r_{t,j}$ is not autocorrelated beyond lag one in this case while $Var(r_{t,j}) = O(m^{-1})$. As a result, K_t^{BNHLS} is robust to the type of endogenous noise assumed here. For simplicity and with no loss of generality, we shall thus focus below on the asymptotic behavior of K_t^{BNHLS} under $\beta_0 = \beta_1 = 0$. We have the following theorem.

Theorem 3 *Assume $\beta_0 = \beta_1 = 0$ and that E1 to E4 are satisfied with $\delta \neq 0$. Further let $k(x) = 1 - x$ (the Bartlett kernel). As m goes to infinity and $H = m^b$ for some $b \in (0,1)$, we have:*

$$\begin{aligned}
K_t(r^*) - IV_t &= O_p(H^{1/2}m^{-1/2}),\\
Var\left[K_t(r^*, \Delta u)\right] &\simeq \frac{2\omega_0}{H} + 4\sum_{h=1}^{L}(\omega_{m,h} - \omega_{m,h+1})\left[1 - \frac{(h+1)^2}{H^2}\right] \text{ and}\\
K_t(\Delta u) &= -\varepsilon_{t,0}^2 + \varepsilon_{t,m}^2 - \frac{4}{H}\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-H} - \frac{2}{H}\sum_{j=1}^{m}\varepsilon_{t,j}\varepsilon_{t,j-H-1}\\
&\quad -\frac{2}{H}\sum_{h=2}^{H-1}(\varepsilon_{t,0}\varepsilon_{t,-h} - \varepsilon_{t,m}\varepsilon_{t,m-h}) + \frac{2}{H}(\varepsilon_{t,0}\varepsilon_{t,-H} - \varepsilon_{t,m}\varepsilon_{t,m-H}).
\end{aligned}$$

where we recall that $\omega_{m,L+1} = 0$ in the expression of $Var\left[K_t(r^*, \Delta u)\right]$.

In the IID noise case, we have $\omega_{m,h} = 0$ for all $h \geq 1$. Hence setting $H \propto m^{2/3}$ yields immediately the same result as in Barndorff-Nielsen and al (2008a) up to the end effects:

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

The estimator K_t^{BNHLS} is thus consistent for IV_t if we are willing to neglect the end effects $-\varepsilon_{t,0}^2 + \varepsilon_{t,m}^2$ [5]. In the dependent case, we have:

$$\frac{Var\left[K_t(r^*, \Delta u)\right]}{\omega_0} \to \frac{4\omega_{m,L}}{\omega_0}\left[1 - \frac{(L+1)^2}{H^2}\right] + O\left(m^{-\alpha}\sum_{h=1}^{L}\left[1 - \frac{(h+1)^2}{H^2}\right]\right) \tag{26}$$

as m goes to infinity, where:

$$\begin{aligned}
\sum_{h=1}^{L}\left[1 - \frac{(h+1)^2}{H^2}\right] &= O(m^{\delta})\\
\omega_0 - \omega_{m,L} &= O(m^{\delta-\alpha})
\end{aligned}$$

The first approximation stems from E4 while the second follows from (12). This implies that:

$$\frac{Var\left[K_t(r^*, \Delta u)\right]}{\omega_0} \to 4 + O(m^{\delta-\alpha}) = O(1) \tag{27}$$

This shows that the asymptotic variance of $Var\left[K_t(r^*, \Delta u)\right]$ is proportional to ω_0. In summary,

K_t^{BNHLS} is not consistent in the strict sense when the noise obeys our model, but this estimator delivers its best performance at the highest frequency.

4 Shrinkage Estimators for the IV

Basically, a shrinkage estimator is an optimal linear combination of several estimators (see Hansen, 2007). In the current paper, an estimator is said to be optimal if it minimizes the mean square error (MSE). We have seen in the previous section that the asymptotic variance of our best estimator, K_t^{BNHLS}, is proportional to ω_0. While ω_0 can be arbitrary small, it does not converge to zero. In the current context, an application of a variance reduction technique is fully justified. To further motivate shrinkage estimators for the IV, we examine the contribution of the discretization error and the microstructure noise to the MSEs of the three estimators considered in the previous section. More precisely, we try to understand the trade-offs at play as one moves from a biased estimator to an unbiased estimator on the one hand, and from an unbiased estimator to a consistent estimator on the other hand.

4.1 Discretization Error versus Microstructure Noise

In this subsection, we examine the relative contribution of the discretization error and the microstructure noise to the MSE of an arbitrary estimator $\widehat{IV}_t$ that satisfies (13). We first consider the bias term:

$$E\left[\widehat{IV}_t\right] - IV_t = E\left[f_u\left(\{u_{t,j}\}_{j=1}^m\right)\right]. \tag{28}$$

As the additive terms in (13) are uncorrelated, the variance of $\widehat{IV}_t$ is given by:

$$\begin{aligned} Var\left[\widehat{IV}_t\right] &= Var\left[f_{r^*}\left(\{r_{t,j}^*\}_{j=1}^m\right)\right] + Var\left[f_{r^*,u}\left(\{r_{t,j}^*, u_{t,j}\}_{j=1}^m\right)\right] \\ &\quad + Var\left[f_u\left(\{u_{t,j}\}_{j=1}^m\right)\right]. \end{aligned}$$

Hence the overall MSE is:

$$\begin{aligned} MSE\left[\widehat{IV}_t\right] &= Var\left[f_{r^*}\left(\{r_{t,j}^*\}_{j=1}^m\right)\right] + Var\left[f_{r^*,u}\left(\{r_{t,j}^*, u_{t,j}\}_{j=1}^m\right)\right] \\ &\quad + Var\left[f_u\left(\{u_{t,j}\}_{j=1}^m\right)\right] + E\left[f_u\left(\{u_{t,j}\}_{j=1}^m\right)\right]^2. \end{aligned} \tag{29}$$

As $f_{r^*,u}\left(\left\{r_{t,j}^*, 0\right\}_{j=1}^m\right) = f_u\left(\{0\}_{j=1}^m\right) = 0$, the above MSE reduces to the variance of $f_{r^*}\left(\left\{r_{t,j}^*\right\}_{j=1}^m\right)$ when there is no noise in the data. Based on this argument, we adopt the following definition.

Definition 4 *The contribution of the microstructure noise to the MSE of $\widehat{IV}_t$ is:*

$$\begin{aligned} MSE_u\left[\widehat{IV}_t\right] &= Var\left[f_{r^*,u}\left(\{r_{t,j}^*, u_{t,j}\}_{j=1}^m\right)\right] + Var\left[f_u\left(\{u_{t,j}\}_{j=1}^m\right)\right] \\ &\quad + E\left[f_u\left(\{u_{t,j}\}_{j=1}^m\right)\right]^2. \end{aligned} \tag{30}$$

Accordingly, we define the MSE due to discretization as:

$$MSE_{r^*}\left[\widehat{IV}_t\right] = Var\left[f_{r^*}\left(\{r_{t,j}^*\}_{j=1}^m\right)\right]. \tag{31}$$

This definition imputes to the microstructure noise the part of the MSE of $\widehat{IV}_t$ that vanishes when there is actually no microstructure noise in the data. In Table 1, we examine the expression of $f_{r^*}\left(\left\{r^*_{t,j}\right\}_{j=1}^m\right)$ for the three estimators listed in the examples. It is seen that this expression includes more and more terms as one moves from the top to the bottom of the table. In fact, $RV_t^{(AC,m,1)}$ kills of the bias of its ancestor $RV_t^{(m)}$ at the expense of a higher discretization error. Likewise, K_t^{BNHLS} brings consistency upon conceding a higher discretization error with respect to the unbiased estimator $RV_t^{(AC,m,1)}$.

	$f_{r^*}\left(\left\{r^*_{t,j}\right\}_{j=1}^m\right)$	$Var\left[f_{r^*}\left(\left\{r^*_{t,j}\right\}_{j=1}^m\right)\right]$
$RV_t^{(m)}$	$\sum_{j=1}^m r^{*2}_{t,j}$	$2\sum_{j=1}^m \sigma^{*4}_{t,j}$
$RV_t^{(AC,m,1)}$	$\sum_{j=1}^m r^{*2}_{t,j} + \sum_{j=1}^m r^*_{t,j}\left(r^*_{t,j+1} + r^*_{t,j-1}\right)$	$2\sum_{j=1}^m \sigma^{*4}_{t,j}$ $+4\sum_{j=1}^m \sigma^{*2}_{t,j}\sigma^{*2}_{t,j-1} + O(m^{-2})$
K_t^{BNHLS}	$\sum_{j=1}^m r^{*2}_{t,j} + \sum_{j=1}^m r^*_{t,j}\left(r^*_{t,j+1} + r^*_{t,j-1}\right)$ $+\sum_{h=2}^H k\left(\frac{h-1}{H}\right)\sum_{j=1}^m r^*_{t,j}\left(r^*_{t,j+h} + r^*_{t,j-h}\right)$	$2\sum_{j=1}^m \sigma^{*4}_{t,j} + 4\sum_{j=1}^m \sigma^{*2}_{t,j}\sigma^{*2}_{t,j-1}$ $+4\sum_{h=2}^H k\left(\frac{h-1}{H}\right)\sum_{j=1}^m \sigma^{*2}_{t,j}\sigma^{*2}_{t,j-h}$ $+O(Hm^{-2})$

Table 1: Part of the MSE due to discretization

We now turn to discuss the MSE due to IID microstructure noise. Unlike $RV_t^{(m)}$ whose bias equals $2mE\left[u^2_{t,j}\right]$, the estimators $RV_t^{(AC,m,1)}$ and K_t^{BNHLS} are unbiased. As a consequence, the MSE of $RV_t^{(m)}$ increases at rate m^2 while those of $RV_t^{(AC,m,1)}$ and K_t^{BNHLS} are only linear in m. On the other hand, the consistency of K_t^{BNHLS} ensures that its variance eventually becomes smaller than that of $RV_t^{(AC,m,1)}$ as m goes to infinity. But there is at least two situations where $RV_t^{(AC,m,1)}$ can have lower variance than K_t^{BNHLS}. The first situation is the one in which the sampling frequency m is not large enough to make the asymptotic results for K_t^{BNHLS} reliable. In fact, the variance of K_t^{BNHLS} can be arbitrarily high in fixed frequency although it diminishes as m goes to infinity. The second situation is the case where the variance of the microstructure noise is so small that it contributes very little to the MSE. In this case, the MSE of the estimators basically reduces to the variance of $f_{r^*}\left(\left\{r^*_{t,j}\right\}_{j=1}^m\right)$ which happens to be larger for K_t^{BNHLS}.

Our intuitions are supported by a simulation study by Gatheral and Oomen (2007). These authors implemented twenty realized measures that aim to estimate the IV. Their main finding is that even though inconsistent, kernel-type estimators like $RV_t^{(AC,m,1)}$ often deliver good performances in term of MSE at sampling frequencies commonly encountered in practice. This result stems from the fact that an inconsistent estimator necessarily delivers its best performance at moderate frequency while a consistent estimator may require quite high frequency data in order to perform well[6]. Unfortunately, there is no clear rule indicating the minimum sampling frequency required for the asymptotic theory of K_t^{BNHLS} to be usable. Moreover, the microstructure noise is not observed so that it is difficult to tell whether or not its size is small compared to the efficient returns. In the next section, we propose to combine linearly the two unbiased estimators $RV_t^{(AC,m,1)}$ and K_t^{BNHLS} in order to achieve an optimal signal-to-noise trade off.

4.2 Shrinkage Realized Kernels

Let us consider the decomposition:

$$K_t^{BNHLS} = \theta_{1,t}^{(L)} + \theta_{2,t}^{(L)} \tag{32}$$

where

$$\theta_{1,t}^{(L)} = \gamma_{t,0}(r) + \sum_{h=1}^{L+1} k\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right) \text{ and} \tag{33}$$

$$\theta_{2,t}^{(L)} = \sum_{h=L+2}^{H} k\left(\frac{h-1}{H}\right)\left(\gamma_{t,h}(r) + \gamma_{t,-h}(r)\right), \tag{34}$$

Note that $\theta_{1,t}^{(L)}$ is a smoothed version of $RV_t^{(AC,m,L+1)}$ and is thus unbiased for the IV when $k(x) = 1 - x$.

We consider a linear combinations of the form:

$$\begin{aligned} K_t^{\varpi} &= \varpi K_t^{BNHLS} + (1-\varpi)\theta_{1,t}^{(L)}, \ \varpi \in \mathbb{R}, \\ &= \theta_{1,t}^{(L)} + \varpi\theta_{2,t}^{(L)}. \end{aligned} \tag{35}$$

Note that K_t^{ϖ} is a realized kernels with kernel function given by:

$$\begin{aligned} g(x) &= k(x), \ 0 \leq x \leq \frac{L}{H}, \text{ and} \\ g(x) &= \varpi k(x), \ \frac{L}{H} < x \leq 1 \end{aligned}$$

The kernel function $g(x)$ is discontinuous at $x = \frac{L}{H}$ unless $\varpi = 1$.

