THIELE CENTRE
FOR APPLIED MATHEMATICS IN NATURAL SCIENCE

Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise

Ole E. Barndorff-Nielsen, Peter Reinhard Hansen,
Asger Lunde and Neil Shephard

Research Report

No. 09 | June 2006

Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise

This Thiele Research Report is also Research Report number 478 in the Stochastics Series at Department of Mathematical Sciences, University of Aarhus, Denmark.

Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise*

Ole E. Barndorff-Nielsen
The T.N. Thiele Centre for Mathematics in Natural Science,
Department of Mathematical Sciences,
University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark
oebn@imf.au.dk

Peter Reinhard Hansen
Department of Economics, Stanford University,
Landau Economics Building, 579 Serra Mall,
Stanford, CA 94305-6072, USA
peter.hansen@stanford.edu

Asger Lunde
Department of Marketing and Statistics,
Aarhus School of Business,
Fuglesangs Alle 4, DK-8210 Aarhus V, Denmark
ALunde@asb.dk

Neil Shephard
Nuffield College, University of Oxford, Oxford OX1 1NF, UK
neil.shephard@nuffield.ox.ac.uk

Abstract

This paper shows how to use realised kernels to carry out efficient feasible inference on the ex-post variation of underlying equity prices in the presence of simple models of market frictions. The issue is subtle with only estimators which have symmetric weights delivering consistent estimators with mixed Gaussian limit theorems. The weights can be chosen to achieve the best possible rate of convergence and to have an asymptotic variance which is close to that of the maximum likelihood estimator in the parametric version of this problem. Realised kernels can also be selected to (i) be analysed using endogenously spaced data such as that in databases on transactions, (ii) allow for market frictions which are endogenous, (iii) allow for temporally dependent noise. The finite sample performance of our estimators is studied using simulation, while empirical work illustrates their use in practice.

Keywords: Bipower variation; Long run variance estimator; Market frictions; Quadratic variation; Realised variance.

*An initial version of this paper was circulated in the Autumn of 2004 under the title "Regular and modified kernel-based estimators of integrated variance: the case with independent noise." We thank Yacine Aït-Sahalia, Clive Bowsher, Xin Huang, Ilze Kalnina, Silja Kinnebrock, Jeremy Large, Nour Meddahi, Ulrich Müller, Per Mykland, Mark Podolskij, Joe Romano, George Tauchen and Lan Zhang for valuable comments. Neil Shephard's research is supported by the UK's ESRC. The Ox language of Doornik (2001) was used to perform the calculations reported here.

1 Introduction

In the last five years the harnessing of high frequency financial data has lead to substantial improvements in our understanding of financial volatility. The idea behind this is to use quadratic variation as a measure of the ex-post variation of asset prices. Estimators of increments of this quantity can allow us, for example, to improve forecasts of future volatility and estimate parametric models of time varying volatility. The most commonly used estimator of this type is the realised variance (e.g. Andersen, Bollerslev, Diebold, and Labys (2001), Meddahi (2002) and Barndorff-Nielsen and Shephard (2002)), which the recent econometric literature has shown has good properties when applied to 10 to 30 minute return data for frequently traded assets.

A weakness with realised variance is that it can be unacceptably sensitive to market frictions when applied to returns recorded over shorter time intervals such as 1 minute, or even more ambitiously, 1 second (e.g. Zhou (1996), Fang (1996) and Andersen, Bollerslev, Diebold, and Labys (2000)). In this paper we study the class of *realised kernel* estimators of quadratic variation. We show how to design these estimators to be robust to certain types of frictions and to be efficient.

The problem of estimating the quadratic variation is, in some ways, similar to the estimation of the long-run variance in stationary time series. For example, the realized variance is analogous to the sum-of-squares variance estimator. The moving average filter of Andersen, Bollerslev, Diebold, and Ebens (2001) and Hansen, Large, and Lunde (2005) and the autoregressive filter of Bollen and Inder (2002), are estimators that use pre-whitening techniques — see also Bandi and Russell (2005a). Aït-Sahalia, Mykland, and Zhang (2005) and Oomen (2005) propose parametric estimators. The two scale estimator of Zhang, Mykland, and Aït-Sahalia (2005) was the first consistent nonparametric estimator for stochastic volatility plus noise processes. It is related to the earlier work of Zhou (1996) on scaled Brownian motion plus noise. The multiscale estimator of Zhang (2006) is more efficient than the two scale estimator. An alternative is due to Large (2005), whose alternation estimator applies when prices move by a sequence of single ticks. Finally, Delattre and Jacod (1997) studied the effect of rounding on realised variances.

More formally, our interest will be in inference for the ex-post variation of log-prices over some arbitrary fixed time period, such as a day, using estimators of realised kernel type. In order to focus on the core issue we represent this period as the single interval $[0, t]$. For a continuous time log-price process X and time gap $\delta > 0$ our flat-top *realised kernels* take on the following form

$$\widetilde{K}(X_\delta) = \gamma_0(X_\delta) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\left\{\gamma_h(X_\delta) + \gamma_{-h}(X_\delta)\right\}.$$

Here the non-stochastic $k(x)$ for $x \in [0, 1]$ is a weight function and the h-th realised autocovariance is

$$\gamma_h(X_\delta) = \sum_{j=1}^{n}\left(X_{\delta j} - X_{\delta(j-1)}\right)\left(X_{\delta(j-h)} - X_{\delta(j-h-1)}\right),$$

with $h = -H, \ldots, -1, 0, 1, \ldots, H$ and $n = \lfloor t/\delta \rfloor$. We will think of δ as being small and so $X_{\delta j} - X_{\delta(j-1)}$ represents the j-th high frequency return, while $\gamma_0(X_\delta)$ is the realised variance of X. Here $\widetilde{K}(X_\delta) - \gamma_0(X_\delta)$ is the realised kernel correction to realised variance for market frictions.

We show that if $k(0) = 1$, $k(1) = 0$ and $H = cn^{2/3}$ then the resulting estimator is asymptotically mixed Gaussian, converging at rate $n^{1/6}$. Here c is a estimable constant which can be optimally chosen as a function of k, the variance of the noise and a function of the volatility path, to minimise the asymptotic variance of the estimator. The special case of a so-called flat-top Bartlett kernel, where $k(x) = 1 - x$, is particularly interesting as its asymptotic distribution is the same as that of the two scale estimator.

When we additionally require that $k'(0) = 0$ and $k'(1) = 0$ then by taking $H = cn^{1/2}$ the resulting estimator is asymptotically mixed Gaussian, converging at rate $n^{1/4}$, which we know is the fastest possible rate. When $k(x) = 1 - 3x^2 + 2x^2$ this estimator has the same asymptotic distribution as the multiscale estimator.

We use our novel realised kernel framework to make three innovations to the literature: (i) we design a kernel to have an asymptotic variance which is smaller than the multiscale estimator, (ii) we design $\widetilde{K}(X_\delta)$ for data with endogenously spaced data, such as that in databases on transactions (see Renault and Werker (2005) for the importance of this), (iii) we cover the case where the market frictions are endogenous. All of these results are new and the last two of them are essential from a practical perspective.

Clearly these realised kernels are related to so-called HAC estimators discussed by, for example, Gallant (1987), Newey and West (1987), and Andrews (1991). The flat-top of the kernel, where a unit weight is imposed on the first autocovariance, is related to the flat-top literature initiated by Politis and Romano (1995) and Politis (2005). However, the realised kernels are not scaled by the sample size, which has a great number of technical implications and makes their analysis subtle.

The econometric literature on realised kernels was started by Zhou (1996) who proposed $\widetilde{K}(X_\delta)$ with $H = 1$. This suffices for unbiasedness under a simple model for frictions where the population values of higher-order autocovariances of the market frictions are zero. However, the estimator is inconsistent. Hansen and Lunde (2006) use realised kernel type estimators, with $k(x) = 1$ for general H to characterize the second order properties of market microstructure noise. Again these are inconsistent estimators. Some analysis of the finite sample performance of a type of inconsistent realised kernel is provided by Bandi and Russell (2005b), who focus on the selection of H in the case where $k(x) = 1 - x$, the Bartlett kernel.

In Section 2 we detail our notation and assumptions about the efficient price process, market frictions and realised kernels. In Section 3 we give a central limit theory for $\gamma_h(X_\delta)$. Section 4 then looks at the corresponding properties of realised kernels. In Section 5 we study the effect irregularly spaced data has on our theory and extend the analysis of realised kernels to the case where the noise is temporally dependent and endogenous. Section 6 performs a Monte Carlo experiment to assess the accuracy of our feasible central limit theory. In Section 7 we apply the theory to some data taken from the New York stock exchange and in Section 8 we draw conclusions. A lengthy Appendix details the proofs of the results given in the paper.

2 Notation, definitions and background

2.1 Semimartingales and quadratic variation

The fundamental theory of asset prices says that the log-price at time t, Y_t, must, in a frictionless arbitrage free market, obey a *semimartingale* process (written $Y \in \mathcal{SM}$) on some filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq T^*}, P\right)$, where $T^* \leq 0$. Introductions to the economics and mathematics of semimartingales are given in Back (1991) and Protter (2004). It is unusual to start the clock of a semimartingale before time 0, but this raises no technical difficulty and eases the exposition. We think of 0 as the start of an economic day and sometimes it is useful to use data from the previous day. Alternatively we could define $\gamma_h(X_\delta)$ as using data from time 0 to t by changing the range of the summation to $j = H + 1$ and $n - H$ and then scaling the resulting estimator. All the theoretical properties we discuss in this paper would then follow in the same way as here.

Crucial to semimartingales, and to the economics of financial risk, is the *quadratic variation* (QV) process of $Y \in \mathcal{SM}$. This can be defined as

$$[Y]_t = \text{p}-\lim_{n\to\infty} \sum_{j=1}^{t_j \leq t} \left(Y_{t_j} - Y_{t_{j-1}}\right)^2, \tag{1}$$

(e.g. Protter (2004, p. 66–77) and Jacod and Shiryaev (2003, p. 51)) for any sequence of deterministic partitions $0 = t_0 < t_1 < \cdots < t_n = T$ with $\sup_j\{t_{j+1} - t_j\} \to 0$ for $n \to \infty$. Discussion of the case of stochastic spacing $\{t_j\}$ will be given in Section 5.1.

The most familiar semimartingales are of *Brownian semimartingale* type ($Y \in \mathcal{BSM}$)

$$Y_t = \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_u \mathrm{d}W_u, \tag{2}$$

where a is a predictable locally bounded drift, σ is a càdlàg volatility process and W is a Brownian motion. For reviews of the econometrics of this type of process see, for example, Ghysels, Harvey, and Renault (1996) and Shephard (2005). If $Y \in \mathcal{BSM}$ then

$$[Y]_t = \int_0^t \sigma_u^2 \mathrm{d}u.$$

In some of our asymptotic theory we also assume, for simplicity of exposition, that

$$\sigma_t = \sigma_0 + \int_0^t a_u^{\#} \mathrm{d}u + \int_0^t \sigma_u^{\#} \mathrm{d}W_u + \int_0^t v_u^{\#} \mathrm{d}V_u, \tag{3}$$

where $a^{\#}$, $\sigma^{\#}$ and $v^{\#}$ are adapted càdlàg processes, with $a^{\#}$ also being predictable and locally bounded and V is Brownian motion independent of W. Much of what we do here can be extended to allow for jumps in σ, following the details discussed in Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006), but we will not address that here.

2.2 Assumptions about noise

We write the effects of market frictions as U, so that we observe the process

$$X = Y + U, \tag{4}$$

and think of $Y \in \mathcal{BSM}$ as the efficient price. Our scientific interest will be in estimating $[Y]_t$. In the main part of our work we will assume that $Y \perp\!\!\!\perp U$ where, in general, $A \perp\!\!\!\perp B$ denotes that A and B are independent. From a market microstructure theory viewpoint this is a strong assumption as one may expect U to be correlated with increments in Y. However, the empirical work of Hansen and Lunde (2006) suggests this independence assumption is not too damaging statistically when we analyse data in thickly traded stocks recorded every minute. In Section 5.3 we will show realised kernels are consistent when this assumption is relaxed.