As both estimators are unbiased, the weight ϖ that minimizes the variance of K_t^{ϖ} *conditional on the volatility path* is given:

$$\begin{aligned} \varpi_t^* &= \arg\min_{\varpi} E\left[\left(K_t^{\varpi} - IV_t\right)^2 | \{\sigma\}\right]. \\ &= -\frac{Cov\left(\theta_{1,t}^{(L)}, \theta_{2,t}^{(L)} | \{\sigma\}\right)}{Var\left(\theta_{2,t}^{(L)} | \{\sigma\}\right)} \end{aligned}$$

The resulting $K_t^{\varpi^*}$ is termed "*shrinkage realized kernels*", as it is obtained by shrinking the unbiased estimator in the direction of the consistent estimator. The efficiency gain of the shrinkage estimator with respect to K_t^{BNHLS} is:

$$Var\left(K_t^{BNHLS} | \{\sigma\}\right) - Var\left(K_t^{\varpi^*} | \{\sigma\}\right) = \left(\rho_{1,2,t}\sqrt{Var\left(\theta_{1,t} | \{\sigma\}\right)} + \sqrt{Var\left(\theta_{2,t} | \{\sigma\}\right)}\right)^2 \geq 0.$$

where $\rho_{1,2,t}$ denotes the conditional correlation between $\theta_{1,t}^{(L)}$ and $\theta_{2,t}^{(L)}$. Hence the shrinkage estimator inherits the good properties of K_t^{BNHLS} at high frequency while performing better than $\theta_{1,t}^{(L)}$.

The shrinkage weight ϖ_t^* is in fact unfeasible, as the conditional moments involved in its expression are typically unknown. A simple strategy is to look for a constant shrinkage weight ϖ^* that minimizes the marginal variance of K_t^{ϖ}. By the law of total variance, we have:

$$\begin{aligned} Var_{Total}\left(K_t^{\varpi}\right) &= Var\left[E\left(K_t^{\varpi}\middle|\{\sigma\}\right)\right] + E\left[Var\left(K_t^{\varpi}\middle|\{\sigma\}\right)\right] \\ &= Var\left[IV_t\right] + E\left[Var\left(K_t^{\varpi}\middle|\{\sigma\}\right)\right]. \end{aligned}$$

Therefore, choosing ϖ to minimize the marginal variance of K_t^{ϖ} is equivalent to choosing ϖ to minimize the expected conditional variance of K_t^{ϖ}. We estimate the constant weight by:

$$\widehat{\varpi}^* = -\frac{\frac{1}{T}\sum_{t=1}^{T}\left(\theta_{1,t}^{(L)} - \overline{\theta_{1,T}^{(L)}}\right)\theta_{2,t}^{(L)}}{\frac{1}{T}\sum_{t=1}^{T}\left(\theta_{2,t}^{(L)}\right)^2}. \tag{36}$$

where $\overline{\theta_{1,T}^{(L)}} = \sum_{t=1}^{T}\theta_{1,t}^{(L)}$. Even though ϖ^* does not converge to ϖ_t^*, it achieves *on average* the goal assigned to the ideal weight ϖ_t^*.

To guess the asymptotic behavior of the constant ϖ^*, we write it as follows:

$$\varpi^* = \frac{1 - \rho_{1,2}x}{1 - 2\rho_{1,2}x + x^2}$$

where $x = \sqrt{\frac{Var\left(K_t^{BNHLS}\right)}{Var\left(\theta_{1,t}^{(L)}\right)}}$, $\rho_{1,2}$ is the unconditional correlation between $\theta_{1,t}^{(L)}$ and K_t^{BNHLS} and the variances are also unconditional. When K_t^{BNHLS} is consistent or $O_p(1)$, we have $z = O\left(m^{-1}\right)$ so that ϖ^* converges to one at rate m. Hence $K_t^{\varpi^*}$ and K_t^{BNHLS} are asymptotically equivalent. Hence the efficiency gain of $K_t^{\varpi^*}$ over K_t^{BNHLS} is higher at moderate of low frequency.

Here is advocated a simple framework where the loss function is the MSE and the shrinkage weight is constant and independent of the estimators K_t^{BNHLS} and $RV_t^{(AC,m,1)}$. Stein (1956) derived a shrinkage estimator for the mean of a multivariate normal distribution that outperforms the usual empirical mean. The Stein estimator is obtained by shrinking the empirical mean toward zero using a shrinkage weight that is a nonlinear in the empirical mean itself. Other shrinkage estimators involving different loss functions are discuss in Hansen (2007, 2008). More specifically, our estimator $K_t^{\varpi^*}$ is related to the estimator proposed in Ghysels, Mykland and Renault (2008) that consists of shrinking the current period estimator of IV_t toward its optimal forecast from the previous period. Finally, the shrinkage method can be used independently of the postulated microstructure noise model.

5 Inference on the Microstructure Noise Parameters

From now one, the notation $\gamma_{t,h}$ is used for $\gamma_{t,h}(r)$ where the latter is defined in (25). We note from (20) that:

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

where we recall that $\omega_{m,h}$ is the h^{th} autocovariance of $\varepsilon_{t,j}$ when observed at frequency m.

Let $b_t^{(m)} = E\left[RV_t^{(m)} - IV_t\right]$ denote the bias of the realized volatility computed at the record frequency. It follows from Lemma 7 in appendix that when $q = 1$, we have:

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

The endogenous parameters β_0 and β_1 hidden in the expression of the bias $b_t^{(m)}$ are unfortunately unidentified. We shall thus focus on the estimation of the bias as a whole rather that tackling β_0 and β_1 individually. In subsection 4.1, we discuss the estimation of $b_t^{(m)}$ and $\{\omega_{m,h}\}_{h=0}^{L}$ while in subsection 4.2 we deal with the memory parameters (L, α, δ).

5.1 Estimation of the Correlogram

By neglecting the $O(m^{-1})$ end terms in the expression of the bias $b_t^{(m)}$, we obtain the following moment conditions:

$$E\left[RV_t^{(m)} - b_t^{(m)} - IV_t\right] = 0, \text{ and} \tag{37}$$

$$E\left[b_t^{(m)} + \left(\gamma_{t,1} + \gamma_{t,-1}\right) - 2m\left(\omega_{m,1} - \omega_{m,2}\right)\right] = 0. \tag{38}$$

We also have:

$$E\left[\left(\gamma_{t,h+1} + \gamma_{t,-h-1}\right) - 2m\left(-\omega_{m,h} + 2\omega_{m,h+1} - \omega_{m,h+2}\right)\right] = 0,\ 1 \le h \le L. \tag{39}$$

Given that $\omega_{m,h} = 0$ for $h > L$, we have $L + 2T$ moment conditions to estimate $L + 2T$ parameters.

Estimating these parameters by the method of moments is straightforward. First solving for $\omega_{m,L}$ and then proceeding by backward substitution yields:

$$\widehat{\omega}_{m,h} = -\frac{1}{2Tm}\sum_{s=1}^{T}\sum_{l=1}^{L-h+1} l\left(\gamma_{s,h+l} + \gamma_{s,-h-l}\right),\ h = 1, ...L, \tag{40}$$

$$\widehat{b}_t^{(m)} = -\gamma_{t,1} - \gamma_{t,-1} - \frac{1}{T}\sum_{s=1}^{T}\sum_{l=2}^{L+1}\left(\gamma_{s,l} + \gamma_{s,-l}\right) \text{ and} \tag{41}$$

$$\overline{RV}_t^{(AC,m,L+1)} = \gamma_{t,0} + \gamma_{t,1} + \gamma_{t,-1} + \frac{1}{T}\sum_{s=1}^{T}\sum_{l=2}^{L+1}\left(\gamma_{s,l} + \gamma_{s,-l}\right), \tag{42}$$

where $\widehat{\omega}_{m,h}$, $\widehat{b}_t^{(m)}$ and $\overline{RV}_t^{(AC,m,L+1)}$ are unbiased estimators of $\omega_{m,h}$, $b_t^{(m)}$ and IV_t respectively[7]. It is seen that $\overline{RV}_t^{(AC,m,L+1)}$ is a bias corrected version of the standard realized variance which uses the data available at all periods to estimate the IV of each period. To estimate the variance ω_0, we use the expression of the bias of the RV sampled at the highest frequency. We have:

$$\widehat{\omega}_0 = \frac{1}{2mT}\sum_{t=1}^{T}\widehat{b}_t^{(m)} + \widehat{\omega}_{m,1} \tag{43}$$

To estimate the covariance matrix of $\widehat{\omega}_m = (\widehat{\omega}_{m,0}, \widehat{\omega}_{m,1}, ..., \widehat{\omega}_{m,L})'$, we define:

$$\gamma_{t,j,(2,L+1)} = \left(\gamma_{t,j,2}, ..., \gamma_{t,j,L+1}\right)',$$

where $\gamma_{t,j,h} = \frac{1}{2} r_{t,j} \left(r_{t,j-h} + r_{t,j+h}\right)$ for all t and h. Then we have:

$$(\widehat{\omega}_{m,1}, ..., \widehat{\omega}_{m,L})' = \frac{1}{mT} \sum_{t=1}^{T} \sum_{j=1}^{m} P^{-1} \gamma_{t,j,(2,L+1)},$$

where P is the $L \times L$ matrix with elements: $P_{i,i} = -1$, $P_{i,i+1} = 2$, $P_{i,i+2} = -1$, and $P_{i,j} = 0$ otherwise $1 \leq i, j \leq L$. If we further define:

$$\widehat{\omega}_{t,j,0} = -\frac{1}{2} \sum_{h=1}^{L+1} \left(\gamma_{t,j,h} + \gamma_{t,j,-h}\right) + \left(P^{-1} \gamma_{t,j,(2,L+1)}\right)_1 \text{ and} \tag{44}$$

$$(\widehat{\omega}_{t,j,1}, ..., \widehat{\omega}_{t,j,L})' = \left(P^{-1} \gamma_{t,j,(2,L+1)}\right)', \tag{45}$$

with $\left(P^{-1} \overline{\gamma}_{t,(2,L+1)}\right)_1$ being the first element of $P^{-1} \overline{\gamma}_{t,(2,L+1)}$, then we are able to write:

$$\widehat{\omega}_{m,h} = \frac{1}{mT} \sum_{t=1}^{T} \sum_{j=1}^{m} \widehat{\omega}_{t,j,h}, \text{ for all } h.$$

We have the following convergence result.

Theorem 5 *Define the subsampled variance $\widehat{Q}_h$ as:*

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

Then under Assumptions E1, E2 and E5, we have:

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

as T goes to infinity and m is fixed.

This Central Limit Theorem holds under general conditions (See Politis, Romano and wolf 1997, 1999). The steps of the proof are the same as for the Theorem 1 in Ubukata and Oya (2009). However, our result stresses that m is fixed and only T goes to infinity. This precision is important because the number of daily observations available to estimate an autocovariance of order h (for fixed h) is finite event if m goes to infinity.

The knowledge of L is required to estimate the correlogram of the microstructure noise. A simple way to estimate L is to perform a significance test for $\omega_{m,h}$. Under the null hypothesis that $\omega_{m,h} = 0$, then

$$\widehat{\tau}_h = \frac{(mT)^{1/2} \widehat{\omega}_{m,h}}{\sqrt{\widehat{Q}_h}} \to N(0,1) \tag{46}$$

The above statistics diverges under the alternative. Our estimator $\widehat{L}$ of L is the maximum lag at which the null is rejected. Note that $\widehat{L}$ is consistent for L to the extent that this significance test is powerful, that is:

$$h \leq L \Leftrightarrow \Pr\left(|\widehat{\tau}_h| > 1.96\right) \to 1$$

We now turn to discuss the estimation of α and δ under the assumption that E3 and E4 are satisfied.

5.2 Assessing the values of α and δ

Assumption E3 stipulates that ω_0 is constant for all m while $\omega_{m,h} - \omega_{m,h+1} = O(m^{-\alpha})$ for $h = 0, ..., L-1$. We can thus write:

$$\frac{\omega_{m,h} - \omega_{m,h+1}}{\omega_0} \simeq C_h m^{-\alpha}, \text{ with } C_h \in \left[\underline{C}, \overline{C}\right].$$

Taking the logs of both side of the equality and averaging over h yields:

$$\alpha \simeq \frac{-1}{L \log m} \sum_{h=0}^{L-1} \log\left(\frac{\omega_{m,h} - \omega_{m,h+1}}{\omega_0}\right) + \frac{1}{L \log m} \sum_{h=0}^{L-1} \log C_h,$$

where $\frac{1}{L\log m}\sum_{h=1}^{L-1} \log C_h \in \left[\frac{\log \underline{C}}{\log m}, \frac{\log \overline{C}}{\log m}\right]$ so that this term goes to zero as m goes to infinity. If m is fixed but sufficiently high to make $\frac{\log \underline{C}}{\log m}$ and $\frac{\log \overline{C}}{\log m}$ negligible, a good estimator of α is given by:

$$\widehat{\alpha} = \frac{1}{\log m}\left[\log \widehat{\omega}_0 - \frac{1}{L}\sum_{h=0}^{L-1} \log\left(\widehat{\omega}_{m,h} - \widehat{\omega}_{m,h+1}\right)\right]. \tag{47}$$

The estimator $\widehat{\alpha}$ enjoys the consistency property of $\widehat{\omega}_{m,h}$, being a smooth function of the latter which is consistent according to Theorem 5. Using the Delta method, we obtain:

$$\sqrt{mT}\log m\left(\widehat{\alpha} - \alpha\right) \to N\left(0, (\nabla\alpha)' Q_\omega (\nabla\alpha)\right), \tag{48}$$

as T goes to infinity and m is fixed, where

$$\begin{aligned}
Q_\omega &= \frac{m}{T}\sum_{t=1}^{T}\left(\frac{1}{m}\sum_{j=1}^{m}\underline{\omega}_{t,j} - \underline{\omega}\right)\left(\frac{1}{m}\sum_{j=1}^{m}\underline{\omega}_{t,j} - \underline{\omega}\right)', \\
\underline{\omega}_{t,j} &= \left(\widehat{\omega}_{t,j,0}, \widehat{\omega}_{t,j,1}, ..., \widehat{\omega}_{t,j,L}\right)' \text{ and} \\
\underline{\omega} &= \frac{1}{mT}\sum_{t=1}^{T}\sum_{j=1}^{m}\underline{\omega}_{t,j}\underline{\omega}_{t,j}.
\end{aligned}$$

The elements of the vector $\nabla\alpha$ are given by:

$$\begin{aligned}
(\nabla\alpha)_1 &= \frac{1}{\omega_0} - \frac{1}{L\left(\omega_0 - \omega_1\right)}, \\
(\nabla\alpha)_h &= \frac{1}{L\left(\omega_{m,h-2} - \omega_{m,h-1}\right)} - \frac{1}{L\left(\omega_{m,h-1} - \omega_{m,h}\right)};\ 2 \leq h \leq L \text{ and} \\
(\nabla\alpha)_{L+1} &= \frac{1}{L\left(\omega_{L-1} - \omega_L\right)}.
\end{aligned}$$

Provided that the conditions of Theorem 5 hold, the asymptotic distribution (48) is valid even when E3 and E4 are not satisfied. This is true because the distribution is derived under fixed m.

To estimate δ, we exploit Assumption E4 according to which $L = Cm^{\delta}$. Our estimator of δ is:

$$\widehat{\delta} = \frac{\log \widehat{L}}{\log m}, \tag{49}$$

where $\widehat{L}$ stems from the significance test based on (46) and m is large enough to make $\frac{\log C}{\log m}$ negligible. The estimator $\widehat{\delta}$ inherits the asymptotic properties of $\widehat{L}$. Finally, note that both $\widehat{\alpha}$ and $\widehat{\delta}$ are Hill (1975) type estimators.

6 Monte Carlo Simulation

The aim in this subsection is to assess the performance of the shrinkage estimator of IV and the quality of the estimators of $\{\omega_{m,l}\}_{l=0}^{\widehat{L}}$ by simulations.

6.1 Simulation Design

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

$$dp_t^* = \sigma_t dW_{1,t} \text{ and} \tag{50}$$

$$d\sigma_t^2 = \kappa\left(\alpha - \sigma_t^2\right)dt + \gamma\sigma_t\left[\rho dW_{1,t} + \sqrt{1-\rho^2}dW_{2,t}\right], \tag{51}$$

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

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

where $\rho = 0$ corresponds to the no leverage assumption made in deriving our theoretical results. The case $\rho = -0.5$ is used to check the robustness of our conclusions. In the specification above, the unit period is one year.

We simulated data for $T = 1000$ days using Euler discretization scheme at one second. Assuming that the market opens from 9:30 am to 4:00 pm, this yields 23400 discretization points within each day. We then consider four frequencies at which the price can be observed: 30 seconds, one minute, two minutes and five minutes. This yields four data sets with respectively $m = 780$, 390, 195 and 78 observations per day. Each data set is contaminated with a microstructure noise process simulated according to the following model:

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

where the exogenous noise $\varepsilon_{t,j}$ is an MA(3).

$$\begin{aligned} \varepsilon_{t,j} &= v_{t,j} + \alpha_1 v_{t,j-1} + \alpha_2 v_{t,j-2} + \alpha_3 v_{t,j-3} \text{ and} \\ & v_{t,j} \overset{IID}{\sim} N(0,\ \alpha_0). \end{aligned}$$

We set the following values for the noise parameter:

$$\begin{aligned}\beta_0 &= 0.5;\ \beta_1 = 0.5;\\ \alpha_1 &= 0.5;\ \alpha_2 = 0.2;\ \alpha_3 = 0.05.\end{aligned}$$

In order to make this simulation design less arbitrary, we will vary α_0 in order to increase or decrease the autocovariances of $\varepsilon_{t,j}$. In fact, we have:

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

Because the link between ω_0 and α_0 is one-to-one, we will directly vary ω_0 within the range:

$$\omega_0 \in \left\{2.25\times10^{-8}; 2.5\times10^{-7}; 2.25\times10^{-6}; 2.5\times10^{-5}\right\}.$$

The value $\omega_0 = 2.5\times10^{-7}$ has been used in Zhang and al. (2005) at five minute sampling frequency while $\omega_0 = 2.25\times10^{-6}$ has served in Ait-Sahalia and al. (2005) at frequencies that range from one minute to thirty minutes.

We consider three IV estimators: the unbiased estimator $\theta_{1,t}^{(L)}$ (Equation (33)), the consistent estimator K_t^{BNHLS} with Bartlett kernel and the shrinkage estimator $K_t^{\varpi^*}$ with constant weight. After several trials, the bandwidth $H = \lfloor 0.4m^{2/3}\rfloor$ seems to work well for K_t^{BNHLS}. In order to estimate the noise autocovariances, a first guess of the maximum lag L is needed. Although $L = 3$ for the simulated noise, we set $\widehat{L} = 4$ in the subsequent computations. A discussion on how to formulate the first guess of L in practice is provided in the empirical section.

6.2 Simulation Results

First, we consider the volatility signature plots, that is, the curve of $\frac{1}{T}\sum_{t=1}^{T} RV_t^{(m_q)}$ plotted against $q = \frac{m}{m_q}$. Figure 2.1 describes one simulated sample without noise while Figure 2.2 describes a noisy version of the same data. It is seen that the volatility signature plots (at the top) are quite informative about the presence of the noise.

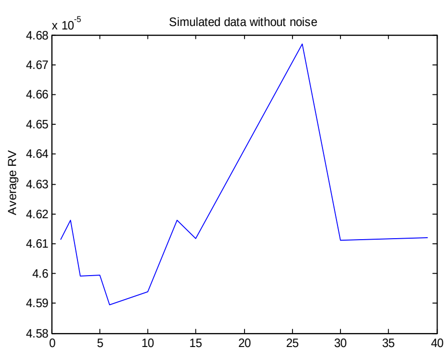

Figure 2.1. Data with no noise.

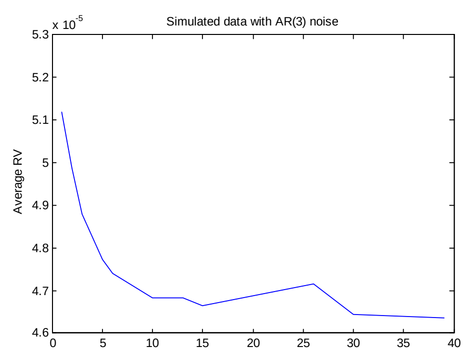

Figure 2.2. Data with MA(3) noise.

Figure 2: Volatility Signature Plots.

For any arbitrary estimator $\widehat{IV}_t$ of IV_t, the empirical MSE of $\widehat{IV}_t$ is given by:

$$MSE(\widehat{IV}_t) = \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{IV}_t - IV_t\right)^2. \tag{52}$$

This MSE converges to the marginal variance of $\widehat{IV}_t$. Table 2 displays the MSE of $\theta_{1,t}^{(L)}$, K_t^{BNHLS} and $K_t^{\varpi^*}$ for the efficient price model with no leverage while Table 3 shows the results when leverage is included. Interestingly, we note that the introduction of leverage slightly reduces the variance in all the scenarios. Otherwise, the two tables display qualitatively similar results.

Variance of $\varepsilon_{t,j}$	Frequency	MSE ($\times 10^{-6}$)			Shrinkage weight
ω_0	m	K_t^{BNHLS}	$\theta_{1,t}^{(L)}$	$K_t^{\varpi^*}$	$\widehat{\varpi}^*$
2.25×10^{-8}	780	0.0016	0.0009	0.0009	0.2525
	390	0.0022	0.0018	0.0015	0.3988
	195	0.0029	0.0028	0.0025	0.4721
	78	0.0050	0.0052	0.0047	0.6036
2.5×10^{-7}	780	0.0017	0.0017	0.0012	0.4962
	390	0.0022	0.0022	0.0017	0.5125
	195	0.0030	0.0033	0.0026	0.5719
	78	0.0051	0.0055	0.0049	0.6438
2.25×10^{-6}	780	0.0049	0.0165	0.0048	0.9263
	390	0.0045	0.0113	0.0044	0.8955
	195	0.0050	0.0095	0.0049	0.8387
	78	0.0071	0.0092	0.0070	0.8546
2.5×10^{-5}	780	0.3572	1.6563	0.3571	1.0040
	390	0.2314	0.7405	0.2314	1.0014
	195	0.1597	0.3862	0.1596	0.9845
	78	0.1126	0.1674	0.1126	0.9714

Table 2: Evaluating the performance of the shrinkage estimators of IV_t by Monte Carlo: Case with no Leverage Effect.

Variance of $\varepsilon_{t,j}$	Frequency	MSE ($\times 10^{-6}$)			Shrinkage weight
ω_0	m	K_t^{BNHLS}	$\theta_{1,t}^{(L)}$	$K_t^{\varpi^*}$	$\widehat{\varpi}^*$
2.25×10^{-8}	780	0.0016	0.0008	0.0007	0.2223
	390	0.0021	0.0016	0.0014	0.3172
	195	0.0029	0.0028	0.0025	0.4583
	78	0.0048	0.0055	0.0047	0.7701
2.5×10^{-7}	780	0.0016	0.0014	0.0010	0.4359
	390	0.0022	0.0022	0.0017	0.4867
	195	0.0030	0.0031	0.0026	0.5216
	78	0.0050	0.0057	0.0049	0.7724
2.25×10^{-6}	780	0.0047	0.0173	0.0046	0.9108
	390	0.0044	0.0113	0.0042	0.8696
	195	0.0048	0.0088	0.0047	0.8685
	78	0.0065	0.0084	0.0064	0.8679
2.5×10^{-5}	780	0.3461	1.5821	0.3461	0.9997
	390	0.2217	0.8107	0.2217	1.0065
	195	0.1554	0.3740	0.1554	0.9976
	78	0.1083	0.1529	0.1077	0.9249

Table 3: Evaluating the performance of the shrinkage estimators of IV_t by Monte Carlo: Case with Leverage Effect.

The efficient return data has been contaminated with a Gaussian MA(3) microstructure noise driven by the same parameters regardless of the sampling frequency. Hence for a given ω_0, the signal-to-noise ratio deteriorates as the sampling frequency increases. Likewise for a fixed sampling frequency, the signal-to-noise ratio deteriorates as ω_0 increases. It turns out that the optimal weight ϖ^* allocated to the consistent estimator heavily depends on the variance of the microstructure noise. In general, ϖ^* is increasing in ω_0. In large ω_0 scenarios, the weight is close to one and decreases very slowly as m increases. By contrast, the weight is smaller in small ω_0 scenarios and increases quite fast as m decreases. Overall, the relative efficiency gain of the shrinkage estimator over the consistent estimator is large when m is large and ω_0 is small. Note that compared to the consistent estimator K_t^{BNHLS}, the MSE of $K_t^{\varpi^*}$ is smaller by more than one half in the scenario ($\omega_0 = 2.25 \times 10^{-7}$, $m = 780$) and by about one third for ($\omega_0 = 2.25 \times 10^{-7}$, $m = 390$).

Not surprisingly, the inconsistent estimator $\theta_{1,t}^{(L)}$ performs better than the consistent estimator in small ω_0 scenarios ($\omega_0 = 2.25 \times 10^{-7}$). In the large ω_0 scenarios ($\omega_0 > 2.25 \times 10^{-7}$), $\theta_{1,t}^{(L)}$ is not preferred to K_t^{BNHLS} at all the sampling frequencies while the best performance of $\theta_{1,t}^{(L)}$ is achieved at lower frequencies. This is consistent with the fact that the larger the noise variance ω_0, the lower the frequency that achives the optimal signal-to-noise ratio for $\theta_{1,t}^{(L)}$. For a discussion on optimal sampling frequencies in the IID noise context, see for example Bandi and Russell (2006).

Tables 4.1 and 4.2 show the estimation results for the correlogram of the noise in the scenarios ($\omega_0 = 2.25 \times 10^{-7}$; $m = 780$) and ($\omega_0 = 2.25 \times 10^{-7}$; $m = 390$) respectively. In these tables, the column labeled "True" contains the true values of the parameters. The estimates are computed using the Equation (44) while the standard deviations are computed from Equation (45) with ten lags. Firstly, we note that the estimator of ω_0 is biased upward and the bias decreases as the record frequency increases. In fact, the bias of $\widehat{\omega}_0$ is due to the presence of endogenous noise. Under the assumption that $\sigma^*_{t,qk}$ is stationary, the unconditional bias of $\widehat{\omega}_0$ is given by:

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

Hence while $\widehat{\omega}_0$ is biased for the variance of the exogenous noise, it does reflect the actual size of the total noise contaminating the price.