Furthermore we mostly work under a white noise assumption about the U process ($U \in \mathcal{WN}$) which we assume has

$$\mathrm{E}(U_t) = 0, \quad \mathrm{Var}(U_t) = \omega^2, \quad \mathrm{Var}(U_t^2) = \lambda^2\omega^4, \quad U_t \perp\!\!\!\perp U_s$$

for any $t, s, \lambda \in \mathbb{R}^+$. This white noise assumption is unsatisfactory from a number of viewpoints (e.g. Phillips and Yu (2006) and Kalnina and Linton (2006)) but is a useful starting point if we think of the market frictions as operating in tick time (e.g. Bandi and Russell (2005c), Zhang, Mykland, and Aït-Sahalia (2005) and Hansen and Lunde (2006)). A feature of $U \in \mathcal{WN}$ is that $[U]_t = \infty$. Thus $U \notin \mathcal{SM}$ and so in a frictionless market would allow arbitrage opportunities. Hence it only makes sense to add processes of this type when there are frictions to be modelled. In Section 5.2 we will show our kernel can be made to be consistent when the $U \in \mathcal{WN}$ assumption is dropped. This type of property has been achieved earlier by the two scale estimator of Aït-Sahalia, Mykland, and Zhang (2006).

2.3 Defining the realised autocovariation process

We measure returns over time spans of length δ. Define, for any processes X and Z,

$$\gamma_h(Z_\delta, X_\delta)_t = \sum_{j=1}^{n} \left(Z_{j\delta} - Z_{(j-1)\delta}\right)\left(X_{(j-h)\delta} - X_{(j-h-1)\delta}\right), \qquad h = -H, \ldots, -1, 0, 1, 2, \ldots, H.$$

We call $\gamma_h(X_\delta) = \gamma_h(X_\delta, X_\delta)$ the realised autocovariation process, while noting that

$$\gamma_h(X_\delta) = \gamma_h(Y_\delta) + \gamma_h(U_\delta) + \widetilde{\gamma}_h(Y_\delta, U_\delta), \quad \text{where} \quad \widetilde{\gamma}_h(Y_\delta, U_\delta) = \gamma_h(Y_\delta, U_\delta) + \gamma_{-h}(Y_\delta, U_\delta). \tag{5}$$

The case of realised QV has a special notation $[X_\delta] = \gamma_0(X_\delta)$. The daily increments of this process are called realised variances, their square root the realised volatility. Realised volatility has a very long history. It appears in, for example, Rosenberg (1972), Merton (1980) and French, Schwert, and Stambaugh (1987), with Merton (1980) making the implicit connection with the case where $\delta \downarrow 0$ in the pure scaled Brownian motion plus drift case. For more general processes a closer connection between realised QV and QV, and its use for econometric purposes, was made in Andersen, Bollerslev, Diebold, and Labys (2001), Comte and Renault (1998) and Barndorff-Nielsen and Shephard (2002).

2.4 Defining the realised kernel

We study the realised kernel

$$\widetilde{K}(X_\delta) = \gamma_0(X_\delta) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\widetilde{\gamma}_h(X_\delta), \quad \widetilde{\gamma}_h(X_\delta) = \gamma_h(X_\delta) + \gamma_{-h}(X_\delta), \tag{6}$$

when $k(0) = 1$ and $k(1) = 0$, noting that $\widetilde{K}(X_\delta) = \widetilde{K}(Y_\delta) + \widetilde{K}(U_\delta) + 2\widetilde{K}(Y_\delta, U_\delta)$. Throughout we will write, using $^\intercal$ to denote a transpose,

$$\begin{aligned}
\gamma(X_\delta) &= \{\gamma_0(X_\delta), 2\gamma_1(X_\delta), \ldots, 2\gamma_H(X_\delta)\}^\intercal, \\
\widetilde{\gamma}(X_\delta) &= \{\gamma_0(X_\delta), \widetilde{\gamma}_1(X_\delta), \ldots, \widetilde{\gamma}_H(X_\delta)\}^\intercal, \\
\widetilde{\gamma}(Y_\delta, U_\delta) &= (\gamma_0(Y_\delta, U_\delta), \widetilde{\gamma}_1(Y_\delta, U_\delta), \ldots, \widetilde{\gamma}_H(Y_\delta, U_\delta))^\intercal.
\end{aligned}$$

An implication of our analysis will be that the asymmetric kernel

$$K(X_\delta) = \gamma_0(X_\delta) + 2\sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\gamma_h(X_\delta) \tag{7}$$

is inconsistent and so should be avoided in high frequency financial econometrics.

2.5 Maximum likelihood estimator of QV

In order to put non-parametric results in context, it is helpful to have a parametric benchmark. In this subsection we recall the behaviour of the maximum likelihood (ML) estimator of $\sigma^2 = [Y]_1$ when $Y_t = \sigma W_t$ and where the noise is Gaussian. All the results we state here are already known.

Given $Y \perp\!\!\!\perp U$ and taking $t = 1$ it follows that

$$\begin{pmatrix} X_{1/n} - X_0 \\ X_{2/n} - X_{1/n} \\ \vdots \\ X_1 - X_{(n-1)/n} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \frac{\sigma^2}{n}I + \begin{pmatrix} 2\omega^2 & \bullet & \bullet & \bullet \\ -\omega^2 & 2\omega^2 & \bullet & \bullet \\ 0 & -\omega^2 & 2\omega^2 & \bullet \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}\right).$$

Let $\hat{\sigma}^2_{\mathrm{ML}}$ and $\hat{\omega}^2_{\mathrm{ML}}$ denote the ML estimators. Their asymptotic properties are given from classical results about the MA(1) process. By adopting the expression given in Aït-Sahalia, Mykland, and Zhang (2005, Proposition 1) to our notation, we have that for $\omega^2 > 0$

$$\left\{ \begin{array}{l} n^{1/4}\left(\hat{\sigma}^2_{\mathrm{ML}} - \sigma^2\right) \\ n^{1/2}\left(\hat{\omega}^2_{\mathrm{ML}} - \omega^2\right) \end{array} \right\} \xrightarrow{L} N\left(0, \begin{pmatrix} 8\omega\sigma^3 & 0 \\ 0 & 2\omega^4 \end{pmatrix}\right). \tag{8}$$

This shows that $\hat{\sigma}^2_{\mathrm{ML}}$ converges at quite a slow rate. This is a familiar result from the work of, for example, Stein (1987) and Gloter and Jacod (2001a, 2001b).

The special case where there is no market microstructure noise, (i.e. the true value of $\omega^2 = 0$) results in faster rates of convergence for $\hat{\sigma}^2_{\mathrm{ML}}$, since

$$n^{1/2}\left(\hat{\sigma}^2_{\mathrm{ML}} - \sigma^2\right) \xrightarrow{L} N\left(0, 6\sigma^4\right). \tag{9}$$

When ω^2 is also known a priori to be zero, and so is not estimated, then

$$n^{1/2}\left(\hat{\sigma}^2_{\mathrm{ML}} - \sigma^2\right) \xrightarrow{L} N\left(0, 2\sigma^4\right). \tag{10}$$

3 Central limit theory for $\gamma(X_\delta)$ and $\widetilde{\gamma}(X_\delta)$

3.1 Core result

Here we will study the large sample behaviour of the contributions to $\gamma(X_\delta)$. These results will be used in the next Section to derive the properties of $\widetilde{K}_w(X_\delta)$ and so to select k to produce good estimators of $[Y]$. Throughout this paper $\stackrel{Ls}{\to}$ will denote convergence in law stably, which will be discussed in some detail in a moment.

Theorem 1 *Suppose that $Y \in \mathcal{BSM}$ and (3) holds, then as $\delta \downarrow 0$ for the Y component alone*

$$\delta^{-1/2}\begin{pmatrix} [Y_\delta]_t - \int_0^t \sigma_u^2 \mathrm{d}u \\ \gamma_1(Y_\delta) \\ \vdots \\ \gamma_H(Y_\delta) \end{pmatrix} \stackrel{Ls}{\to} MN\left(0, A \times \int_0^t \sigma_u^4 \mathrm{d}u\right), \quad A = \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Here MN denotes a mixed normal distribution and $\gamma_h(Y_\delta) - \gamma_{-h}(Y_\delta) = O_p(\delta)$.

If $U \in \mathcal{WN}$ and $Y \perp\!\!\!\perp U$ then $\widetilde{\gamma}(Y_\delta, U_\delta) \stackrel{Ls}{\to} MN\left(0, 2\omega^2[Y]B\right)$, where B is a $(H+1) \times (H+1)$ symmetric matrix with block structure

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \quad B_{22} = \begin{pmatrix} 2 & \bullet & \bullet & \bullet \\ -1 & 2 & \bullet & \bullet \\ \ddots & \ddots & \ddots & \bullet \\ \cdots & 0 & -1 & 2 \end{pmatrix},$$

$$B_{11} = \begin{pmatrix} 1 & \bullet \\ -1 & 2 \end{pmatrix}, \qquad B_{21} = \begin{pmatrix} 0 & -1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix},$$

$B_{12} = B_{21}^\intercal$. Here B_{22} is a $(H-1) \times (H-1)$ symmetric matrix.

Finally, when $U \in \mathcal{WN}$ and writing $n = \lfloor t/\delta \rfloor$, for $n \geq H$

$$\mathrm{E}\left\{\gamma(U_\delta)\right\} = \mathrm{E}\left\{\widetilde{\gamma}(U_\delta)\right\} = 2\omega^2 n\left(1, -1, 0, 0, \ldots, 0\right)^\intercal,$$

$$\mathrm{Cov}\left\{\gamma(U_\delta)\right\} = 4\omega^4\left(nC + D\right), \quad \mathrm{Cov}\left\{\widetilde{\gamma}(U_\delta)\right\} = 4\omega^4\left(nC + \widetilde{D}\right). \tag{11}$$

Here the $(H+1) \times (H+1)$ symmetric matrices C, D and $\widetilde{D}$ have block structure

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}, \quad D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}, \quad \widetilde{D} = \begin{pmatrix} \widetilde{D}_{11} & \widetilde{D}_{12} \\ \widetilde{D}_{21} & \widetilde{D}_{22} \end{pmatrix},$$

where the $(H-1) \times (H-1)$ and $(H-1) \times 2$ dimensional matrices are

$$C_{22} = \begin{pmatrix} 6 & \bullet & \bullet & \bullet & \bullet \\ -4 & 6 & \bullet & \bullet & \bullet \\ 1 & -4 & 6 & \bullet & \bullet \\ 0 & 1 & -4 & 6 & \bullet \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}, \qquad C_{21} = \begin{pmatrix} 1 & -4 \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix},$$

$$D_{22} = \begin{pmatrix} -2 & \bullet & \bullet & \bullet & \bullet \\ 2 & -2 & \bullet & \bullet & \bullet \\ -1 & 2 & -2 & \bullet & \bullet \\ 0 & -1 & 2 & -2 & \bullet \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}, \qquad D_{21} = \begin{pmatrix} -1 & 2 \\ 0 & -1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix},$$

$$\widetilde{D}_{22} = \begin{pmatrix} -7 & \bullet & \bullet & \bullet & \bullet & \bullet \\ 6 & -10 & \bullet & \bullet & \bullet & \bullet \\ -2 & 8 & -13 & \bullet & \bullet & \bullet \\ 0 & -2.5 & 10 & -16 & \bullet & \bullet \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{H}{2} & 2H & -3H-1 \end{pmatrix}, \quad \widetilde{D}_{21} = \begin{pmatrix} -1 & 4 \\ 0 & -\frac{3}{2} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix},$$

where $C_{12} = C_{21}^{\intercal}$, $D_{12} = D_{21}^{\intercal}$ *and* $\widetilde{D}_{12} = \widetilde{D}_{21}^{\intercal}$. *The* 2×2 *matrices* C_{11}, D_{11} *and* $\widetilde{D}_{11}$ *are*

$$C_{11} = \begin{pmatrix} 1+\lambda^2 & -2-\lambda^2 \\ -2-\lambda^2 & 5+\lambda^2 \end{pmatrix}, \qquad D_{11} = \begin{pmatrix} -\lambda^2/2 & 1+\lambda^2/2 \\ 1+\lambda^2/2 & -2 \end{pmatrix},$$
$$\widetilde{D}_{11} = \begin{pmatrix} -\lambda^2/2 & \lambda^2/2+1 \\ \lambda^2/2+1 & -\lambda^2/2-7/2 \end{pmatrix}.$$

Lastly $\gamma_h(U_\delta)_t - \gamma_{-h}(U_\delta)_t = O_p(1)$.