	True (x10^{-7})	Estimate (x10^{-7})	Std. Dev. (x10^{-7})	Student-t
$\widehat{\omega}_0$	2.2500	3.8423	0.4218	9.1095
$\widehat{\omega}_1$	1.0619	1.0722	0.1224	8.7595
$\widehat{\omega}_2$	0.3917	0.3946	0.0512	7.7072
$\widehat{\omega}_3$	0.0870	0.0932	0.0210	4.4365
$\widehat{\omega}_4$	0.0000	0.0082	0.0098	0.8432

Table 4.1. $m = 780$

	True (x10^{-7})	Estimate (x10^{-7})	Std. Dev. (x10^{-7})	Student-t
$\widehat{\omega}_0$	2.2500	5.4195	0.5989	9.0489
$\widehat{\omega}_1$	1.0619	1.0885	0.1386	7.8517
$\widehat{\omega}_2$	0.3917	0.4109	0.0689	5.9660
$\widehat{\omega}_3$	0.0870	0.1006	0.0375	2.6806
$\widehat{\omega}_4$	0.0000	0.0021	0.0212	0.0995

Table 4.2. $m = 390$

Table 4: Estimated correlogram of the noise (Simulated Data).

The results suggest that the autocovariances estimators $\{\widehat{\omega}_l\}_{l=1}^4$ are unbiased. The Student-t statistics displayed in the last column indicate that the null hypothesis $\omega_4 = 0$ cannot be rejected at level 5%. This suggests that upon formulating a good initial guess of L, a standard t-test can be reliably used to assess the significance of the noise autocovariances.

7 Empirical Application

In the first subsection, we describe the data and discuss some methodological aspects of the empirical study. The results are presented in the second subsection.

7.1 Data and Methodology

For this application, we use the data on twelve stocks listed in the Dow Jones Industrial[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 is open from 9:30 am to 4:00 pm, and this results in $m = 390$ observations per day. There are a few missing observations (less than 5 missing data per day) which we filled in using the previous tick method.

While our theoretical model assumes no jumps, the conclusions of many studies strongly suggest its presence in observed prices (see e.g Eraker (2004)). By assuming that the jumps are uncorrelated with both the efficient price and the noise, we can perform our analysis by ignoring their presence. This does not affect the estimators of the noise parameters, but the estimators $K_t^{\varpi^*}$, K_t^{BNHLS}

and $\theta_{1,t}^{(L)}$ are now designed the total quadratic variation which is equal to the IV_t plus the jump contribution. To deal with outliers, we follow an intuition given in Barndorff-Nielsen and al (2008b)[9] 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. The resulting $r_{t,j}^{NEW}$ is treated as our initial observed return $r_{t,j} \equiv r_{t,j}^{NEW}$. This approach assumes that a jump must cannot be 50 times larger than the absolute median of the data. It has three main advantages. Firstly, it preserves the structure of dependence of the microstructure noise which is of interest in our analysis. Secondly, the process $\left| r_{t,j}^{OLD} \right|$ contains substantial information about the normal range of the data, including the jumps. And thirdly, the median is known to be robust to outliers. Figure 3 show examples of the impact of this preprocessing on the data.

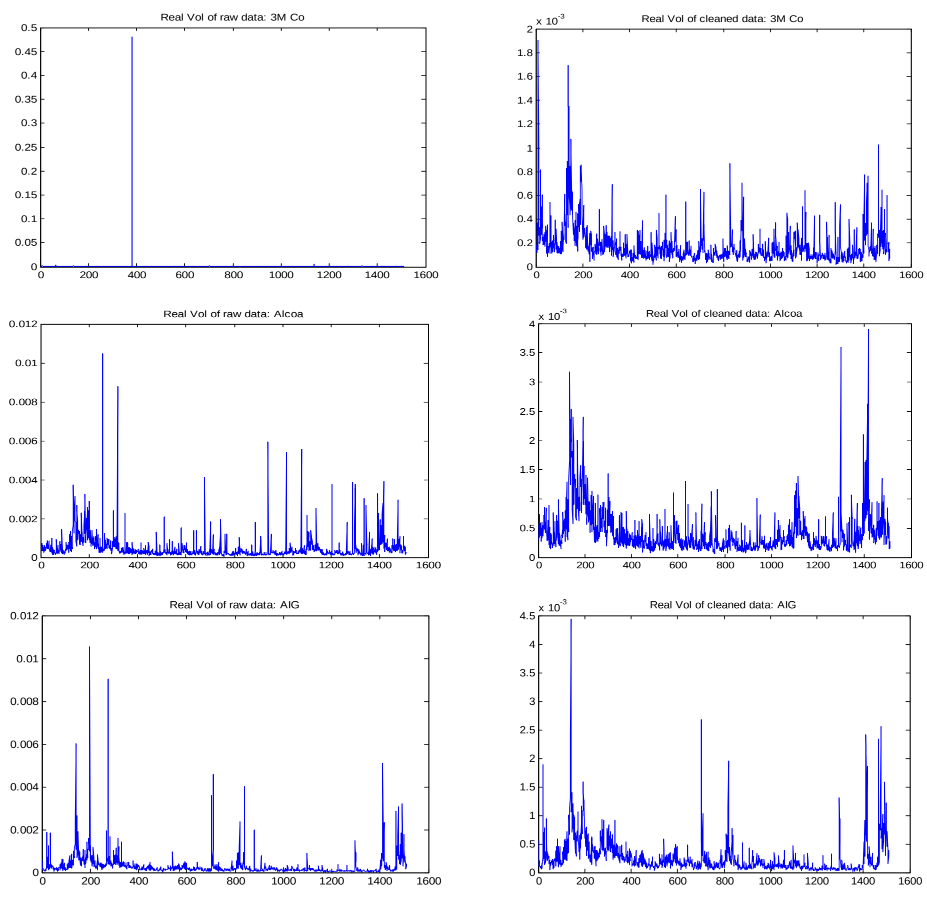

Figure 3: Preprocessing the data. Left: Realized volatility of $r_{t,j}^{OLD}$. Right: Realized volatility of $r_{t,j}^{NEW}$.

We have suggested that L can be estimated by testing the significance of $\{\omega_{m,h}\}_{h=1}^{L}$. However, the computation of these autocovariances requires the prior knowledge of L. We circumvent this vicious circle by using the following information criterion to obtain an initial guess of L:

$$\widehat{L} = \underset{0 \leq l \leq H-1}{\arg\min} \left\{ \Delta(l) = \frac{1}{T} \sum_{t=1}^{T} \left(K_t^{H,T} - \overline{RV}_t^{(AC,m,l+1)} \right)^2 \right\}, H \propto m^{2/3}, \tag{53}$$

where $\overline{RV}_t^{(AC,m,l+1)}$ is defined as in (42) and:

$$K_t^{H,T} = RV_t^{(Ac,m,1)} + \frac{1}{T} \sum_{s=1}^{T} \sum_{h=2}^{H} \left(1 - \frac{h-1}{H} \right) \left(\gamma_{s,h} + \gamma_{s,-h} \right).$$

To see the intuition underlying the choice of this criterion, note that $\Delta(l)$ satisfies:

$$E\left[\Delta(l)\right] = Var\left(K_t^{H,T} - \overline{RV}_t^{(AC,m,l+1)} \right) + \left[E\left(K_t^{H,T} - \overline{RV}_t^{(AC,m,l+1)} \right) \right]^2$$

where the moments are taken unconditionally. On the one hand, $\overline{RV}_t^{(AC,m,l+1)}$ is obtained by truncating the expression of $\widehat{IV}_t$ to l autocovariance terms and is thus unbiased for IV_t when $l \geq L$. On the other hand, $K_t^{H,T}$ is a smoothed version of $\overline{RV}_t^{(AC,m,H)}$ which is also unbiased for IV_t due to $L < H \propto m^{2/3}$. Hence $E\left(K_t^{H,T} - \overline{RV}_t^{(AC,m,l+1)} \right)$ is decreasing in l in the area $l < L$ and equals zero in the area $l \geq L$. As the variance of $K_t^{H,T} - \overline{RV}_t^{(AC,m,l+1)}$ is increasing in l, there is a trade-off between bias and variance that results in a L-shaped curve $\Delta(l)$. See figure 4.1.

The initial guess $\widehat{L}$ given by (53) is used to compute the estimators of $\omega_{m,h}, h = 1, ..., \widehat{L}+1$. If the significance test indicates that the last two or three noise autocovariances are not significantly different from zero, then the initial guess becomes our final estimator. Otherwise, we increment $\widehat{L}$ by $+1$ and repeat the process until the significance test fails to reject the null for the last autocovariance.

7.2 Empirical Results

We follow three basic steps in conducting this empirical study. In the first step, we estimate the memory parameter L. Next, the estimator $\widehat{L}$ is used to compute the estimators of $\{\omega_{m,h}\}_{h=1}^{L}$, α and δ along with the relevant Student-t statistics. Finally, we compute the shrinkage estimator $K_t^{\varpi^*}$.

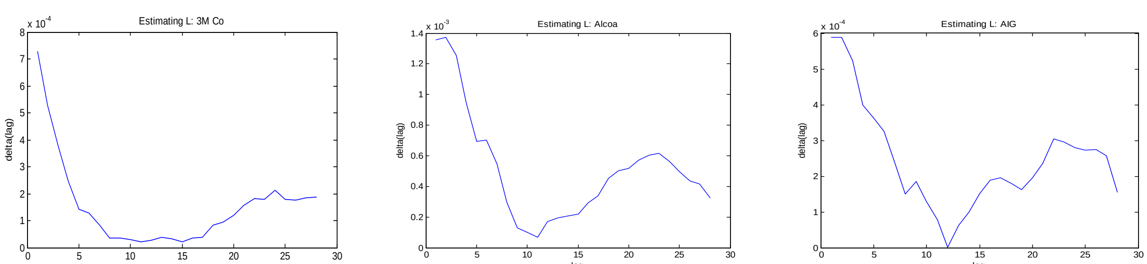

Figure 4.1: Plot of $\Delta(l)$ against l. The minimum of $\Delta(l)$ is used as the first guess of L.

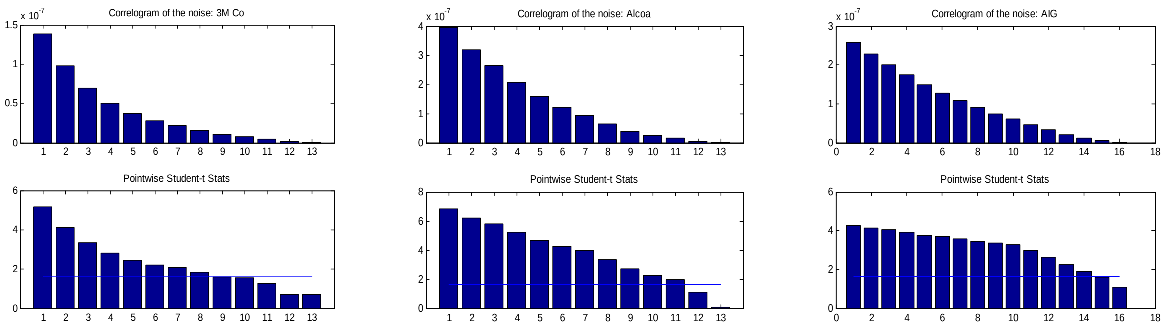

Figure 4.2: The correlogram of the noise $\{\omega_{m,h}\}_{h=0}^{L}$ (top) and the pointwise associated Student stats (bottom).

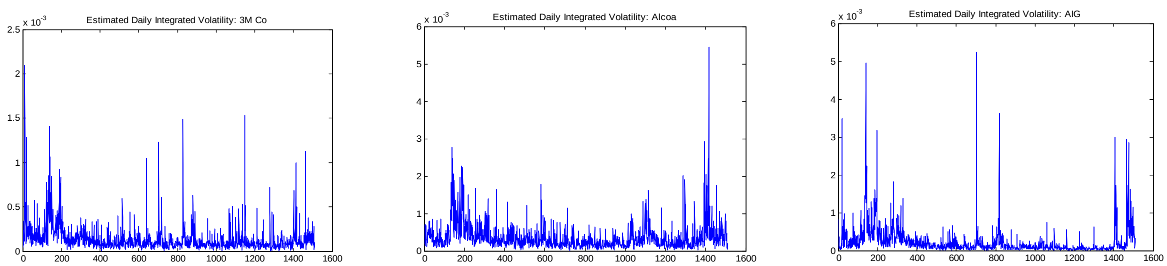

Figure 4.3: Estimated Daily Integrated Volatility $K_t^{\varpi^*}$.

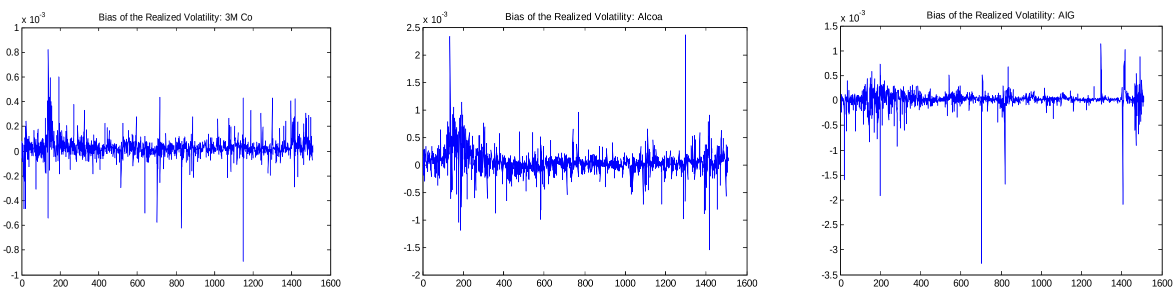

Figure 4.4: Estimated Bias of the Realized Volatility $RV^{(m)}$.

Figure 4: Estimation Results for 3M Co, Alcoa and AIG.

Figure 4.1. shows the plots of $\Delta(l)$ against L while Figure 4.2 shows the estimated noise autocovariances along with the significance tests for the assets 3M Co, Alcoa and AIG. This figures suggets that the initial guess of L slightly overestimates the value predicted by the significance test. The estimated values of L for the other assets are displayed in Table 5. For all the stocks considered, our results suggest that the noise is L-dependent with values of L lying between 5 minutes (American Express) and 14 minutes (AIG and General Electric). The finding that the noise is autocorrelated is not new in the literature[10]. However, we contribute to the discussion by showing that there is a vicious circle raised by the determination of L and we propose a way to solve this.

Figure 4.3 and Figure 4.4 show respectively the time series of $K_t^{\varpi^*}$ and the resulting estimated bias of the RV $\widetilde{b}_t = RV^{(m)} - K_t^{\varpi^*}$. This alternative formula is preferred for the bias because it has less variance compared to the natural method of moment estimator $\widehat{b}_t^{(m)}$ given in (41). To compute the realized kernels, we set $H = 30$ for the bandwidth except for the American Express index (AXP) which necessitates $H = 10$. These bandwidth values appear to produce better results than $(390)^{2/3} \simeq 53$. Figure 4.4 suggests that the sign of $\widetilde{b}_t$ is not constant through time. It turns out that when the correlogram is positive as we found for 3M Co, Alcoa and AIG, a negative bias can only be due to a negative correlation between the noise and the latent return. This suggests that either β_0 or β_1 is negative.

	$\widehat{L}$	$\widehat{\alpha}$ ($\widehat{\sigma}_{\widehat{\alpha}}$)	$\widehat{\delta}$
3M Co (MMM)	12	0.5301 (0.1841)	0.4165
Alcoa Inc (AA)	12	0.4797 (0.0278)	0.4165
American International Group (AIG)	14	0.4715 (0.0100)	0.4423
Americal Express (AXP)	4	0.3151 (0.0087)	0.2324
Dupont and Dupont (DD)	12	0.4805 (0.0111)	0.4165
Walt Disney (DIS)	9	0.4773 (0.0132)	0.3683
General Electric (GE)	14	0.5303 (0.0267)	0.4423
General Motors (GM)	13	0.5117 (0.0667)	0.4299
IBM	12	0.4802 (0.0126)	0.4165
Intel Corp. (INTC)	11	0.5048 (0.0239)	0.4019
Hewlett-Packard (HPQ)	12	0.4942 (0.0146)	0.4165
Microsoft (MSFT)	11	0.4960 (0.0212)	0.4019

Table 5: Estimates of L, α and δ for twelve stocks listed in the DJI. $\widehat{\sigma}_{\widehat{\alpha}}$ is the estimated standard deviation of $\widehat{\alpha}$

Table 5 shows the estimates $\widehat{\alpha}$ and $\widehat{\delta}$. It is seen that $\widehat{\delta} < \widehat{\alpha} < 2/3$ for all the assets. In our framework, the fact that the inequality $\widehat{\delta} < \widehat{\alpha}$ is satisfied indicates that the noise process has finite variance, while $\widehat{\delta} < 2/3$ indicates that the estimator of Barndorff-Nielsen and al. (2008a) delivers its best performance at the highest available frequency. Finally, note that the value of $\widehat{\delta}$ can still be used as a measure of persistence of the microstructure noise even if assumption E4 is not satisfied.

8 Conclusion

This paper proposes a flexible semi-parametric model for the market microstructure noise. We specify the microstructure noise as the sum of two terms. The first term is correlated with the latent return and the second term is exogenous. The exogenous noise is modeled as an L-dependent process, where L is allowed to increase with the frequency at which the prices are recorded. In light of this model, we study the properties of common realized measures that aim to estimate the integrated volatility.

We propose a new shrinkage realized kernels which is an optimal linear combination of the consistent realized kernels of Barndorff-Nielsen and al (2008a) and an unbiased estimator constructed for this purpose. It is shown theoretically that the shrinkage estimator has lower variance than the consistent estimator in small samples while both estimators are asymptotically equivalent in large samples. The Monte Carlo simulations show that the relative efficiency gain of the shrinkage realized kernels over the standard realized kernel is substantial in situations where the variance of the microstructure noise is small. When the variance of the noise is large, the inconsistent estimator is markedly dominated.

Finally, we propose a framework to assess the true values of the noise parameters via the observed returns. Unfortunately, the endogeneity parameters are not identified. Our empirical findings about the noise confirm the conclusions of Hansen and Lunde (2006): there is strong evidence that the noise is autocorrelated and correlated with the latent returns. If our Assumption E3-E3 are true, then the rate at which L increases with the sampling frequency is in general slower than $\sqrt{m}$.

Acknowledgement 6 *We wish to thank Torben Andersen, Valentina Corradi, Prosper Dovonon, Rene Garcia, Silvia Goncalves, Ilze Kalnina, Nour Meddahi, the Associate Editor and two anonymous referee for helpful comments. An earlier draft of this paper has circulated under the title "Assessing the Nature of Pricing Inefficiencies via Realized Measures".*

Notes

[1]See Andersen, Bollerslev, Diebold and Labys (2000); Andersen, Bollerslev, Diebold and Ebens (2001).

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

[3]See e.g Barndorff-Nielsen, Graversen, Jacod and Shephard (2006).

[4]In the current context, an endogenous noise is a noise that is correlated with the efficient price or return.

[5]See BNHLS (2007) for the treatment of these end effects in practice.

[6]See Ait-Sahalia, Mykland and Zhang (2005) and Bandi and Russell (2008) for optimal sampling frequencies of some inconsistent estimators.

[7]When the data are non equally spaced, the expressions of the autocorrelation estimators are more tedious. See for example Ubukata and Oya (2009).

[8]The data we use in this paper have been purchased from a private provider who has ensured its accuracy by comparision with three other independent financial data providers. Please see Section 9 for the preprocessing details.

[9]For quote data, BNHLS (2008b) suggest to delete entries for which the spread is more that 50 times the median spread on that day.

[10]See for example Hansen and Lunde (2006) and Ubukata and Oya (2009).

References

[1] Ait-Sahalia, Y., Mykland, P.A., Zhang, L. (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., Mykland, P.A., Zhang, L. (2006) "Ultra high frequency volatility estimation with dependent microstructure noise", *NBER Working Paper* No. 11380.

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

[4] Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2002) "Parametric and nonparametric measurement of volatility", In Y. Aït-Sahalia and LP Hansen (Eds) *Handbook of Financial Econometrics*, Elsevier, North-Holland.

[5] Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001) "The distribution of realized stock return volatility", *Journal of Financial Economics* 61, 43-76.

[6] Andersen, T. G., Bollerslev, T. and Diebold, F. X. and Labys, P. (2000) "Great realizations", *Risk Magazine* 13, 105-108.

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

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

[9] Andrew, D.W.K. and Monahan, J.C. (1992) "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator", *Econometrica* 60, 953-966.

[10] Bandi, F. M. and Russell J. R. (2006) "Separating microstructure noise from volatility" *Journal of Financial Economics* 79 (2006) 655–692.

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

[12] Barndorff-Nielsen, O. E., Graversen, S.E., Jacod, J. and Shephard, N. (2006) "Limit theorems for realized power variation in econometrics", *Econometric Theory* 22, 677-719.

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

[14] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008b) "Realised kernels in practice: trades and quotes", *Econometrics Journal*, forthcoming.

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

[16] Beveridge, S. and Nelson, C. (1981) "A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to the measurement of the business cycle", *Journal of Monetary Economics* 7, 151-174.

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

[18] Corradi, V., Distaso, W., and Swanson, N. R. (2008) "Predictive density estimators for daily volatility based on the use of realized measures", *Journal of Econometrics,* forthcoming.

[19] Eraker, B. (2004) "Do stock prices and volatility jump? reconciling evidence from spot and option prices ", *Journal of Finance*, 59, 1367-1403.

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

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

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

[23] Gonçalves, S. and Meddahi, N. (2009) "Boostrapping realized volatility", *Econometrica,* 77, 283-306.

[24] Hansen, B. E. (2007) "Least squares model averaging" *Econometrica* 75, 1175-1189.

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

[26] 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.

[27] Hasbrouck, J. (1996) "Modelling market microstructure time series", In G.S. Maddala and C.R. Rao (Eds) *Handbook of Statistics.* Elsevier, North-Holland.

[28] 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.

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

[30] Jacod, J. (1994) "Limit of random measures associated with the increments of a brownian semimartingale" No 120, *Laboratoire de Probabilities, Université Pierre et Marie Curie, Paris.*

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

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

[33] Kalnina, I. and Linton, O. (2008) "Estimating quadratic variation consistently in the presence of correlated measurement error", *Journal of Econometrics,* 147, 47-59.

[34] Meddahi, N. (2002) "A theoretical comparison between integrated and realized volatility", *Journal of Applied Econometrics* 17, 475-508.

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

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

[37] Podolskij, M. and Vetter, M. (2006) "Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps", Bernoulli 15(3), 634-658.

[38] Politis, D. N., Romano, J. P. and Wolf, M. (1997) "Subsampling for heteroskedatic time series", Journal of Econometrics 81, 281-317.

[39] Roll, R. (1984) "A simple measure of the effective bid-ask spread in an efficient market", *Journal of Finance 39,* 1127-1139.

[40] Stein, C. (1956), "Inadmissibility of the usual estimator for the mean of a multivariate distribution", Proc. Third Berkeley Symp. Math. Statist. Prob., 1, pp. 197–206 (Univ. of Calif. Press, 1956)

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

[42] Stoll, H. R. (2000) "Frictions", *Journal of Finance* 55, 1479-1514.

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

[44] Watson, M. (1986) "Univariate detrending methods with stochastic trend", *Journal of Monetary Economics* 18, 49-75.

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

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

Appendix: Proofs

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

Lemma 7 *Assume that* $r_{t,j} = (1 + a_{t,j}) r^*_{t,j} - a_{t,j-1} r^*_{t,j-1} + (\varepsilon_{t,j} - \varepsilon_{t,j-1})$ *for some deterministic sequence* $\{a_{t,j}\}$, $j = 1, ..., m$. *Let* $\widetilde{r}_{t,k}$ *be the serie of non-overlapping sums of* q *consecutive observations of* $r_{t,j}$*:*

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

for $k = 1, ..., m_q$ *and some positive integer* $q \geq 1$ *such that* $m_q = m/q$, *with the convention that* $\sum_{j=qk-q+1}^{qk-1} r^*_{t,j} = 0$ *if* $q = 1$. *Then we have:*

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

Proof of Lemma 7:

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

where

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

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

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

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

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

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

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

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

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

Only squared terms have nonzero expectation:

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

where $\omega_{m,q} = E\left[\varepsilon_{t,j}\varepsilon_{t,j-q}\right]$ is independent of t and j. Also, all the terms involved in the expression of $RV^{(m_q)}$ are uncorrelated and thus:

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

where

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

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

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

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

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

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

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

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

Hence:

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

The following Lemma will be used in the proof of Theorem 2.