3.2 Comments

3.2.1 Stable convergence

The concept and role of stable convergence may be unfamiliar to some readers and we therefore add some words of explanation. The concise mathematical definition is as follows. Let $X_n^{\#}$ denote a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. Then we say that $X_n^{\#}$ *converges stably in law* if there exists a probability measure μ on $(\Omega \times \mathbb{R}, \mathcal{F} \times \mathcal{B})$ (where $\mathcal{B}$ denotes the Borel σ-algebra on $\mathbb{R}$) such that for every bounded random variable V on $(\Omega, \mathcal{F}, P)$ and every bounded and continuous function f on $\mathbb{R}$ we have that, for $n \to \infty$,

$$\mathrm{E}\left(Vf\left(X_n^{\#}\right)\right) \to \int V(\omega) f(x) \mu(\mathrm{d}\omega, \mathrm{d}x).$$

If $X_n^{\#}$ converges stably in law then, in particular, it converges in distribution (or in law or weak convergence), the limiting law being $\mu(\Omega, \cdot)$. Accordingly, one says that $X_n^{\#}$ converges stably to some random variable $X^{\#}$ if there exists a probability measure μ, as above, such that $X^{\#}$ has law $\mu(\Omega, \cdot)$. This concept and its extension to stable convergence of processes is discussed in Jacod and Shiryaev (2003, pp. 512–518). For earlier discussions see, for example, Rényi (1963), Aldous and Eagleson (1978), Hall and Heyde (1980, pp. 56–58) and Jacod (1997). An early use of this concept in econometrics was Phillips and Ouliaris (1990). It is used extensively in, for example, Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006).

However, this formalisation does not reveal the key property of stable convergence which is that $X_n^{\#} \xrightarrow{Ls} X^{\#}$ stably implies that for any random variable Z, the pair $(Z, X_n^{\#})$

converges in law to $\left(Z, X^{\#}\right)$. Consider the following simple example of the above result. Let

$$X_n^{\#} = \delta^{-1/2}\left([Y_\delta]_t - \int_0^t \sigma_u^2 \mathrm{d}u\right)$$

and $Z = \sqrt{\int_0^t \sigma_u^4 \mathrm{d}u}$. Our focus is on $X_n^{\#}/\sqrt{Z}$ and our convergence in law stably implies that

$$\delta^{-1/2}\left([Y_\delta]_t - \int_0^t \sigma_u^2 \mathrm{d}u\right) / \sqrt{\int_0^t \sigma_u^4 \mathrm{d}u} \xrightarrow{L} N(0,2). \tag{12}$$

Without the convergence in law stably, (12) could not be deduced. The following Lemma is helpful in using this concept.

Lemma 1 *Let $Y_n^{\#}$ and $Z_n^{\#}$ $(n = 1, 2, \dots)$ be sequences of random variables, defined on some probability space $(\Omega, \mathcal{F}, P)$, and suppose that $Y_n^{\#}$ converges stably to a random variable $Y^{\#}$ and that $Z_n^{\#}$ converges in probability to 0, i.e. we have $Y_n^{\#} \xrightarrow{Ls} Y^{\#}$ and $Z_n^{\#} \xrightarrow{p} 0$ for $n \to \infty$. (Here $Y^{\#}$ may be defined on an extension of $(\Omega, \mathcal{F}, P)$.) Then $X_n^{\#} = Y_n^{\#} + Z_n^{\#} \xrightarrow{Ls} Y^{\#}$.*

3.2.2 Related results

The asymptotic distribution

$$\delta^{-1/2}\left([Y_\delta]_t - [Y]_t\right) \xrightarrow{Ls} MN\left(0, 2\int_0^t \sigma_u^4 \mathrm{d}u\right) \tag{13}$$

appears in Jacod (1994), Jacod and Protter (1998) and Barndorff-Nielsen and Shephard (2002). This estimator has the efficiency of the ML estimator (10) in the pure Brownian motion case.

The extension of the limiting results to deal with more general realised autocovariances is new. We do this here in terms of the $\widetilde{\gamma}_h = \gamma_h + \gamma_{-h}$ as they will be key later

$$\delta^{-1/2}\begin{pmatrix} [Y_\delta]_t - \int_0^t \sigma_u^2 \mathrm{d}u \\ \widetilde{\gamma}_1(Y_\delta) \\ \vdots \\ \widetilde{\gamma}_H(Y_\delta) \end{pmatrix} \xrightarrow{Ls} MN\left(0, \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 4 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 4 \end{pmatrix} \times \int_0^t \sigma_u^4 \mathrm{d}u\right).$$

We now turn our attention to some simple kernels. The most interesting special case is

$$\delta^{-1/2}\left([Y_\delta]_t + \widetilde{\gamma}_1(Y_\delta) - [Y]_t\right) \xrightarrow{Ls} MN\left(0, 6\int_0^t \sigma_u^4 \mathrm{d}u\right), \tag{14}$$

which achieves the bound (9).

The main impact of the noise is through the $\widetilde{\gamma}(U_\delta)$ term. The mean and variance of $[U_\delta] = \gamma_0(U_\delta)$ was studied by, for example, Fang (1996), Bandi and Russell (2005c) and Zhang, Mykland, and Aït-Sahalia (2005). Note that both the mean and variance of $[U_\delta]$ explode as $n \to \infty$. Of course these features are passed onto $[X_\delta]$ making it inconsistent, thus motivating this literature. The bias of $[U_\delta]$ is exactly balanced by that of $\widetilde{\gamma}_1(U_\delta)$, so producing the unbiased but inconsistent estimator $[X_\delta] + \widetilde{\gamma}_1(X_\delta)$ with (e.g. Zhou (1996)) $\mathrm{E}([U_\delta] + \widetilde{\gamma}_1(U_\delta)) = 0$ and $\mathrm{Var}([U_\delta] + \widetilde{\gamma}_1(U_\delta)) = 4\omega^4 (2n - 1.5)$.

4 Behaviour of kernels

4.1 Core result

In this Section we derive the asymptotic behaviour of arbitrary realised kernels. In Section 4.3 we derive a way of choosing the number of terms to use in the kernel, which is indexed by ω^2 and $\int_0^t \sigma_u^4 \mathrm{d}u$. Subsequently we provide estimators of these quantities, implying the feasible asymptotic distribution of the realised kernel can be applied in practice to form confidence intervals for $[Y]$.

The asymptotic behaviour of the realised kernel is determined by the asymptotic behaviour of quadratic forms in the A, B, C and D or $\widetilde{D}$ matrices.

Theorem 2 *Write*

$$w = \left(1, 1, k\left(\frac{1}{H}\right), \ldots, k\left(\frac{H-1}{H}\right)\right)^{\intercal}.$$

Assume that the kernel weight function $k(x)$ is four times continuously differentiable and write, as usual, derivatives using primes. As H increases, so the flat-top kernels have

$$\begin{aligned}
w^{\intercal}Aw &= Hk_{\bullet}^{0,0} + O(1), \\
w^{\intercal}Bw &= -H^{-1}\left\{k'(0) + k_{\bullet}^{0,2}\right\} + O(H^{-2}), \\
w^{\intercal}Cw &= H^{-2}\left\{k'(0)^2 + k'(1)^2\right\} + H^{-3}\left\{k'''(0) + k_{\bullet}^{0,4}\right\} + O(H^{-4}), \\
w^{\intercal}Dw &= \lambda^2/2 - 2H^{-1}\left\{k'(0) + k_{\bullet}^{0,2}\right\} + O(H^{-2}), \\
w^{\intercal}\widetilde{D}w &= -H^{-1}\left\{k'(0) + \frac{1}{2}k'(0)^2 + k_{\bullet}^{0,2}\right\} + O(H^{-2}),
\end{aligned}$$

where

$$k_{\bullet}^{0,0} = \int_0^1 k(x)^2 \mathrm{d}x, \quad k_{\bullet}^{0,2} = \int_0^1 k(x)k''(x)\mathrm{d}x, \quad k_{\bullet}^{0,4} = \int_0^1 k(x)k''''(x)\mathrm{d}x.$$

The λ^2 term in $w^{\intercal}Dw$ cannot be forced to zero either as $H \to \infty$ or $n \to \infty$. This means that the asymmetric realised kernels (7) are always inconsistent unless $\lambda = 0$. From now on we entirely focus on $\widetilde{K}(X_\delta)$.

The large n and large H asymptotic distribution of

$$\widetilde{K}(X_\delta) - \int_0^t \sigma_u^2 \mathrm{d}u$$

is mixed normal with a zero mean and variance of

$$\begin{aligned}
4Hn^{-1}k_{\bullet}^{0,0}t\int_0^t \sigma_u^4 \mathrm{d}u &- 4H^{-1}\left\{k'(0) + k_{\bullet}^{0,2}\right\}\left\{2\omega^2\int_0^t \sigma_u^2 \mathrm{d}u + \omega^4\right\} \\
&+ 4\omega^4 n\left[H^{-2}\left\{k'(0)^2 + k'(1)^2\right\} + H^{-3}\left\{k'''(0) + k_{\bullet}^{0,4}\right\}\right] - 4\omega^4 H^{-1}\frac{1}{2}k'(0)^2.
\end{aligned} \tag{15}$$

If we now relate H to n there is an important special case. When

$$k'(0) = 0 \quad \text{and} \quad k'(1) = 0, \tag{16}$$

then setting $H = cn^{1/2}$ we produce the result

$$\begin{aligned} &n^{1/4}\left\{\widetilde{K}(X_\delta) - \int_0^t \sigma_u^2 \mathrm{d}u\right\} \\ &\quad \overset{Ls}{\to} MN\Bigg[0, 4ck_\bullet^{0,0}t\int_0^t \sigma_u^4\mathrm{d}u \\ &\qquad - 8c^{-1}k_\bullet^{0,2}\omega^2\left(\int_0^t \sigma_u^2\mathrm{d}u + \frac{\omega^2}{2}\right) + 4\omega^4 c^{-3}\left\{k'''(0) + k_\bullet^{0,4}\right\}\Bigg]. \end{aligned} \tag{17}$$

We saw in (8) that this is the best rate of convergence that can be achieved for this problem.

Whether or not (16) holds, when we set $H = cn^{2/3}$ we have

$$n^{1/6}\left\{\widetilde{K}(X_\delta) - \int_0^t \sigma_u^2\mathrm{d}u\right\} \overset{Ls}{\to} MN\left[0, 4ck_\bullet^{0,0}t\int_0^t \sigma_u^4\mathrm{d}u + 4\omega^4c^{-2}\left\{k'(0)^2 + k'(1)^2\right\}\right]. \tag{18}$$

If (16) does hold then we get the very simple result that

$$n^{1/6}\left\{\widetilde{K}(X_\delta) - \int_0^t \sigma_u^2\mathrm{d}u\right\} \overset{Ls}{\to} MN\left(0, 4ck_\bullet^{0,0}t\int_0^t \sigma_u^4\mathrm{d}u\right). \tag{19}$$

4.2 Special cases with $n^{1/6}$

When $H = cn^{2/3}$ we have the asymptotic distribution given in (18). For this class of kernels the value of c which minimises the asymptotic variance in (18) is

$$c = d\frac{\omega^{4/3}}{\left(t\int_0^t \sigma_u^4\mathrm{d}u\right)^{1/3}}, \quad \text{where} \quad d = \left[\frac{2\left\{k'(0)^2 + k'(1)^2\right\}}{k_\bullet^{0,0}}\right]^{1/3}.$$

Then the lower bound for the asymptotic variance is

$$4d\omega^{4/3}\left(t\int_0^t \sigma_u^4\mathrm{d}u\right)^{2/3}\left[k_\bullet^{0,0} + d^{-3}\left\{k'(0)^2 + k'(1)^2\right\}\right] = 6dk_\bullet^{0,0}\omega^{4/3}\left(t\int_0^t \sigma_u^4\mathrm{d}u\right)^{2/3}. \tag{20}$$

Hence $dk_\bullet^{0,0}$ controls the asymptotic efficiency of estimators in this class.