Lemma 8 *Under the assumptions of Theorem 2, we have:*

$$
\begin{aligned}
E\left[RV_t^{(AC,m,1)}\right] &= IV_t + \left(2a_{t,m} + a_{t,m}^2\right)\sigma_{t,m}^{*2} - \left(2a_{t,0} + a_{t,0}^2\right)\sigma_{t,0}^{*2} \\
Var\left[RV_t^{(AC,m,1)}\right] &= 2\textstyle\sum_{j=1}^m \sigma_{t,j}^{*4} + 4\sum_{j=1}^m \left(1 + a_{t,j} + a_{t,j}a_{t,j-1}\right)^2 \sigma_{t,j}^{*2}\sigma_{t,j-1}^{*2} \\
&\quad + 4\textstyle\sum_{j=1}^m \left(1 + a_{t,j}\right)^2 a_{t,j-2}^2 \sigma_{t,j}^{*2}\sigma_{t,j-2}^{*2} + 8\omega_0 \sum_{j=1}^m \left(1 + a_{t,j}\right)^2 \sigma_{t,j}^{*2} \\
&\quad + 8\omega_0 \textstyle\sum_{j=1}^m a_{t,j}^2 \sigma_{t,j}^{*2} + 8m\omega_0^2 + 2\left(E\left[\varepsilon_{t,j}^4\right] - \omega_0^2\right) + 2\left(2a_{t,0} + a_{t,0}^2\right)^2 \sigma_{t,0}^{*4} \\
&\quad + 2\left(2a_{t,m} + a_{t,m}^2\right)^2 \sigma_{t,m}^{*4} + 2\left(2a_{t,m} + a_{t,m}^2\right)\sigma_{t,m}^{*4} + 4a_{t,-1}^2 a_{t,0}^2 \sigma_{t,-1}^{*2}\sigma_{t,0}^{*2} \\
&\quad - 8a_{t,m-1}a_{t,m}\left(1 + a_{t,m} + a_{t,m}a_{t,m-1}\right)\sigma_{t,m-1}^{*2}\sigma_{t,m}^{*2} \\
&\quad + 4a_{t,m-1}^2 a_{t,m}^2 \sigma_{t,m-1}^{*2}\sigma_{t,m}^{*2} + 8\omega_0\left(\sigma_{t,m-1}^{*2} - \sigma_{t,0}^{*2}\right) \\
&\quad + 8\omega_0\left(a_{t,-1}^2\sigma_{t,-1}^{*2} + 2a_{t,0}^2\sigma_{t,0}^{*2} + a_{t,m}\sigma_{t,m}^{*2}\right) \\
&\quad - 8\omega_0\left(a_{t,m-1}\sigma_{t,m-1}^{*2} + a_{t,m-1}^2\sigma_{t,m-1}^{*2}\right).
\end{aligned}
$$

Proof of Lemma 8: We first note that:

$$
\begin{aligned}
RV_t^{(AC,m,1)} &= \sum_{j=1}^m r_{t,j}^2 + 2\sum_{j=1}^m r_{t,j}r_{t,j-1} \\
&= (I) + (II) + (III) + (IV) + (V) + (VI) + (VII) + (VIII) + (IX),
\end{aligned}
$$

where

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

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

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

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

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

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

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

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

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

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

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

The calculation of that variance is simplified by noting that only the terms (IV) to (IX) are possibly correlated. Thus we have:

$$
\begin{aligned}
Var((I)) &= 2\textstyle\sum_{j=1}^m \sigma_{t,j}^{*4} + 2\left(2a_{t,0} + a_{t,0}^2\right)^2 \sigma_{t,0}^{*4} + 2\left(2a_{t,m} + a_{t,m}^2\right)^2 \sigma_{t,m}^{*4} \\
&\quad + 2\left(2a_{t,m} + a_{t,m}^2\right)\sigma_{t,m}^{*4} \\
Var((II)) &= 4\textstyle\sum_{j=1}^m \left(1 + a_{t,j} + a_{t,j}a_{t,j-1}\right)^2 \sigma_{t,j}^{*2}\sigma_{t,j-1}^{*2} + 4a_{t,-1}^2 a_{t,0}^2 \sigma_{t,-1}^{*2}\sigma_{t,0}^{*2} \\
&\quad + 4a_{t,m-1}^2 a_{t,m}^2 \sigma_{t,m-1}^{*2}\sigma_{t,m}^{*2} \\
&\quad - 8a_{t,m-1}a_{t,m}\left(1 + a_{t,m} + a_{t,m}a_{t,m-1}\right)\sigma_{t,m-1}^{*2}\sigma_{t,m}^{*2} \\
Var((III)) &= 4\textstyle\sum_{j=1}^m \left(1 + a_{t,j}\right)^2 a_{t,j-2}^2 \sigma_{t,j}^{*2}\sigma_{t,j-2}^{*2} \\
Var((IV)) &= 8\omega_0 IV_t + 8\omega_0\left(a_{t,0}^2\sigma_{t,0}^{*2} + a_{t,m}^2\sigma_{t,m}^{*2}\right) + 16\omega_0 a_{t,m}\sigma_{t,m}^{*2} \\
2Cov\left((IV),(V)\right) &= 8\textstyle\sum_{j=1}^m \left(1 + a_{t,j}\right) E\left[\left(\varepsilon_{t,j} - \varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1} - \varepsilon_{t,j-2}\right)\right] E\left(r_{t,j}^{*2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= -8\omega_0 IV_t - 8\omega_0 \textstyle\sum_{j=1}^{m} a_{t,j}\sigma_{t,j}^{*2} \\
2Cov\left((IV),(VI)\right) &= 8\textstyle\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 IV_t + 8\omega_0\sigma_{t,m-1}^{*2} \\
2Cov\left((IV),(VII)\right) &= -8\textstyle\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 \textstyle\sum_{j=1}^{m}\left(1+a_{t,j}\right)^2 \sigma_{t,j}^{*2} \\
2Cov\left((V),(VI)\right) &= 2Cov\left((V),(VII)\right) = \\
&2Cov((V),(VIII)) = 2Cov((V),(IX)) = 0 \\
Var((VI)) &= 8\omega_0 IV_t - 8\omega_0\sigma_{t,0}^{*2} \\
2Cov\left((VI),(VII)\right) &= -8\textstyle\sum_{j=1}^{m-2} a_{t,j} E[\left(\varepsilon_{t,j+1}-\varepsilon_{t,j}\right)\left(\varepsilon_{t,j+2}-\varepsilon_{t,j+1}\right)] E\left(r_{t,j}^{*2}\right) \\
&= 8\omega_0\textstyle\sum_{j=1}^{m} a_{t,j}\sigma_{t,j}^{*2} - 8\omega_0\left(a_{t,m-1}\sigma_{t,m-1}^{*2} + a_{t,m}\sigma_{t,m}^{*2}\right) \\
2Cov((VI),(VIII)) &= 2Cov((VI),(IX)) = 0 \\
Var((VII)) &= 8\omega_0\textstyle\sum_{j=1}^{m} a_{t,j}^2\sigma_{t,j}^{*2} + 8\omega_0\left(a_{t,-1}^2\sigma_{t,-1}^{*2} + a_{t,0}^2\sigma_{t,0}^{*2} - a_{t,m-1}^2\sigma_{t,m-1}^{*2} - a_{t,m}^2\sigma_{t,m}^{*2}\right) \\
Cov((VII),(VIII)) &= Cov((VII),(IX)) = 0 \\
Var((VIII)) &= Var\left[2\textstyle\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)\right] \\
&= 4Var\left[2\textstyle\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_{t,j-1}^2\right] \\
&= 4mE[\varepsilon_{t,j}^4] + 16m\omega_0^2 - 8\omega_0^2 \\
2Cov\left((VIII),(IX)\right) &= 4Cov\left[\textstyle\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right),\ \sum_{k=1}^{m}\left(\varepsilon_{t,k}-\varepsilon_{t,k-1}\right)^2\right] \\
&= -\left(8m-4\right)\left(E[\varepsilon_{t,j}^4] + \omega_0^2\right)
\end{aligned}
$$

since we have:

$$
\begin{aligned}
&E[\left(\varepsilon_{t,j+k}-\varepsilon_{t,j+k-1}\right)^2\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)] = -2\omega_0^2\ \forall\ k \geq 1 \\
&E[\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)^3\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)] = -E[\varepsilon_{t,j}^4] - 3\omega_0^2\ (\text{for } k=j) \\
&E[\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)^3] = -E[\varepsilon_{t,j}^4] - 3\omega_0^2\ (\text{for } k=j-1) \\
&E[\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)\left(\varepsilon_{t,j-k-1}-\varepsilon_{t,j-k-2}\right)^2] = -2\omega_0^2\ \forall\ k \geq 1 \\
&\Rightarrow E\left[\left(\textstyle\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)\right)\left(\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)^2\right)\right] \\
&\qquad = (-2m+1)E[\varepsilon_{t,j}^4] + (-2m^2-2m+1)\omega_0^2
\end{aligned}
$$

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

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

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

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

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

The sum of all these terms gives:

$$
\begin{aligned}
Var\left[RV_t^{(AC,m,1)}\right] &= 2\textstyle\sum_{j=1}^{m}\sigma_{t,j}^{*4} + 4\sum_{j=1}^{m}\left(1+a_{t,j}+a_{t,j}a_{t,j-1}\right)^2\sigma_{t,j}^{*2}\sigma_{t,j-1}^{*2} \\
&+4\textstyle\sum_{j=1}^{m}\left(1+a_{t,j}\right)^2 a_{t,j-2}^2\sigma_{t,j}^{*2}\sigma_{t,j-2}^{*2} + 8\omega_0\sum_{j=1}^{m}\left(1+a_{t,j}\right)^2\sigma_{t,j}^{*2} \\
&+8\omega_0\textstyle\sum_{j=1}^{m} a_{t,j}^2\sigma_{t,j}^{*2} + 8m\omega_0^2 + 2\left(E\left[\varepsilon_{t,j}^4\right] - \omega_0^2\right) + 2\left(2a_{t,0}+a_{t,0}^2\right)^2\sigma_{t,0}^{*4} \\
&+2\left(2a_{t,m}+a_{t,m}^2\right)^2\sigma_{t,m}^{*4} + 2\left(2a_{t,m}+a_{t,m}^2\right)\sigma_{t,m}^{*4} + 4a_{t,-1}^2a_{t,0}^2\sigma_{t,-1}^{*2}\sigma_{t,0}^{*2} \\
&-8a_{t,m-1}a_{t,m}\left(1+a_{t,m}+a_{t,m}a_{t,m-1}\right)\sigma_{t,m-1}^{*2}\sigma_{t,m}^{*2} \\
&+4a_{t,m-1}^2a_{t,m}^2\sigma_{t,m-1}^{*2}\sigma_{t,m}^{*2} + 8\omega_0\left(\sigma_{t,m-1}^{*2} - \sigma_{t,0}^{*2}\right) \\
&+8\omega_0\left(a_{t,-1}^2\sigma_{t,-1}^{*2} + 2a_{t,0}^2\sigma_{t,0}^{*2} + a_{t,m}\sigma_{t,m}^{*2}\right)
\end{aligned}
$$

$$-8\omega_0\left(a_{t,m-1}\sigma^{*2}_{t,m-1}+a^2_{t,m-1}\sigma^{*2}_{t,m-1}\right).\blacksquare$$

Proof of Theorem 1: Substituting for $a_{t,j}=\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,j}}$ in Lemma 7, we get for the expectation:

$$E\left[RV_t^{(m_q)}\right]=IV_t+2m_q\left(\omega_0-\omega_{m,q}\right)+2\sum_{k=1}^{m_q}\left[\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}+\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2\right]\sigma^{*2}_{t,qk}$$
$$+\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,0}}\right)^2\sigma^{*2}_{t,0}-\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,m}}\right)^2\sigma^{*2}_{t,m}$$

Hence:

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

For the variance, we have:

$$Var\left[RV^{(m_q)}\right]=2\sum_{k=1}^{m_q}\left[\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2+\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2\right]^2\sigma^{*4}_{t,qk}$$
$$+2\sum_{k=1}^{m_q}\left(\sum_{l=qk-q+1}^{qk-1}\sum_{j=qk-q+1}^{qk-1}\sigma^{*2}_{t,j}\sigma^{*2}_{t,l}\right)+Var\left[\sum_{k=1}^{m_q}\left(\varepsilon_{t,kq}-\varepsilon_{t,kq-q}\right)^2\right]$$
$$+4\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2\sigma^{*2}_{t,j}\sigma^{*2}_{t,qk}$$
$$+4\sum_{k=1}^{m_q}\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk-q}}\right)^2\sigma^{*2}_{t,qk-q}\sigma^{*2}_{t,qk}$$
$$+4\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk-q}}\right)^2\sigma^{*2}_{t,j}\sigma^{*2}_{t,qk-q}$$
$$+8\left(\omega_0-\omega_{m,q}\right)\sum_{k=1}^{m_q}\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2\sigma^{*2}_{t,qk}$$
$$+8\left(\omega_0-\omega_{m,q}\right)\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma^{*2}_{t,j}+8\left(\omega_0-\omega_{m,q}\right)\sum_{k=1}^{m_q}\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk-q}}\right)^2\sigma^{*2}_{t,qk-q}$$
$$+2\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,0}}\right)^4\sigma^{*4}_{t,0}-2\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,m}}\right)^4\sigma^{*4}_{t,m}-4\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,m}}\right)^2\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,m}}\right)^2\sigma^{*4}_{t,m}.$$

In details, we have:

$$2\sum_{k=1}^{m_q}\left[\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2+\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2\right]^2\sigma^{*4}_{t,qk}=$$
$$2\left(1+4\beta_0+8\beta_0^2+8\beta_0^3+4\beta_0^4\right)\sum_{k=1}^{m_q}\sigma^{*4}_{t,qk}+\frac{8\beta_1}{\sqrt{m}}\left(1+4\beta_0+6\beta_0^2+4\beta_0^3\right)\sum_{k=1}^{m_q}\sigma^{*3}_{t,qk}$$
$$+\frac{16\beta_1^2}{m}\left(1+3\beta_0+3\beta_0^2\right)\sum_{k=1}^{m_q}\sigma^{*2}_{t,qk}+\frac{16\beta_1^3}{m\sqrt{m}}\left(1+2\beta_0\right)\sum_{k=1}^{m_q}\sigma^*_{t,qk}+\frac{8\beta_1^4}{qm}.$$

$$4\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2\sigma^{*2}_{t,j}\sigma^{*2}_{t,qk}=4\left(1+2\beta_0+\beta_0^2\right)\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma^{*2}_{t,j}\sigma^{*2}_{t,qk}$$
$$+\frac{8\beta_1}{\sqrt{m}}\left(1+\beta_0\right)\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma^{*2}_{t,j}\sigma^*_{t,qk}+\frac{4\beta_1^2}{m}IV_t.$$

$$4\sum_{k=1}^{m_q}\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk}}\right)^2\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma^*_{t,qk-q}}\right)^2\sigma^{*2}_{t,qk-q}\sigma^{*2}_{t,qk}=$$
$$4\beta_0^2\left(1+\beta_0\right)^2\sum_{k=1}^{m_q}\sigma^{*2}_{t,qk-q}\sigma^{*2}_{t,qk}+\frac{8\beta_1}{\sqrt{m}}\beta_0\left(1+\beta_0\right)^2\sum_{k=1}^{m_q}\sigma^*_{t,qk-q}\sigma^{*2}_{t,qk}$$
$$+\frac{8\beta_1}{\sqrt{m}}\beta_0^2\left(1+\beta_0\right)\sum_{k=1}^{m_q}\sigma^{*2}_{t,qk-q}\sigma^*_{t,qk}+\frac{16\beta_1^2}{m}\beta_0\left(1+\beta_0\right)\sum_{k=1}^{m_q}\sigma^*_{t,qk-q}\sigma^*_{t,qk}$$

$+\frac{4\beta_1^2}{m}(1+\beta_0)^2\sum_{k=1}^{m_q}\sigma_{t,qk}^{*2}+\frac{4\beta_1^2}{m}\beta_0^2\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*2}$
$+\frac{8\beta_1^3}{m\sqrt{m}}(1+\beta_0)\sum_{k=1}^{m_q}\sigma_{t,qk}^{*}+\frac{8\beta_1^3}{m\sqrt{m}}\beta_0\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*}+\frac{4\beta_1^4}{qm}.$

$4\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,qk-q}^{*}}\right)^2\sigma_{t,j}^{*2}\sigma_{t,qk-q}^{*2}=4\beta_0^2\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk-q}^{*2}$
$+\frac{8\beta_1}{\sqrt{m}}\beta_0\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk-q}^{*}+\frac{4\beta_1^2}{m}IV_t.$

$8(\omega_0-\omega_{m,q})\sum_{k=1}^{m_q}\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,qk}^{*}}\right)^2\sigma_{t,qk}^{*2}=8(\omega_0-\omega_{m,q})(1+\beta_0)^2\sum_{k=1}^{m_q}\sigma_{t,qk}^{*2}$
$+\frac{16\beta_1}{m}(1+\beta_0)\sum_{k=1}^{m_q}\sigma_{t,qk}^{*}+\frac{8\beta_1^2}{q}(\omega_0-\omega_{m,q}).$

$8(\omega_0-\omega_{m,q})\sum_{k=1}^{m_q}\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,qk-q}^{*}}\right)^2\sigma_{t,qk-q}^{*2}=8(\omega_0-\omega_{m,q})\beta_0^2\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*2}$
$+\frac{16\beta_1}{\sqrt{m}}(\omega_0-\omega_{m,q})\beta_0\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*}+\frac{8\beta_1^2}{q}(\omega_0-\omega_{m,q}).$

Also, $Var\left[\sum_{k=1}^{m_q}(\varepsilon_{t,kq}-\varepsilon_{t,kq-q})^2\right]=O(m_q)$. We thus define:

$$\kappa=\frac{1}{m_q}Var\left[\sum_{k=1}^{m_q}(\varepsilon_{t,kq}-\varepsilon_{t,kq-q})^2\right].$$

The sum of all these terms yields:

$Var\left[RV^{(m_q)}\right]=m_q\kappa+\frac{16\beta_1^2}{q}(\omega_0-\omega_{m,q})+\frac{12\beta_1^4}{qm}$
$+8\left[\frac{(3+5\beta_0)\beta_1^3}{m\sqrt{m}}+\frac{2(1+\beta_0)\beta_1}{\sqrt{m}}+\frac{2\beta_0\beta_1}{\sqrt{m}}(\omega_0-\omega_{m,q})+\frac{\beta_0\beta_1^3}{m\sqrt{m}}\right]\sum_{k=1}^{m_q}\sigma_{t,qk}^{*}$
$+4\left(1+2\beta_0+2\beta_0^2\right)\left[\frac{7\left(1+2\beta_0+2\beta_0^2\right)\beta_1^2}{m}+2(\omega_0-\omega_{m,q})\right]\sum_{k=1}^{m_q}\sigma_{t,qk}^{*2}$
$+8\frac{\left(1+4\beta_0+6\beta_0^2+4\beta_0^3\right)\beta_1}{\sqrt{m}}\sum_{k=1}^{m_q}\sigma_{t,qk}^{3}+2\sum_{k=1}^{m_q}\left(\sum_{j=qk-q+1}^{qk}\sigma_{t,j}^{*2}\right)^2$
$+\frac{16\beta_0(1+\beta_0)\beta_1^2}{m}\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*}\sigma_{t,qk}^{*}+\frac{8(1+\beta_0)\beta_1}{\sqrt{m}}\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk}^{*}$
$+\frac{8\beta_0^2(1+\beta_0)\beta_1}{\sqrt{m}}\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*2}\sigma_{t,qk}^{*}+\frac{8\beta_0(1+\beta_0)^2\beta_1}{\sqrt{m}}\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*}\sigma_{t,qk}^{*2}$
$+\frac{8\beta_0\beta_1}{\sqrt{m}}\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk-q}^{*}+2\left(4\beta_0+8\beta_0^2+8\beta_0^3+4\beta_0^4\right)\sum_{k=1}^{m_q}\sigma_{t,qk}^{4}$
$+4\left(2\beta_0+\beta_0^2\right)\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk}^{*2}+4\beta_0^2\sum_{k=1}^{m_q}\sum_{j=qk-q+1}^{qk-1}\sigma_{t,j}^{*2}\sigma_{t,qk-q}^{*2}$
$+4\beta_0^2(1+\beta_0)^2\sum_{k=1}^{m_q}\sigma_{t,qk-q}^{*2}\sigma_{t,qk}^{*2}+8(\omega_0-\omega_{m,q})\left(\beta_0^2+\frac{2\beta_0\beta_1}{\sqrt{m}}\right)+Q_m.$

where

$Q_m=2\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,0}^{*}}\right)^4\sigma_{t,0}^{*4}-2\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,m}^{*}}\right)^4\sigma_{t,m}^{*4}-4\left(\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,m}^{*}}\right)^2\left(1+\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,m}^{*}}\right)^2\sigma_{t,m}^{*4}$
$+\frac{8\beta_1^3\beta_0}{m\sqrt{m}}\left(\sigma_{t,0}^{*}-\sigma_{t,qm_q}^{*}\right)+\frac{4\beta_1^2\beta_0^2}{m}\left(\sigma_{t,0}^{*2}-\sigma_{t,qm_q}^{*2}\right)=O(m^{-1}).\blacksquare$

Proof of Theorem 2: Substituting for $a_{t,j}=\beta_0+\frac{\beta_1}{\sqrt{m}\sigma_{t,j}}$ in Lemma 8, yield:

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

For the variance, we have:

$Var\left[RV_t^{(AC,m,1)}\right]=2\sum_{j=1}^{m}\sigma_{t,j}^{*4}+4\sum_{j=1}^{m}\left(1+a_{t,j}+a_{t,j}a_{t,j-1}\right)^2\sigma_{t,j}^{*2}\sigma_{t,j-1}^{*2}$
$+4\sum_{j=1}^{m}(1+a_{t,j})^2a_{t,j-2}^2\sigma_{t,j}^{*2}\sigma_{t,j-2}^{*2}+8\omega_0\sum_{j=1}^{m}(1+a_{t,j})^2\sigma_{t,j}^{*2}$
$+8\omega_0\sum_{j=1}^{m}a_{t,j}^2\sigma_{t,j}^{*2}+8m\omega_0^2+2\left(E\left[\varepsilon_{t,j}^4\right]-\omega_0^2\right)+R_m.$

where

$R_m = 2\left(2a_{t,0}+a_{t,0}^2\right)^2\sigma_{t,0}^{*4} + 2\left(2a_{t,m}+a_{t,m}^2\right)^2\sigma_{t,m}^{*4} + 2\left(2a_{t,m}+a_{t,m}^2\right)\sigma_{t,m}^{*4}$
$\quad +4a_{t,-1}^2a_{t,0}^2\sigma_{t,-1}^{*2}\sigma_{t,0}^{*2} - 8a_{t,m-1}a_{t,m}\left(1+a_{t,m}+a_{t,m}a_{t,m-1}\right)\sigma_{t,m-1}^{*2}\sigma_{t,m}^{*2}$
$\quad +4a_{t,m-1}^2a_{t,m}^2\sigma_{t,m-1}^{*2}\sigma_{t,m}^{*2} + 8\omega_0\left(\sigma_{t,m-1}^{*2}-\sigma_{t,0}^{*2}\right)$
$\quad +8\omega_0\left(a_{t,-1}^2\sigma_{t,-1}^{*2} + 2a_{t,0}^2\sigma_{t,0}^{*2} + a_{t,m}\sigma_{t,m}^{*2} - a_{t,m-1}\sigma_{t,m-1}^{*2} - a_{t,m-1}^2\sigma_{t,m-1}^{*2}\right).$

$R_m = 4\beta_1^4 + 3\beta_1^2\omega_0 + O\left(\beta_0\beta_1 m^{-1/2}\right).$

$4\sum_{j=1}^m\left(1+a_{t,j}+a_{t,j}a_{t,j-1}\right)^2\sigma_{t,j}^{*2}\sigma_{t,j-1}^{*2} = \frac{4\beta_1^4}{m} + \frac{8\beta_0\beta_1^3}{m\sqrt{m}}\sum_{j=1}^m\sigma_{t,j}^*$
$\quad +\frac{8\beta_1^2}{m}\left(1+\frac{\beta_0\beta_1}{\sqrt{m}}\right)\sum_{j=1}^m\sigma_{t,j-1}^* + \frac{8\beta_1^2}{m}\left(1+2\beta_0+2\beta_0^2\right)\sum_{j=1}^m\sigma_{t,j}^*\sigma_{t,j-1}^*$
$\quad +\frac{4\beta_0^2\beta_1^2}{m}\sum_{j=1}^m\sigma_{t,j}^{*2} + \frac{4\beta_1^2(1+\beta_0)^2}{m}\sum_{j=1}^m\sigma_{t,j-1}^{*2} + \frac{8\beta_0\beta_1}{\sqrt{m}}\left(1+\beta_0+\beta_0^3\right)\sum_{j=1}^m\sigma_{t,j}^{*2}\sigma_{t,j-1}^*$
$\quad +\frac{8\beta_1}{\sqrt{m}}\left(1+2\beta_0+2\beta_0^2+\beta_0^3\right)\sum_{j=1}^m\sigma_{t,j-1}^{*2}\sigma_{t,j}^* + 4\left(1+2\beta_0+3\beta_0^2+2\beta_0^3+\beta_0^4\right)\sum_{j=1}^m\sigma_{t,j-1}^{*2}\sigma_{t,j}^{*2}.$

$4\sum_{j=1}^m\left(1+a_{t,j}\right)^2a_{t,j-2}^2\sigma_{t,j}^{*2}\sigma_{t,j-2}^{*2} = \frac{4\beta_1^4}{m} + \frac{8\beta_0\beta_1^3}{m\sqrt{m}}\sum_{j=1}^m\sigma_{t,j-2}^*$
$\quad +\frac{8(1+\beta_0)\beta_1^3}{m\sqrt{m}}\sum_{j=1}^m\sigma_{t,j}^* + \frac{4\beta_0^2\beta_1^2}{m}\sum_{j=1}^m\sigma_{t,j-2}^{*2} + \frac{8\beta_1(1+\beta_0)\beta_0^2}{\sqrt{m}}\sum_{j=1}^m\sigma_{t,j-2}^{*2}\sigma_{t,j}^*$
$\quad +\frac{4\beta_1^2(1+\beta_0)^2}{m}\sum_{j=1}^m\sigma_{t,j}^{*2} + \frac{16\beta_0\beta_1^2(1+\beta_0)}{m}\sum_{j=1}^m\sigma_{t,j}^*\sigma_{t,j-2}^*$
$\quad +\frac{8\beta_1\beta_0(1+\beta_0)^2}{\sqrt{m}}\sum_{j=1}^m\sigma_{t,j}^{*2}\sigma_{t,j-2}^* + 4\beta_0^2\left(1+\beta_0\right)^2\sum_{j=1}^m\sigma_{t,j}^{*2}\sigma_{t,j-2}^{*2}.$