Three flat-top cases of this setup are analysed in Table 1[1]. The flat-top Bartlett kernel puts $k(x) = 1 - x$, Epanechnikov kernel puts $k(x) = 1 - x^2$, while the second order kernel has $k(x) = 1 - 2x + x^2$. The Bartlett kernel has the same asymptotic distribution as the two scale estimator. It is more efficient than the Epanechnikov kernel but less good than the second order kernel.

4.3 Special cases with $n^{1/4}$

When $H = cn^{1/2}$ the asymptotic distribution is given in (17). Given preliminary estimates of ω^2, $\int_0^t \sigma_u^2\mathrm{d}u$ and $\int_0^t \sigma_u^4\mathrm{d}u$ it is a simple matter to numerically find a value of c which minimises the asymptotic variance. To gain some understanding think of $Y = \sigma W$

[1] The results on specific k functions in this Section were calculated using the computer algebra package Maple.

	$k(x)$	$k(0)$	$k(1)$	$k'(0)$	$k'(1)$	$k_\bullet^{0,0}$	d	$dk_\bullet^{0,0}$
Bartlett	$1-x$	1	0	-1	-1	$\frac{1}{3}$	2.28	0.763
2nd order	$1-2x+x^2$	1	0	-2	0	$\frac{1}{5}$	3.42	0.683
Epanechnikov	$1-x^2$	1	0	0	-2	$\frac{8}{15}$	2.46	1.31

Table 1: *Properties of some $n^{1/6}$ flat-top realised kernels. Bartlett kernel has the same asymptotic distribution as the two scale estimator. $dk_\bullet^2$ measures the relative asymptotic efficiency of the realised kernels in this class.*

and $t=1$, while we will ignore the $\omega^2/2$ as it will have a small impact. Then the task simplifies to minimising

$$4\sigma^4\left(ck_\bullet^{0,0}-2c^{-1}k_\bullet^{0,2}\xi+\xi^2c^{-3}f\right),$$

where $\xi=\omega^2/\sigma^2$ and $f=k'''(0)+k_\bullet^{0,4}$. Writing $x=c^2$ the first order condition is $k_\bullet^{0,0}x^2+2k_\bullet^{0,2}\xi x-3\xi^2f=0$. Taking the square root of the positive root yields

$$\widehat{c}=\frac{\omega}{\sigma}d,\quad d=\sqrt{\frac{1}{k_\bullet^{0,0}}\left\{-k_\bullet^{0,2}+\sqrt{\left(k_\bullet^{0,2}\right)^2+3k_\bullet^{0,0}f}\right\}}.$$

At this optimal point the asymptotic variance is

$$4\left(dk_\bullet^{0,0}-2d^{-1}k_\bullet^{0,2}+d^{-3}f\right)\sigma^3\omega=g\sigma^3\omega.$$

From (8) we should expect that $g\geq 8$.

	$k(x)$	$k_\bullet^{0,0}$	$k_\bullet^{0,2}$	f	d	g
Cubic kernel	$1-3x^2+2x^3$	0.371	-1.2	12	3.68	9.04
5-th order kernel	$1-10x^3+15x^4-6x^5$	0.391	-1.42	17.1	3.70	10.2
6-th order kernel	$1-15x^4+24x^5-10x^6$	0.471	-1.55	22.8	3.97	12.1
7-th order kernel	$1-21x^5+35x^6-15x^7$	0.533	-1.71	31.8	4.11	13.9
8-th order kernel	$1-28x^6+48x^7-21x^8$	0.582	-1.87	43.8	4.31	15.7
Parzen	$\begin{cases}1-6x^2+6x^3 & 0\leq x\leq 1/2\\ 2(1-x)^3 & 1/2\leq x\leq 1\end{cases}$	0.269	-1.5	24	4.77	8.54
Tukey-Hanning	$\{1+\cos(\pi x)\}/2$	0.375	-1.23	12.1	3.70	9.18
Mod. Tukey-Hanning	$\{1-\cos\pi(1-x)^2\}/2$	0.218	-1.71	41.7	5.74	8.29

Table 2: *Properties of some $n^{1/4}$ flat-top realised kernels, where $f=k'''(0)+k_\bullet^{0,4}$. The cubic kernel has the same asymptotic distribution as the multiscale estimator. g measures the relative asymptotic efficiency of the realised kernels in this class — 8 being the parametric efficiency bound.*

Eight flat-top cases of this setup are analysed in Table 2. The first is derived by thinking of a cubic kernel $k(x)=1+ax+bx^2+dx^3$, where a,b,d are constants. We can choose a, b, d by imposing the conditions (16) and that $k(0) = 1$ and $k(1) = 0$. The resulting cubic kernel has $k(x) = 1 - 3x^2 + 2x^3$, which has some of the features of cardinal cubic splines (e.g. Park and Schowengerdt (1983)) and quadratic mother kernels (e.g. Phillips, Sun, and Jin (2003)). It is also noteworthy as it has the same asymptotic distribution as the multiscale estimator. The flat-top Tukey-Hanning kernel puts $k(x) = \{1 + \cos(\pi x)\}/2$. Another interesting estimator is the flat-top Parzen kernel[2], which places

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \leq x \leq 1/2 \\ 2(1-x)^3 & 1/2 \leq x \leq 1. \end{cases}$$

We call our final k the modified Tukey-Hanning kernel

$$k(x) = \left\{1 - \cos \pi (1-x)^2\right\}/2. \tag{21}$$

Table 2 shows that the performance of the Tukey-Hanning kernel is almost identical to that of the cubic kernel. The Parzen kernel outperforms the cubic kernel, but is not as good as (21). Both kernels fail to reach the parametric efficiency bound, but are very close and tend to select more lags than the cubic kernel.

It is important to ask whether the approximation suggested by Theorem 2 and our special cases thereof provides a useful guide to finite sample behaviour? Table 3 gives $\mathrm{Var}\big\{n^{1/4}\widetilde{K}(X_\delta)\big\}/\omega$ listed against n in the Brownian motion plus noise case for a variety of values of ω^2 when $\sigma = 1$. The most empirically realistic value for ω^2 is around 0.001 for the types of data we study later in this paper. The Table also includes results for an optimal selection of k, computed numerically. This indicates that there does exist a realised kernel which can achieve the ML efficiency bound of 8 in this case. More generally the Table shows that the asymptotics provides a good approximation to the finite sample case, especially when n is over $1,000$ and when ω^2 is moderate to large. The Table also shows that even though the Bartlett kernel converges at the slow $n^{1/6}$ rate, it is only mildly inefficient even when n is $4,000$. When ω^2 is small the asymptotic variance provides a poor approximation in all cases unless n is $4,000$ or so. Of course, in that case the realised kernels are quite precise as the asymptotic variance is proportional to $\omega\sigma^3$.

4.4 Non-flat-top kernels

The flat-top constraint is imposed on these kernels to make them unbiased. If we remove the flat-top constraint then the realised kernel becomes

$$\overline{K}(X_\delta) = \gamma_0(X_\delta) + \sum_{h=1}^{H} k\left(\frac{h}{H}\right)\left\{\gamma_h(X_\delta) + \gamma_{-h}(X_\delta)\right\},$$

where we assume $k(0) = 1$ and $k(1) = 0$. Now the bias in the Bartlett case $k(x) = 1 - x$ is $O(n/H) = O(n^{1/3})$. In the cubic case it is $O(n/H^2) = O(1)$, which is better but not satisfactory. To remove the flat-top condition we need a kernel which is a higher

[2]The Parzen kernel is not everywhere differentiable and so the above formulas do not immediately apply. However, we can simply split the integrals into $(0, 1/2)$ and $(1/2, 1)$ and carry out the integrals over those regions. Using this argument delivers the results we give here.

	$\omega^2 = 0.1$					$\omega^2 = 0.01$				
n	Opt	Mod	Par	3-rd	Bart	Opt	Mod	Par	3-rd	Bart
256	8.52	9.11	9.39	9.60	10.7	9.63	10.6	10.8	10.7	10.6
$1,024$	8.30	8.76	9.03	9.37	11.9	8.73	9.43	9.73	9.81	10.3
$4,096$	8.19	8.58	8.85	9.26	13.9	8.34	8.86	9.13	9.40	10.9
$16,384$	8.14	8.49	8.76	9.21	16.8	8.17	8.58	8.84	9.22	12.5
$65,536$	8.12	8.45	8.71	9.19	20.6	8.08	8.44	8.70	9.13	14.8
$1,048,576$	8.10	8.41	8.68	9.17	31.9	8.02	8.33	8.59	9.07	22.2
∞		8.29	8.54	9.04	∞		8.29	8.54	9.04	∞
	$\omega^2 = 0.001$					$\omega^2 = 0.0001$				
n	Opt	Mod	Par	3-rd	Bart	Opt	Mod	Par	3-rd	Bart
256	15.1	15.4	16.2	16.1	16.9	38.7	38.8	38.8	38.8	38.8
$1,024$	10.8	11.8	12.1	12.1	11.7	21.0	21.1	21.2	23.2	21.5
$4,096$	9.22	10.0	10.3	10.4	10.5	13.2	14.0	15.0	14.9	14.0
$16,384$	8.55	9.19	9.47	9.61	10.4	10.1	11.1	11.6	11.3	11.0
$65,536$	8.26	8.73	9.00	9.31	11.3	8.93	9.69	10.0	10.0	10.2
$1,048,576$	8.06	8.40	8.66	9.10	15.8	8.20	8.64	8.90	9.25	11.9
∞		8.29	8.54	9.04	∞		8.29	8.54	9.04	∞

Table 3: *Flat-top realised kernels.* $Var\left\{n^{1/4}\widetilde{K}(X_\delta)\right\}/\omega$ *listed against* n*. Asymptotic lower bound is 8. 'Opt' refers to* k *selected numerically to minimise the finite sample variance of an unbiased realised kernel. '3-rd' refers to* $1 - 3x^2 + 2x^3$*. 'Mod' denotes modified Tukey-Hanning (21).*

n	$\omega^2=0.01$ Var	Bias^2	$\omega^2=0.001$ Var	Bias^2	$\omega^2=0.0001$ Var	Bias^2	$\omega^2=0.01$ Var	Bias^2	$\omega^2=0.001$ Var	Bias^2	$\omega^2=0.0001$ Var	Bias^2
	5-th order kernel						6-th order kernel					
256	9.97	5.28	8.34	33.1	13.8	4.19	11.9	0.10	13.1	1.33	13.8	4.19
$1,024$	10.1	3.47	9.74	45.4	10.7	33.5	12.0	0.02	12.3	1.22	13.1	15.8
$4,096$	10.2	1.97	10.0	34.9	9.90	189	12.0	0.00	12.0	0.48	11.5	43.0
$16,384$	10.2	1.05	10.1	31.0	9.88	461	12.1	0.00	12.0	0.11	12.1	10.4
$65,536$	10.2	0.57	10.2	17.2	10.1	322	12.1	0.00	12.0	0.02	12.0	3.41
$262,144$	10.2	0.29	10.2	9.07	10.2	254	12.1	0.00	12.0	0.00	12.0	0.71
$1,048,576$	10.2	0.15	10.2	4.65	10.2	138	12.1	0.00	12.0	0.00	12.0	0.11
∞	10.2	0.00	10.2	0.00	10.2	0.00	12.1	0.00	12.1	0.00	12.1	0.00
	7-th order kernel						8-th order kernel					
256	13.6	0.00	14.7	0.27	13.8	4.19	15.0	0.00	15.9	0.05	13.8	4.19
$1,024$	13.8	0.00	13.8	0.09	15.5	6.88	15.5	0.00	15.1	0.00	17.4	2.80
$4,096$	13.9	0.00	13.7	0.01	12.7	8.80	15.6	0.00	15.3	0.00	13.8	1.66
$16,384$	13.9	0.00	13.8	0.00	13.7	0.55	15.7	0.00	15.6	0.00	15.1	0.02
$65,536$	13.9	0.00	13.9	0.00	13.7	0.05	15.7	0.00	15.6	0.00	15.6	0.00
$262,144$	13.9	0.00	13.9	0.00	13.9	0.00	15.7	0.00	15.7	0.00	15.6	0.00
$1,048,576$	13.9	0.00	13.9	0.00	13.9	0.00	15.7	0.00	15.7	0.00	15.7	0.00
∞	13.9	0.00	13.9	0.00	13.9	0.00	15.7	0.00	15.7	0.00	15.7	0.00