$8\omega_0\sum_{j=1}^m\left(1+a_{t,j}\right)^2\sigma_{t,j}^{*2} = 8\omega_0\beta_1^2 + \frac{16\omega_0\beta_1(1+\beta_0)}{\sqrt{m}}\sum_{j=1}^m\sigma_{t,j}^*$
$\quad +8\omega_0\left(1+\beta_0\right)^2\sum_{j=1}^m\sigma_{t,j}^{*2}.$

$8\omega_0\sum_{j=1}^m a_{t,j}^2\sigma_{t,j}^{*2} = 8\omega_0\beta_1^2 + \frac{16\omega_0\beta_1\beta_0}{\sqrt{m}}\sum_{j=1}^m\sigma_{t,j}^* + 8\omega_0\beta_0^2\sum_{j=1}^m\sigma_{t,j}^{*2}.$

Hence:

$Var\left[RV_t^{(AC,m,1)}\right] = 8m\omega_0^2 + 2\sum_{j=1}^m\sigma_{t,j}^{*4} + 2\left(E\left[\varepsilon_{t,j}^4\right]-\omega_0^2\right)$
$+\frac{\beta_1^4+6\beta_1^2\omega_0}{m} + \frac{8\beta_1^4}{m^2} + \frac{8\beta_1}{\sqrt{m}}\left[\frac{(\beta_0+1)^2\beta_1^2}{m} + \frac{\beta_1}{\sqrt{m}} + 2\omega_0\left(1+2\beta_0\right)\right]\sum_{j=1}^m\sigma_{t,j}^*$
$+8\left[\frac{\beta_0^2\beta_1^2}{m} + \left(\frac{\beta_1^2}{m}+\omega_0\right)\left(1+\beta_0\right)^2 + 2\omega_0\beta_0^2\right]\sum_{j=1}^m\sigma_{t,j}^{*2}$
$+\frac{8\beta_1^2}{m}\left(1+2\beta_0+2\beta_0^2\right)\sum_{j=1}^m\sigma_{t,j}^*\sigma_{t,j-1}^* + \frac{16\beta_0\beta_1^2}{m}\left(1+\beta_0\right)\sum_{j=1}^m\sigma_{t,j}^*\sigma_{t,j-2}^*$
$+\frac{8\beta_0\beta_1}{\sqrt{m}}\left(1+\beta_0+\beta_0^3\right)\sum_{j=1}^m\sigma_{t,j}^{*2}\sigma_{t,j-1}^* + \frac{8\beta_1}{\sqrt{m}}\left(1+2\beta_0+2\beta_0^2+\beta_0^3\right)\sum_{j=1}^m\sigma_{t,j-1}^{*2}\sigma_{t,j}^*$
$+\frac{8\beta_1(1+\beta_0)\beta_0^2}{\sqrt{m}}\sum_{j=1}^m\sigma_{t,j-2}^{*2}\sigma_{t,j}^* + \frac{8\beta_1\beta_0(1+\beta_0)^2}{\sqrt{m}}\sum_{j=1}^m\sigma_{t,j}^{*2}\sigma_{t,j-2}^*$
$+4\left(1+2\beta_0+3\beta_0^2+2\beta_0^3+\beta_0^4\right)\sum_{j=1}^m\sigma_{t,j-1}^{*2}\sigma_{t,j}^{*2} + 4\beta_0^2\left(1+\beta_0\right)^2\sum_{j=1}^m\sigma_{t,j}^{*2}\sigma_{t,j-2}^{*2} + \beta_0 O\left(m^{-1/2}\right).\blacksquare$

Proof of Theorem 3:

The result for $K_t\left(r^*\right)$ follows from Theorem 1 of Barndorff-Nielsen and al (2008a). Below, we examine the term $K_t^{BNHLS}\left(r^*,\Delta\varepsilon\right)$:

$$K_t^{BNHLS}\left(r^*,\Delta\varepsilon\right) = \gamma_{t,0}\left(r^*,\Delta\varepsilon\right) + 2\sum_{h=1}^H k\left(\frac{h-1}{H}\right)\gamma_{t,h}\left(r^*,\Delta\varepsilon\right).$$

Let us define $\Phi = (1, k\left(\frac{0}{H}\right), k\left(\frac{1}{H}\right), ..., k\left(\frac{H-1}{H}\right))'$. Then, we have:

$$K_t^{BNHLS}\left(r^*,\Delta\varepsilon\right) = \Phi'\sum_{j=1}^m r_{t,j}^*\begin{pmatrix}\varepsilon_{t,j}-\varepsilon_{t,j-1}\\ 2\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)\\ ...\\ 2\left(\varepsilon_{t,j-H}-\varepsilon_{t,j-H-1}\right)\end{pmatrix}.$$

Note that:

$$\begin{aligned} Var\left[K_t^{BNHLS}(r^*,\Delta\varepsilon)\right] &= Var\left[E\left[K_t^{BNHLS}(r^*,\Delta\varepsilon)\,|\,\{(\varepsilon_{t,j-h}-\varepsilon_{t,j-h-1})\}_{h=0}^{H}\right]\right] \\ &\quad +E\left[Var\left[K_t^{BNHLS}(r^*,\Delta\varepsilon)\,|\,\{(\varepsilon_{t,j-h}-\varepsilon_{t,j-h-1})\}_{h=0}^{H}\right]\right] \\ &= E\left[Var\left[K_t^{BNHLS}(r^*,\Delta\varepsilon)\,|\,\{(\varepsilon_{t,j-h}-\varepsilon_{t,j-h-1})\}_{h=0}^{H}\right]\right] \\ &= IV_t\Phi' Var\left(\Delta\varepsilon^H\right)\Phi, \end{aligned}$$

where $\Delta\varepsilon^H = (\varepsilon_{t,j}-\varepsilon_{t,j-1}, 2(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}), ..., 2(\varepsilon_{t,j-H}-\varepsilon_{t,j-H-1}))$.

We now compute explicitely $Var\left(\Delta\varepsilon^H\right)$:

$$\begin{aligned} E\left[(\varepsilon_{t,j}-\varepsilon_{t,j-1})^2\right] &= 2(\omega_0-\omega_{m,1}) \\ E\left[(\varepsilon_{t,j}-\varepsilon_{t,j-1})(\varepsilon_{t,j-h}-\varepsilon_{t,j-h-1})\right] &= -\omega_{m,h-1}+2\omega_{m,h}-\omega_{m,h+1} \\ E\left[(\varepsilon_{t,j-h}-\varepsilon_{t,j-h-1})(\varepsilon_{t,j-k}-\varepsilon_{t,j-k-1})\right] &= -\omega_{m,h-k-1}+2\omega_{m,h-k}-\omega_{m,h-k+1}, h>k. \end{aligned}$$

Let $\Delta\omega_{m,h} = \omega_{m,h}-\omega_{m,h+1}$ for all h. Then:

$$Var\left(\Delta\varepsilon^H\right) = 2\times$$
$$\begin{pmatrix} \Delta\omega_0 & \bullet & ... & \bullet & \bullet \\ (-\Delta\omega_0+\Delta\omega_{m,1}) & 2\Delta\omega_0 & ... & \bullet & \bullet \\ (-\Delta\omega_{m,1}+\Delta\omega_{m,2}) & 2(-\Delta\omega_0+\Delta\omega_{m,1}) & ... & \bullet & \bullet \\ ... & 2(-\Delta\omega_{m,1}+\Delta\omega_{m,2}) & ... & ... & ... \\ ... & & ... & 2\Delta\omega_0 & \bullet \\ (-\Delta\omega_{m,H-1}+\Delta\omega_{m,H}) & 2(-\Delta\omega_{m,H-2}+\Delta\omega_{m,H-1}) & ... & 2(-\Delta\omega_0+\Delta\omega_{m,1}) & 2\Delta\omega_0 \end{pmatrix}$$

To ease the calculations, a simplified representation of $Var\left(\Delta\varepsilon^H\right)$ is needed. To that end, let us define:

$$\underset{(H+1\times H+1)}{S^0} = \begin{pmatrix} 1 & -1 & \bullet & ... & \bullet \\ -1 & 2 & -1 & ... & ... \\ 0 & -1 & 2 & ... & \bullet \\ ... & ... & ... & ... & -1 \\ 0 & & 0 & -1 & 2 \end{pmatrix}$$

Also let S^h be the symmetric matrix of size $H+1$ with elements $S_{j,k}^h = 1$ if $j=k+h$ or $j=k-h$, $S_{j,k}^h = -1$ if $j=k+h+1$ or $j=k-h-1$, and $S_{j,k}^h = 0$ otherwise. In fact, S^h is the sparse matrix with ones on the h^{th} diagonals and minus ones on the $h+1^{th}$ diagonals. Finally, let $\widetilde{S}^h$ be the matrix S^h with the nonzero elements of the first row and first column replaced by zero. Then we have:

$$\Phi' Var\left(\Delta\varepsilon^H\right)\Phi = 2(\omega_0-\omega_{m,1})\Phi' S^0\Phi + 2\sum_{h=1}^{L}(\omega_{m,h}-\omega_{m,h+1})\Phi'\left(S^h+\widetilde{S}^h\right)\Phi.$$

As $m\to\infty$ and $H=m^b$ for $b\in(0,1)$, we easily check that:

$$\begin{aligned}
\Phi' S^0 \Phi &= \sum_{h=0}^{H}\left(k\left(\frac{h+1}{H}\right)-k\left(\frac{h}{H}\right)\right)^2 \to \frac{1}{H}\int_0^1 k'(x)^2\,dx=\frac{1}{H}.\\
\Phi' S^h\Phi+\Phi'\widetilde{S}^h\Phi &= k\left(\frac{h}{H}\right)-k\left(\frac{h+1}{H}\right)+\frac{4}{H}\sum_{l=0}^{H-h-1}k\left(\frac{l}{H}\right),\\
&= \frac{1}{H}+4\int_0^{1-\frac{h+1}{H}}k(x)\,dx=\frac{1}{H}+2\left[1-\frac{(h+1)^2}{H^2}\right].
\end{aligned}$$

Focusing on the dominant terms, we have:

$$\begin{aligned}
\Phi' Var\left(\Delta\varepsilon^H\right)\Phi &\simeq \frac{2}{H}\sum_{h=0}^{L}\left(\omega_{m,h}-\omega_{m,h+1}\right)+4\sum_{h=1}^{L-1}\left(\omega_{m,h}-\omega_{m,h+1}\right)\left[1-\frac{(h+1)^2}{H^2}\right]\\
&\quad +4\omega_{m,L}\left[1-\frac{(L+1)^2}{H^2}\right],\\
&= \frac{2\omega_0}{H}+4\sum_{h=1}^{L-1}\left(\omega_{m,h}-\omega_{m,h+1}\right)\left[1-\frac{(h+1)^2}{H^2}\right]+4\omega_{m,L}\left[1-\frac{(L+1)^2}{H^2}\right].
\end{aligned}$$

This yields the second result. The remaining term to examine is thus $K_t^{BNHLS}(\Delta\varepsilon)$. We have:

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

$$\begin{aligned}
RV_t^{(AC,m,1)} &= \textstyle\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)^2+2\sum_{j=1}^{m}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)\\
&= -2\textstyle\sum_{j=1}^{m}\varepsilon_{t,j-2}\left(\varepsilon_{t,j}-\varepsilon_{t,j-1}\right)-\varepsilon_{t,0}^2+\varepsilon_{t,m}^2\\
&= 2\textstyle\sum_{j=1}^{m}\varepsilon_{t,j}\left(\varepsilon_{t,j-1}-\varepsilon_{t,j-2}\right)-\varepsilon_{t,0}^2+\varepsilon_{t,m}^2+2\left(\varepsilon_{t,0}\varepsilon_{t,-1}-\varepsilon_{t,m}\varepsilon_{t,m-1}\right).
\end{aligned}$$

And for $h\geq 2$, we have:

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

Summing over H yields:

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

Finally, we have:

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

■

Proof of Theorem 5:

As $\widehat{\omega}_{t,j,h}$ is a linear combination of a finite number of terms of type $\gamma_{t,j,h}$ (see Equations 44 and 45), our Assumption E5 ensures that:

$$
Var\left(T^{-1/2}m^{-1/2}\sum_{t=1}^{T}\sum_{j=1}^{m}\widehat{\omega}_{t,j,h}\right)\to Q_h,\text{as } T\to\infty,
$$

for some Q_h that depends only on h. Next, we note that our Assumptions E1 and E2 replicate the Assumption 1 of Ubukata and Oya (2009), which together with E5 ensure that their Lemma 1 holds, that is:

$$
\frac{(mT)^{1/2}\left(\widehat{\omega}_{m,h}-\omega_{m,h}\right)}{Q_h}\to N\left(0,1\right),\ \text{as } T\to\infty.
$$

By letting T alone go to inifinity, we ensure that $\frac{m}{mT}\to 0$ as mT goes to infinity. Lemma 2 of Ubukata and Oya (2009) then guarantee under the same assumptions that $\widehat{Q}_h\to Q_h$ in L^2. Finally, we replace Q_h by $\widehat{Q}_h$ above to obtain the desired result.