Table 4: *Finite sample value of $Var\{n^{1/4}\widetilde{K}(X_\delta)\}/\omega$ listed against n and scaled squared bias for various order cases. In the n=256 case, when ω^2 is very small H is selected to be zero and so the realised kernel becomes the RV.*

polynomial near zero and is symmetric, so the bias becomes negligible. For this we add the additional constraint that $k''(0) = k''(1) = 0$. Simple polynomials of this type

$$k(x) = 1 + ax^j + bx^{j+1} + cx^{j+2}, \quad j = 3, 4, \ldots$$

yield $c = -\left(j + j^2\right)/2$, $b = 2j + j^2$, $a = -1 - 3j/2 - j^2/2$. Examples of this include

$$k(x) = \begin{cases} 1 - 10x^3 + 15x^4 - 6x^5, & j = 3 \\ 1 - 15x^4 + 24x^5 - 10x^6, & j = 4 \\ 1 - 21x^5 + 35x^6 - 15x^7, & j = 5 \\ 1 - 28x^6 + 48x^7 - 21x^8, & j = 6. \end{cases} \tag{22}$$

The bias of these estimators is $O(n/H^j) = O(n^{-(j-2)/2})$ which has no impact on its asymptotic distribution when $j \geq 3$ and should become more robust in finite samples as j increases. We call the j-th case the $j+2$-th order kernel. Table 2 shows that these estimators are less efficient than (21) realised kernel. Table 4 shows the corresponding finite sample behaviour for this realised kernel. In addition to the scaled variance, we also report the scaled squared bias

$$\left\{n^{1/4}\left(\mathrm{E}\widetilde{K}(X_\delta) - 1\right)\right\}^2 \Big/ \omega = 4n^{5/2}\omega^3\left[2\left\{1 - k\left(\frac{1}{cn^{1/2}}\right)\right\}\right]^2.$$

The Table shows the bias is small when ω^2 is large and so does not create a distortion for the inference procedure for this realised kernel. However, for small ω^2 the bias dramatically swamps the variance and so inference would be significantly affected.

4.5 Estimation of $\int_0^t \sigma_u^4 \mathrm{d}u$ and ω^2

Some of our limit theories depend upon integrated quarticity $\int_0^t \sigma_u^4 \mathrm{d}u$ and the noise's variance ω^2. We now discuss estimators of these quantities.

To estimate ω^2 Oomen (2005) suggested using the unbiased $\widetilde{\omega}^2 = -\widetilde{\gamma}_1(X_\delta)/2n$, while, for example, Bandi and Russell (2005a) suggest $\widehat{\omega}^2 = [X_\delta]/2n$ which has a bias of $\int_0^t \sigma_u^2 \mathrm{d}u/2n$. Using Theorem 1 we have that

$$\operatorname{Var}\left\{n^{1/2}\left(\widetilde{\omega}^2 - \omega^2\right)\right\} = \omega^4\left(5 + \lambda^2\right), \quad \operatorname{Var}\left\{n^{1/2}\left(\widehat{\omega}^2 - \omega^2\right)\right\} = \omega^4\left(1 + \lambda^2\right).$$

In the Gaussian case $\lambda^2 = 2$, and so $\widetilde{\omega}^2$ and $\widehat{\omega}^2$ have variances which are around 3.5 and 1.5 times that of the ML estimator in the parametric case given in (8). Although it is possible to derive a kernel style estimator to estimate ω^2 efficiently, we resist the temptation to do so here as the statistical gains are minor.

Estimating integrated quarticity reasonably efficiently is a tougher problem than estimating QV. We do not know of any existing research which has solved this problem. Define the subsampled squared returns, for some $\delta^{\#} > 0$,

$$x_{j,\cdot}^2 = \frac{1}{S}\sum_{s=0}^{S-1}\left(X_{\delta^{\#}\left(j+\frac{s}{S}\right)} - X_{\delta^{\#}\left(j-1+\frac{s}{S}\right)}\right)^2,$$

$j = 1, 2, \ldots, n$. This allows us to define a bipower variation estimator of integrated quarticity

$$\left\{X_{\delta^{\#}}, \omega^2; S\right\}^{[2,2]} = \left(\delta^{\#}\right)^{-1}\sum_{j=1}^{\lfloor t/\delta^{\#}\rfloor}\left(x_{j,\cdot}^2 - 2\omega^2\right)\left(x_{j-2,\cdot}^2 - 2\omega^2\right), \quad n = \lfloor t/\delta^{\#}\rfloor.$$

The no noise case of this statistic was introduced by Barndorff-Nielsen and Shephard (2004) and Barndorff-Nielsen and Shephard (2006) and studied in depth by Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006). See also Mykland (2006).

Detailed calculations show that when $\delta^{\#}$ is small and S is large then the conditional variance of $\left\{X_{\delta^{\#}}, \omega^2; S\right\}^{[2,2]}$ is approximately $72\omega^8 n^3/S^2$, which needs $n^{3/2}/S \to 0$ for consistency[3]. An interesting research problem is how to make this type of estimator more efficient by using kernel type estimators. For now we use moderate values of n and high values of S in our Monte Carlos and empirical work.

The finite sample performance of our estimator can be improved by using the inequality $\int_0^t \sigma_u^4 \mathrm{d}u \geq \frac{1}{t}\left(\int_0^t \sigma_u^2 \mathrm{d}u\right)^2$. This is useful as we have a very efficient estimator of

[3]Let

$$\varepsilon_j = \frac{1}{S}\sum_{s=0}^{S-1}\left[\left(U_{\delta^{\#}\left(j+\frac{s}{S}\right)} - U_{\delta^{\#}\left(j-1+\frac{s}{S}\right)}\right)^2 - 2\omega^2 \right.$$
$$\left. + 2\left(U_{\delta^{\#}\left(j+\frac{s}{S}\right)} - U_{\delta^{\#}\left(j-1+\frac{s}{S}\right)}\right)\left(Y_{\delta^{\#}\left(j+\frac{s}{S}\right)} - Y_{\delta^{\#}\left(j-1+\frac{s}{S}\right)}\right)\right],$$

$\int_0^t \sigma_u^2 du$. Thus our preferred way of estimating integrated quarticity is

$$\widehat{IQ}_{\delta,S} = \max\left[\frac{1}{t}\left(\widetilde{K}(X_\delta)\right)^2, \left\{X_{\delta\#}, \widehat{\omega}^2; S\right\}^{[2,2]}\right].$$

5 Relaxing some of the assumptions

5.1 Effect of endogenous and stochastically spaced data

So far our analysis has been based on measuring prices at regularly spaced intervals of length δ. In some ways it is more natural to work with returns measured in tick time and so it would be attractive if we could extend the above theory to cover stochastically spaced data. The convergence result inside QV is known to hold under very wide conditions that allow the spacing to be stochastic and endogenous. This is spelt out in, for example, Protter (2004, pp. 66-77) and Jacod and Shiryaev (2003, p. 51). It is important, likewise, to be able to derive central limit theorems for stochastically spaced data without assuming the times of measurement are independent of the underlying $\mathcal{BSM}$. This is emphasised by Renault and Werker (2005) in both their theoretical and empirical work.

Let $Y \in \mathcal{BSM}$ and assume we have measurements at times $t_j = T_{\delta j}$, $j = 1, 2, \ldots, n$, where $0 = t_0 < t_1 < \cdots < t_n = T_1$ and where T is a stochastic process of the form $T_t = \int_0^t \tau_u^2 du$, with τ having strictly positive, càdlàg sample paths. Then we can construct a new process $Z_t = Y_{T_t}$, so at the measurement times $Z_{\delta j} = Y_{T_{\delta j}}$ $j = 1, 2, \ldots, n$. Performing the analysis on observations of Z made at equally spaced times then allows one to analyse irregularly spaced data on Y. The following argument shows that $Z \in \mathcal{BSM}$ with spot volatility $\sigma_{T_t}\tau_t$ and so the analysis is straightforward. In particular, the feasible CLT

---

then

$$\begin{aligned}
R &= \left[\sum_{j=1}^n \left(x_{j,\cdot}^2 - 2\omega^2\right)\left(x_{j-2,\cdot}^2 - 2\omega^2\right)\right] - \sum_{j=1}^n y_{j,\cdot}^2 y_{j-2,\cdot}^2 = \sum_{j=1}^n y_{j,\cdot}^2 \varepsilon_{j-2} + \sum_{j=1}^n \varepsilon_j y_{j-2,\cdot}^2 + \sum_{j=1}^n \varepsilon_j \varepsilon_{j-2} \\
&\simeq \sum_{j=1}^n \varepsilon_j \left(y_{j-2,\cdot}^2 + y_{j+2,\cdot}^2\right) + \sum_{j=1}^n \varepsilon_j \varepsilon_{j-2}.
\end{aligned}$$

Now

$$\mathrm{Var}\left(\sum_{j=1}^n \varepsilon_j \left(y_{j-2,\cdot}^2 + y_{j+2,\cdot}^2\right)|Y\right) \simeq \frac{12\omega^4}{S}\sum_{j=1}^n \left(y_{j-2,\cdot}^2 + y_{j+2,\cdot}^2\right)^2 = O(n^{-1}S^{-1}).$$

So

$$\begin{aligned}
\mathrm{Var}(R|Y) &\simeq \mathrm{Var}\left(\sum_{j=1}^n \varepsilon_j\varepsilon_{j-2}|Y\right) = \sum_{j=1}^n \mathrm{Var}\left(\varepsilon_j\varepsilon_{j-2}|Y\right) + 2n\mathrm{Cov}\left(\varepsilon_j\varepsilon_{j-2}, \varepsilon_{j-1}\varepsilon_{j-3}|Y\right) \\
&= \sum_{j=1}^n \mathrm{E}\left(\varepsilon_j^2|Y\right)\mathrm{E}\left(\varepsilon_{j-2}^2|Y\right) + 2n\mathrm{Cov}\left(\varepsilon_j\varepsilon_{j-2}, \varepsilon_{j-1}\varepsilon_{j-3}|Y\right) \\
&\simeq \sum_{j=1}^n \left\{\frac{8\omega^4}{S} + \frac{8\omega^2}{S}\left(y_j^2\right)_\cdot\right\}\left\{\frac{8\omega^4}{S} + \frac{8\omega^2}{S}\left(y_{j-2}^2\right)_\cdot\right\} + \frac{n}{S^2}2\left(2\omega^2\right)^2 + \cdots \\
&= \frac{72\omega^8 n}{S^2} + O(S^{-2}).
\end{aligned}$$

is implemented by recording data every 5 trades, say, but then analysing it as if the spacing had been equidistant.

Write $Z = Y \circ T$ and $S_t = \int_0^t \sigma_u^2 \mathrm{d}u$. We assume that Y and T are adapted to a common filtration $\mathcal{F}_t$, which includes the history of the paths of T_u and $Y \circ T_u$ for $0 \leq u \leq t$. This assumption implies that σ_{u-} is in $\mathcal{F}_t$ for $0 \leq u \leq T_t$. Recall the key result (e.g. Revuz and Yor (1999, p. 181)) $[Z] = S \circ T$, while $Z \in \mathcal{M}_{loc}$. The following proposition shows that $[Z]$ is absolutely continuous and implies by the martingale representation theorem that Z is a stochastic volatility process with spot volatility of $\sigma_{T_t}\tau_t$.

Proposition 1 *Let $\upsilon_t = \sigma_{T_t}\tau_t$ and*

$$\Upsilon_t = \int_0^t \upsilon_u^2 \mathrm{d}u. \tag{23}$$

Then υ is a càdlàg process and $\Upsilon = S \circ T$.

The implication of this for kernels is that we can write

$$Z_t = \int_0^t a_{T_u}\tau_u \mathrm{d}u + \int_0^t \sigma_{T_u}\tau_u \mathrm{d}B_u^{\#},$$

where $B^{\#}$ is Brownian motion. Hence if we define a tick version of the kernel estimator

$$\gamma_h(Z_n)_t = \sum_{j=1}^{\lfloor t/\delta \rfloor} \left(Y \circ T_{\delta j} - Y \circ T_{\delta(j-1)}\right)\left(Y \circ T_{\delta(j-h)} - Y \circ T_{\delta(j-h-1)}\right),$$

$$\widetilde{K}(Z_n)_t = \gamma_0(Z_n)_t + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\left\{\gamma_h(Z_n)_t + \gamma_{-h}(Z_n)_t\right\},$$

then the theory for this process follows from the previous results. Thus using the symmetric kernel allows consistent inference on $[Z]_t = [Y]_{T_t}$.

5.2 Effect of serial dependence

So far we have assumed that $U \in \mathcal{WN}$. Now we will relax this assumption by considering kernels of the type

$$\overline{K}(X_\delta) = \sum_{h=-H}^{H} k\left(\frac{h}{H}\right)\gamma_h(X_\delta),$$

where as usual we write $k(0) = 1$ and $k(\pm 1) = 0$. To analyse these kernels it is helpful to write $\overline{\gamma}_h = \sum_{i=1}^n U_i U_{i-h}$.

Proposition 2 *Suppose U has a zero mean. If as $H \to \infty$*

$$\sum_{h=-H}^{H} a_{h,H} U_h = O_p(1), \qquad \textit{for any} \sum_{h=-H}^{H} a_{h,H}^2 = O(1),$$

then

$$\sum_{h=-H}^{H} k\left(\frac{h}{H}\right)\gamma_h(U_\delta) = \sum_{h=-H}^{H} k\left(\frac{h}{H}\right)\left(2\overline{\gamma}_h - \overline{\gamma}_{h-1} - \overline{\gamma}_{h+1}\right) + O_p(H^{-1/2}).$$

Further, suppose k is twice continuously differentiable, then

$$\sum_{h=-H}^{H} k\left(\frac{h}{H}\right)\gamma_h = -\frac{n}{H^2}\sum_{h=-H}^{H} k''\left(\frac{h}{H}\right)\left(\frac{1}{n}\overline{\gamma}_h\right) + O_p(nH^{-3}) + O_p(H^{-1/2}). \qquad (24)$$

Proposition 3 *We assume $k''(0) = 0$ and that U is an AR(1) process with persistence parameter ρ, then*

$$\sqrt{\frac{n}{H}}\left(\sum_{h=-H}^{H} k''\left(\frac{h}{H}\right)\left(\frac{1}{n}\overline{\gamma}_h\right)\right) \xrightarrow{d} N\left\{0, 4\omega^4\frac{1+\rho^2}{1-\rho^2}\int_0^1 k''(x)^2\mathrm{d}x\right\}.$$

This means that

$$\sum_{h=-H}^{H} k\left(\frac{h}{H}\right)\gamma_h = O_p\left(\frac{n^{1/2}}{H^{3/2}}\right) + O_p\left(\frac{n}{H^3}\right) + O_p(H^{-1/2}).$$

So if $H = cn^{1/2}$ then

$$\sum_{h=-H}^{H} k\left(\frac{h}{H}\right)\gamma_h = O_p\left(n^{-1/4}\right).$$

If we assume that $Y \perp\!\!\!\perp U$ then temporal dependence in U makes no difference to the asymptotic behaviour of $\gamma_h(U_\delta, Y_\delta)$ as $\delta \downarrow 0$ for the limit behaviour is driven by the local martingale difference behaviour of the increments of the Y process. The above results mean that if $H = cn^{2/3}$ then $\overline{K}(U_\delta) = O_p(n^{-1/3})$ which implies that this term has no impact on the asymptotic distribution of $\overline{K}(X_\delta)$. The same is not true when $H = cn^{1/2}$, since then $\overline{K}(U_\delta) = O_p(n^{-1/4})$ and so the rate of convergence of the realised kernel is not changed by serial dependence, but the asymptotic distribution is altered.

5.3 Endogenous noise

One of our key assumptions has been that $Y \perp\!\!\!\perp U$, that is the noise can be regarded as an exogenous process. Hence it is interesting to ask if our realised kernels continue to be consistent when U is endogenous. We do this under a simple linear model of endogeneity

$$U_{\delta i} = \sum_{h=0}^{\overline{H}} \beta_h\left(Y_{\delta(i-h)} - Y_{\delta(i-1-h)}\right) + \overline{U}_{\delta i},$$

where $Y \perp\!\!\!\perp \overline{U}$ and for simplicity we assume that $\overline{U} \in \mathcal{WN}$. Now

$$\gamma_h(Y_\delta, U_\delta) = \sum_{j=0}^{\overline{H}} \beta_j\gamma_{h+j}(Y_\delta) - \sum_{j=0}^{\overline{H}} \beta_j\gamma_{h+j+1}(Y_\delta) + \gamma_h(Y_\delta, \overline{U}_\delta).$$

Hence our asymptotic methods for studying the distribution of realised kernels under exogenous noise can be used to study the impact of endogenous noise on realised kernels through the limit theory we developed for $\gamma_h(Y_\delta)$ and $\gamma_h(Y_\delta, \overline{U}_\delta)$. In particular

$$\gamma_h(Y_\delta, U_\delta) - \gamma_h(Y_\delta, \overline{U}_\delta) = \begin{cases} \beta_0[Y] + O_p(n^{-1/2}), & h = 0, \\ -\beta_0[Y] + O_p(n^{-1/2}), & h = -1, \\ O_p(n^{-1/2}), & |h| \neq 1. \end{cases}$$

Hence flat-top kernels will be robust to this type of endogenous noise. An alternative approach to dealing with endogenous noise has been independently proposed by Kalnina and Linton (2006) using multiscale estimators.

6 Simulation study

6.1 Goal of the study

In this Section we report simulation results which assess the accuracy of the feasible asymptotic approximation for the realised kernel. A much more thorough analysis is provided in a Web Appendix to this paper available from `www.hha.dk/~alunde/bnhls/bnhls.htm`.

Before we turn our attention to feasible asymptotic distributions, we note the Web Appendix also reports on the accuracy of $\widetilde{K}(X_\delta)$ as an estimator of $\int_0^t \sigma_u^2 \mathrm{d}u$ and $\widehat{IQ}_{\delta,S}$ as an estimator of $\int_0^t \sigma_u^4 \mathrm{d}u$. The raw estimator $\widetilde{K}(X_\delta)$ may be negative, in which case we always truncate it at zero (the same technique is used for ML estimators of course). The Web Appendix shows this occurrence is extremely rare, even for small sample sizes, but increases with ω^2.

In this short section our focus will be assessing the infeasible and feasible central limit theories for $\widetilde{K}(X_\delta) - \int_0^t \sigma_u^2 \mathrm{d}u$. Throughout we simulate over the time interval $[0,1]$. We recall the asymptotic variance of $\widetilde{K}(X_\delta)_1$ is given in (15) which we write as ϖ here. This allows us to compute the asymptotic pivot

$$T_{raw} = \frac{\widetilde{K}(X_\delta)_1 - \int_0^1 \sigma_u^2 \mathrm{d}u}{\sqrt{\varpi}} \xrightarrow{L} N(0,1).$$

An alternative is to use the delta method and base the asymptotic analysis on (e.g. Barndorff-Nielsen and Shephard (2002) and Goncalves and Meddahi (2004))

$$T_{\log} = \frac{\log\left\{\widetilde{K}(X_\delta)_1 + d\right\} - \log\left\{\int_0^1 \sigma_u^2 \mathrm{d}u + d\right\}}{\sqrt{\varpi}/\left\{\widetilde{K}(X_\delta)_1 + d\right\}} \xrightarrow{L} N(0,1).$$

The presence of $d \geq 0$ allows for the possibility that $\widetilde{K}(X_\delta)_1$ may be truncated to be exactly zero. By selecting $d = 0.12$ we have the property that $\widetilde{K}(X_\delta)_1 + d$ is not negative in any of our Monte Carlo experiments. In our simulations we have taken $d = 0.2$.

In the infeasible case our simple rule-of-thumb for the choice of H is $H^* = 5.74\omega\sqrt{n/[Y]_1}$, which immediately gives us ϖ for in this case we assume knowledge of the path of σ. In practice this is less interesting than the feasible version, which puts $\widehat{H}^* = 5.74\widehat{\omega}\sqrt{n/[X_{\delta^*}]_1}$, where $[X_{\delta^*}]_1$ is the realised variance estimator based on low frequency data, such as 10 minute returns, which should not be too sensitive to market frictions. Having selected H, in the feasible case we can then compute $\widetilde{K}(X_\delta)$ and $\widehat{IQ}_{\delta,S}$ and so plug these into ϖ, replacing $\int_0^1 \sigma_u^2 \mathrm{d}u$ and $\int_0^1 \sigma_u^4 \mathrm{d}u$ respectively. Monte Carlo results reported in the Appendix suggest taking $S = \sqrt{n}$ in computing $\widehat{IQ}_{\delta,S}$.

6.2 Simulation design

Recall we simulate over the time interval $[0,1]$. We normalize one second to be $1/23400$, so that the interval $[0,1]$ is thought to cover 6.5 hours. The X process is generated using an Euler scheme based on $N = 23,400$ of intervals. We then construct sparsely sampled returns $X_{i/n} - X_{(i-1)/n}$, based on sample sizes n. In our Monte Carlo designs n takes on the values 195, 390, 780, $1,560$, $4,680$, $5,850$, $7,800$, $11,700$ and $23,400$. The case of 1 minute returns is when $n = 390$.

We consider the following SV model, which was also simulated by e.g. Huang and Tauchen (2005) and Goncalves and Meddahi (2004)

$$\mathrm{d}Y_t = \mu \mathrm{d}t + \sigma_t \mathrm{d}W_t, \quad \sigma_t = \exp\left(\beta_0 + \beta_1 \tau_t\right), \quad \mathrm{d}\tau_t = \alpha \tau_t \mathrm{d}t + \mathrm{d}B_t, \quad \mathrm{corr}(\mathrm{d}W_t, \mathrm{d}B_t) = \rho.$$

Here ρ is a leverage parameter. To make the results comparable to our constant volatility simulations reported in our Appendix we impose that $\mathrm{E}\left(\sigma_t^2\right) = 1$ by setting $\beta_0 = \beta_1^2/(2\alpha)$. We utilize the fact that the stationary distribution $\tau_t \sim N\left(0, (-2\alpha)^{-1}\right)$ to restart the process each day. In these experiments we set $\mu = 0.03$, $\beta_1 = 0.125$, $\alpha = -0.025$ and $\rho = -0.3$. The variance of σ is comparable to the empirical results found in e.g. Hansen and Lunde (2005). Finally, the market microstructure effects are modelled through ω^2. This is varied over 0.0001, 0.001 and 0.01, the latter being regarded as a very large effect indeed. These values are taken from the detailed study of Hansen and Lunde (2006).

6.3 Results

Table B.1 shows the Monte Carlo results for the infeasible asymptotic theory for T_{raw}, knowing a priori the value of ϖ. We can see from the Table that the results are rather good, although the asymptotics are slightly underestimating the mass of the distribution in the tails. The mean and standard deviations of T_{raw} show that the T-statistic is slightly overdispersed.

Table B.2 shows the results for the feasible asymptotic theory for T_{raw}. This indicates that the asymptotic theory does eventually kick in but it takes very large samples for it to provide anything like a good approximation. The reason for this is clearly that it is hard to consistently well approximate the integrated quarticity empirically. This result is familiar from the literature on realised volatility where the same phenomena is observed.

Table B.3 shows the results for the log version of the feasible theory based on $T_{\log}$. The accuracy of the asymptotic predictions does not seem to change very much with ω^2 and is much better than in the T_{raw} case. For small sample sizes extreme quantiles suffer from important distortions, but generally the asymptotics perform extremely well.

7 Empirical study

7.1 Analysis of General Electric trades in 2004

In this subsection we implement our efficient, feasible inference procedure for the daily increments of $[Y]$ for the realised kernel estimator on trades of General Electric (GE) shares carried out on the New York Stock Exchange in 2004. A more detailed analysis, including a comparison with results based on data from 2000 and on 29 other major stocks, is provided in our Web Appendix. We should note that the variance of the noise was around 10 times higher in 2000 than in 2004 and so looking over both periods is instructive. This Appendix also details the cleaning we carried out on the data before it was analysed and the precise way we calculated all of our statistics.

Our realised kernel will be implemented on returns recorded every k trades, where k is selected each day so that there are approximately 360 observations a day[4]. This means

[4]As our sample size is quite large it is important to calculate it in tick time in order not to be influenced by the bias effect discussed by Renault and Werker (2005) caused by sampling in calender time.

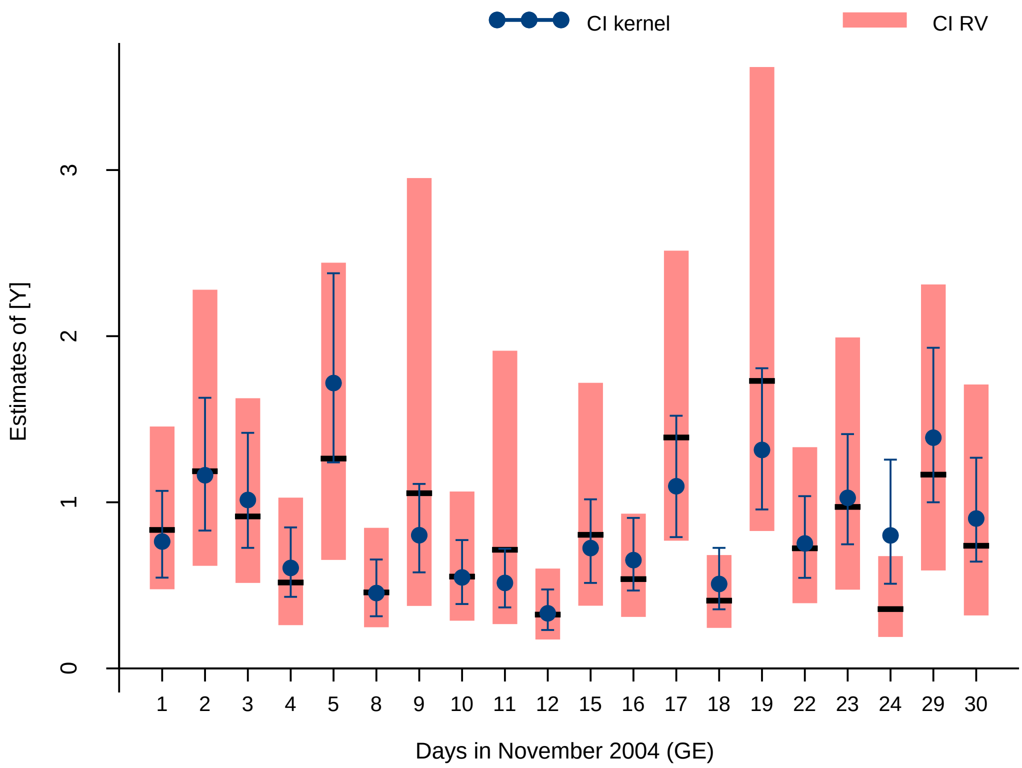

Figure 1: *Confidence intervals for the daily increments to $[Y]$ for General Electrics (GE) in November 2004. Rectangles denote the 95% confidence intervals based on 20 minute returns using the Barndorff-Nielsen and Shephard (2002) feasible realised variance inference method. The other interval corresponds to our realised kernel, sampling in tick times so the period over which returns are calculated is roughly 60 seconds.*

that on average these returns are recorded every 60 seconds. This inference method will be compared to the feasible procedure of Barndorff-Nielsen and Shephard (2002), which ignores the presence of market microstructure effects, based on returns calculated over 20 minutes within each day. This baseline was chosen as Hansen and Lunde (2006) has suggested that the Barndorff-Nielsen and Shephard (2002) method was empirically sound when based on that type of interval for thickly traded stocks.

General Electric shares are traded very frequently on the NYSE. A typical day results in between $1,500$ and $6,000$ trades. For this stock Hansen and Lunde (2006) have presented detailed work which suggests that over 60 second intervals it is empirically reasonable to assume that Y and U are uncorrelated and U is roughly a white noise process. Hence the main assumptions behind the inference procedure for our efficient kernel estimator are roughly satisfied and so we feel comfortable implementing the feasible limit theory on this dataset. We should note that on all the days in 2004 our realised kernel estimator of the daily increments of $[Y]$ was positive. In the 2004 sample period, we found $\widehat{\omega}^2 = [X_{1\text{min}}]/2n$ to be very small and it was used to calculate $\hat{H}^*$ and ϖ. Due to the positive bias in $\widehat{\omega}^2$, this will result in conservative confidence intervals.

Figure 1 shows daily 95% confidence intervals (CIs) for the realised kernel for November 2004 using the modified Tukey-Hanning weights (21) with $H = cn^{1/2}$. Also drawn are the corresponding results for the realised variance. We can see the realised kernel has much shorter CIs. The width of these intervals does change through time, with them tending to be slightly wider in high volatility periods. Over the entire year there are only 3 days when the CIs do not overlap.

Day	Trans	Lower	RV20m	Upper	n	Lower	KV60s	Upper	k	n	H	$\widehat{\omega}^2$
1	4631	0.48	0.83	1.46	20	0.55	0.76	1.07	13	357	4	0.0016
2	4974	0.62	1.19	2.28	20	0.83	1.16	1.63	14	356	4	0.0025
3	4918	0.51	0.92	1.63	20	0.73	1.01	1.42	14	352	4	0.0021
4	5493	0.26	0.52	1.03	20	0.43	0.60	0.85	16	344	4	0.0013
5	5504	0.65	1.26	2.44	20	1.24	1.72	2.38	16	344	3	0.0028
8	4686	0.25	0.46	0.85	20	0.31	0.45	0.66	14	335	4	0.0014
9	4923	0.38	1.05	2.95	20	0.58	0.80	1.11	14	352	3	0.0014
10	4970	0.29	0.55	1.07	20	0.39	0.55	0.77	14	355	4	0.0013
11	4667	0.27	0.71	1.91	20	0.37	0.51	0.72	13	359	4	0.0011
12	4822	0.17	0.32	0.60	20	0.23	0.33	0.48	14	345	4	0.0009
15	4681	0.38	0.80	1.72	20	0.52	0.72	1.02	14	335	4	0.0015
16	4526	0.31	0.54	0.93	20	0.47	0.65	0.91	13	349	4	0.0011
17	5477	0.77	1.39	2.51	20	0.79	1.10	1.52	16	343	3	0.0018
18	4738	0.24	0.41	0.68	20	0.36	0.51	0.73	14	339	4	0.0014
19	5224	0.83	1.73	3.62	20	0.96	1.32	1.81	15	349	3	0.0019
22	5359	0.39	0.72	1.33	20	0.55	0.75	1.04	15	358	3	0.0012
23	5405	0.47	0.97	1.99	20	0.75	1.03	1.41	15	361	3	0.0016
24	4626	0.19	0.36	0.68	20	0.51	0.80	1.26	13	356	3	0.0013
29	4709	0.59	1.17	2.31	20	1.00	1.39	1.93	14	337	3	0.0023
30	4719	0.32	0.74	1.71	20	0.64	0.90	1.27	14	338	4	0.0018

Table 5: *Inference for General Electric (GE) volatility in November 2004. Trans denotes the number of transactions on that day. RV20m is the daily* $[X_{20\ minutes}]$*. KV60s denotes* $\widetilde{K}_w^{\mathrm{TH2}}(X_{ap.\ 1\ min})$*, that is the corresponding corresponding realised kernel calculating returns every* k *observations.* n *is the sample size per day,* H *is the number of lags in the kernel and* $\widehat{\omega}^2 = [X_{1min}; 60]/2n$*.*

Table 5 shows the details of these results for November 2004. The estimates of ω^2 are very small, ranging from about 0.001 to 0.003. These are in the range of the small to medium levels of noise set out in our Monte Carlo designs discussed in the previous Section. The Table shows the sample size for the realised kernel, which is between 335 and 361 intervals of roughly 60 seconds. Typically each interval corresponds to about 15 trades. It records the daily selected value of H that ranges from 3 to 4, which is rather modest and is driven by the fact that ω^2 is quite small.

Table 6 provides summary statistics for some alternative estimators over the entire year. This suggests the other realised kernel estimators have roughly the same average value and that they are quite tightly correlated. The Table also records the summary statistics for the realised variance computed using 20, 5, 1 minute and 10, and 1 second intervals. The last two of these estimators show a substantially higher mean. Interestingly, the realised QV based on 5 minute sampling is most correlated with the realized kernels. This is in line with the optimal sampling frequencies for the realised QV reported in Bandi and Russell (2005a). The realised kernels have a stronger degree of serial dependence than our benchmark realised QV, $[X_{20\ \mathrm{minutes}}]$. This point suggests the realised kernel may be useful when it comes to forecasting, extending the exciting work of Andersen, Bollerslev, Diebold, and Labys (2001). The high serial dependence found in the realised QVs based on the high sampling frequencies suggests a strong dependence in the bias components of these estimators.

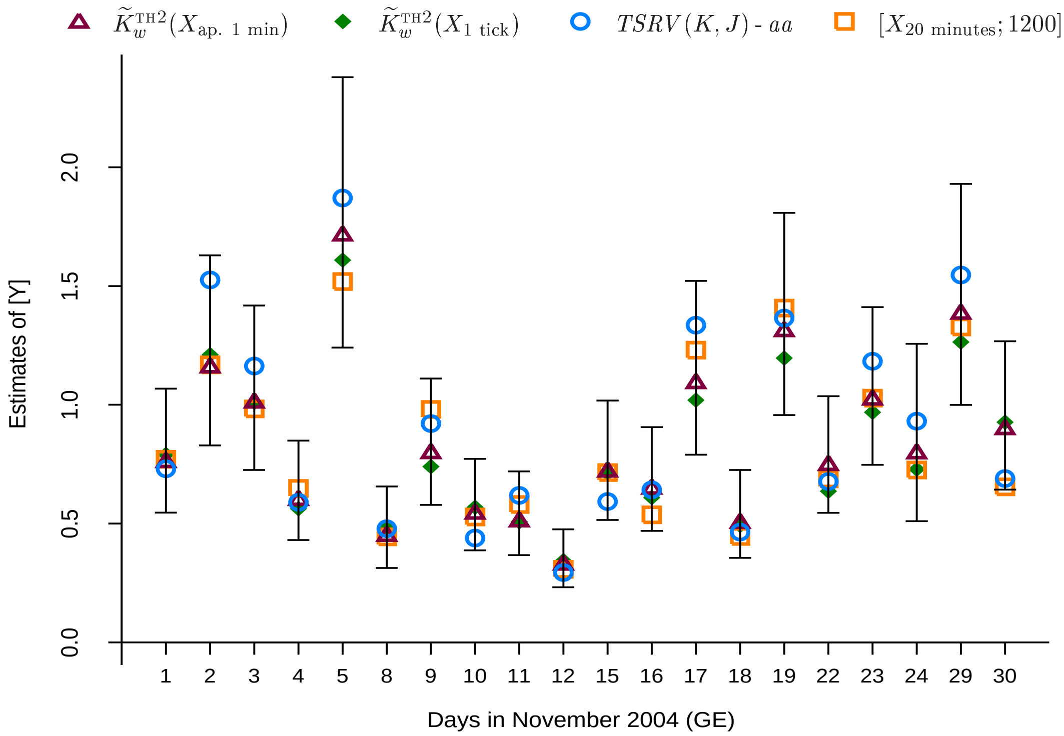

Figure 2: *Four estimators for the daily increments to $[Y]$ for General Electrics in November 2004. The intervals are the confidence intervals for our realised Modified Tukey-Hanning kernel based on returns sampled roughly every 60 seconds. The triangles denote the subsampled version of this realised kernel. Diamonds denote our modified Tukey-Hanning kernel based on all trades, circles represents $[X_{20\ minutes}; 1200]$ (calculates RV over 20 minutes returns, averaged over 1200 times, just changing the initial place prices are recorded). Squares ($TSRV(K, J)$ - aa) denote the bias adjusted Aït-Sahalia, Mykland, and Zhang (2006, eq. (4.22)) two scale estimator.*

7.2 Speculative analysis

The analysis in the previous subsection does not use all of the available data efficiently, for the realised kernel is computed only on every 15 or so trades. This was carried out so that the empirical reality of the GE data matched the assumptions of our feasible central limit theory, allowing us to calculate daily confidence intervals. In this subsection, we give up on the goal of carrying out inference and simply focus on estimating $[Y]$ by employing all of the data. The results in Section 6 suggest our efficient realised kernel can do this, even though the white noise assumption and independence of Y and U are no longer empirically well-grounded assumptions. For these robust estimators we select $H = cn^{2/3}$, where we use the same values for c as in the previous subsection. Inevitably then, the results in this subsection will be more speculative than those given in the previous analysis.

We calculate the realised kernel using every trade on each day, based on returns sampled roughly every 60 seconds, or by applying the kernel weights to returns sampled every trade. The time series of these estimators are drawn in Figure 2, together with the corresponding bias corrected two scale estimator and a subsampled version of the realised variance estimator using 5 minute returns, where the degree of subsampling was selected to exhaust the available data. For the sake of comparison, we also include the confidence intervals from Figure 1. Figure 2 shows that realised kernels give very similar estimates – on some days the estimates are almost identical. The uncorrected two scale estimators and the bias corrected two scale estimator based on white noise are quite biased, but the two scale estimator based on being robust to the white noise assumption[5] (denoted TSRV(K,J)-aa) seems in line with the results for the realised kernels and the subsampled RV estimators. Table 7 provides summary statistics of these estimators. The realised kernels are pretty robust to choice of the design of the weights.

8 Conclusions

In this paper we have provided a detailed analysis of the accuracy of realised kernels as estimators of quadratic variation when an efficient price is obscured by simple market frictions. We show how to make these estimators consistent and derive central limit theorems for the estimators under various assumptions about the kernel weights. Such estimators can be made to converge at the fastest possible rate and are very close to being efficient. They can be made robust to dynamics in the noise process, robust to endogenous market frictions and robust to endogenous spacing in the timing of the data. The last two of these features are new to this literature.

Our efficient feasible central limit theory for our estimators performed satisfactorily in Monte Carlo experiments designed to assess finite sample behaviour. Our kernel was shown to be consistent under rather broad assumptions on the dynamics of the noise term. We have applied the estimator empirically, using 60 second return data on General Electric transaction data for 2004. Feasible inference for our realised kernel is compared with that for a simpler realised variance estimator based on 20 minute returns. The empirical results suggest that the realized kernel estimator is more accurate. Its serial correlation suggests that the realized kernel may be useful for forecasting, following Andersen, Bollerslev, Diebold, and Labys (2001).

There are many possible extensions to this work, e.g. multivariate versions of these results which deal with the scrambling effects discussed by, for example, Hayashi and Yoshida (2005), Bandi and Russell (2005c), Zhang (2005), Sheppard (2005), Voev and Lunde (2005) and Griffin and Oomen (2006) and derive an asymptotically efficient choice of kernel under temporal dependence in U.

[5]In empirical work we found this statistic to be sensitive to the choice of K. To be consistent with our empirical findings J has to be about 15 (yielding returns measured roughly over 1 minute). Aït-Sahalia, Mykland, and Zhang (2006) show K has to much larger than J, but their automatic selection formula for K typically selects K smaller than J, so we initially imposed $K \geq 2J$. That was not enough to get good empirical results and so we finally imposed $K \geq 5J$, which worked well in practice.

References

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies 18*, 351–416.

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2006). Ultra high frequency volatility estimation with dependent microstructure noise. Unpublished paper: Department of Economics, Princeton University.

Aldous, D. J. and G. K. Eagleson (1978). On mixing and stability of limit theorems. *Annals of Probability 6*, 325–331.

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

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2000). Great realizations. *Risk 13*, 105–108.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of exchange rate volatility. *Journal of the American Statistical Association 96*, 42–55. Correction published in 2003, volume 98, page 501.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*, 817–858.

Back, K. (1991). Asset pricing for general processes. *Journal of Mathematical Economics 20*, 371–395.

Bandi, F. M. and J. R. Russell (2005a). Microstructure noise, realized variance, and optimal sampling. Unpublished paper, Graduate School of Business, University of Chicago.

Bandi, F. M. and J. R. Russell (2005b). On the finite sample properties of kernel-based integrated variance estimators. Unpublished paper, Graduate School of Business, University of Chicago.

Bandi, F. M. and J. R. Russell (2005c). Realized covariation, realized beta and microstructure noise. Unpublished paper, Graduate School of Business, University of Chicago.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In Y. Kabanov, R. Lipster, and J. Stoyanov (Eds.), *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*, pp. 33–68. Springer.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, and N. Shephard (2006). Limit theorems for realised bipower variation in econometrics. *Econometric Theory 22*. Forthcoming.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*, 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica 72*, 885–925.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*, 1–30.

Bartlett, M. S. (1946). On the theoretical specification of sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society, Supplement 8*, 27–41.

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

Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance 8*, 291–323.

Delattre, S. and J. Jacod (1997). A central limit theorem for normalized functions of the increments of a diffusion process in the presence of round off errors. *Bernoulli 3*, 1–28.

Fang, Y. (1996). Volatility modeling and estimation of high-frequency data with Gaussian noise. Unpublished Ph.D. thesis, Sloan School of Management, MIT.

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

Gallant, A. R. (1987). *Nonlinear Statistical Models*. New York: John Wiley.

Ghysels, E., A. C. Harvey, and E. Renault (1996). Stochastic volatility. In C. R. Rao and G. S. Maddala (Eds.), *Statistical Methods in Finance*, pp. 119–191. Amsterdam: North-Holland.

Gloter, A. and J. Jacod (2001a). Diffusions with measurement errors. I — local asymptotic normality. *ESAIM: Probability and Statistics 5*, 225–242.

Gloter, A. and J. Jacod (2001b). Diffusions with measurement errors. II — measurement errors. *ESAIM: Probability and Statistics 5*, 243–260.

Goncalves, S. and N. Meddahi (2004). Bootstrapping realized volatility. Unpublished paper, CIRANO, Montreal.

Griffin, J. E. and R. C. A. Oomen (2006). Covariance measurement in the presence of non-synchronous trading and market microstructure noise. Unpublished: Dept. of Statistics, University of Warwick.

Hall, P. and C. C. Heyde (1980). *Martingale Limit Theory and its Applications*. San Diego: Academic Press.

Hansen, P. R., J. Large, and A. Lunde (2005). Moving average-based estimators of integrated variance. *Econometric Reviews*. Submitted.

Hansen, P. R. and A. Lunde (2005). A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics 3*, 525–554.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise (with discussion). *Journal of Business and Economic Statistics 24*, 127–218.

Hayashi, T. and N. Yoshida (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli 11*, 359–379.

Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variation. *Journal of Financial Econometrics 3*, 456–499.

Jacod, J. (1994). Limit of random measures associated with the increments of a Brownian semimartingale. Preprint number 120, Laboratoire de Probabilitiés, Université Pierre et Marie Curie, Paris.

Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. In J. Azema, M. Emery, and M. Yor (Eds.), *Séminaire Probability XXXI, Lecture Notes in Mathematics*, Volume 1655, pp. 232–246. Berlin: Springer Verlag.

Jacod, J. and P. Protter (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability 26*, 267–307.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (2 ed.). Springer: Berlin.

Kalnina, I. and O. Linton (2006). Estimating quadratic variation consistently in the presence of correlated measurement error. Unpublished paper: Department of Economics, LSE.

Large, J. (2005). Estimating quadratic variation when quoted prices jump by a constant increment. Unpublished paper: Nuffield College, Oxford.

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatilities. *Journal of Applied Econometrics 17*, 479–508.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics 8*, 323–361.

Mykland, P. A. (2006). A Gaussian calculus for inference from high frequency data. Unpublished paper: Department of Statistics, University of Chicago.

Newey, W. K. and K. D. West (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–708.

Oomen, R. A. A. (2005). Properties of bias corrected realized variance in calender time and business time. *Journal of Financial Econometrics 3*, 555–577.

Park, S. K. and R. A. Schowengerdt (1983). Image reconstruction by parametric cubic convolution. *Computer Vision, Graphics, and Image Processing 3*, 258–272.

Phillips, P. C. B. and S. Ouliaris (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica 58*, 165–193.

Phillips, P. C. B., Y. Sun, and S. Jin (2003). Long run variance estimation using steep origin kernels without truncation. Unpublished paper: Cowles Foundation, Yale University.

Phillips, P. C. B. and J. Yu (2006). Discussion of Hansen and Lunde (2006). *Journal of Business and Economic Statistics 24*, 202–208.

Politis, D. N. (2005). Higher-order accurate, positive semi-definite estimation of large-sample covariance and spectral density matrices. Unpublished paper: Department of Mathematics, USCD.

Politis, D. N. and J. P. Romano (1995). Bias-corrected nonparametric spectral estimation. *Journal of Time Series Analysis 16*, 67–103.

Protter, P. (2004). *Stochastic Integration and Differential Equations*. New York: Springer-Verlag.

Renault, E. and B. Werker (2005). Stochastic volatility models with transaction time risk. Department of Economics, University of North Carolina.

Rényi, A. (1963). On stable sequences of events. *Sankya, Series A 25*, 293–302.

Revuz, D. and M. Yor (1999). *Continuous Martingales and Brownian motion* (3 ed.). Heidelberg: Springer.

Rosenberg, B. (1972). The behaviour of random variables with nonstationary variance and the distribution of security prices. Working paper 11, Graduate School of Business Administration, University of California, Berkeley. Reprinted in Shephard (2005).

Shephard, N. (2005). *Stochastic Volatility: Selected Readings*. Oxford: Oxford University Press.

Sheppard, K. (2005). Measuring realized covariance. Unpublished paper: Department of Economics, University of Oxford.

Stein, M. L. (1987). Minimum norm quadratic estimation of spatial variograms. *Journal of the American Statistical Association 82*, 765–772.

Voev, V. and A. Lunde (2005). Integrated covariance estimation using high-frequency data in the presence of noise. Unpublished paper, Department of Economics, University of Konstanz.

Zhang, L. (2005). Estimating covariation: Epps effect and microstructure noise. Unpublished paper, Department of Finance, University of Illinois, Chicago.

Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli*. Forthcoming.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*, 1394–1411.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics 14*, 45–52.

Appendix A: Proofs

Proof of Theorem 1. Throughout we take $t = 1$ and so $\delta = 1/n$ for the other cases follow trivially.

PURE $\mathcal{BSM}$ CASE. Write $y_j = Y_{j/n} - Y_{(j-1)/n}$, then the terms we need to study are $\sum_{j=1}^{n} y_j^2$, $\sum_{j=1}^{n} y_j y_{j+1}, \ldots, \sum_{j=1}^{n} y_j y_{j+H}$. This can be written in the form of a set of multipower variation statistics (e.g. Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006))

$$\sum_{j=1}^{n} \prod_{k=0}^{H} g_{l,k}(y_{j+k}) = \sum_{j=1}^{n} g_l(y_j, y_{j+1}, \ldots, y_{j+H}), \quad l = 0, 1, 2, \ldots, H,$$

by selecting the functions $g_{l,k}$ appropriately. In particular, writing $g_{l,k}$ into a matrix form

$$g(x) = \begin{pmatrix} x_0^2 & 1 & 1 & \cdots & 1 \\ x_0 & x_1 & 1 & \cdots & 1 \\ x_0 & 1 & x_2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \ddots & \\ x_0 & 1 & 1 & \cdots & x_H \end{pmatrix}, \quad x \in \mathbb{R}^{H+1}.$$

We satisfy all the conditions in Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006) except there the $g_{l,k}(x)$ are assumed all to be even functions. To see that for our specific form of g this assumption of evenness does not matter, we will look solely at the $\sum_{j=1}^{n} y_j y_{j+1}$ statistic. The other terms then follow immediately by the same argument.

We think of the bipower variation statistic

$$\frac{1}{n} \sum_{j=1}^{n} g_{2,1}\left(\sqrt{n} y_j\right) g_{2,2}\left(\sqrt{n} y_{j+1}\right),$$

where $g_{2,1}(x_0) = x_0$ and $g_{2,2}(x_1) = x_1$. Then using the notation of Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006) that $\rho_\sigma(h) = \mathrm{E}\left(h\left(x\right)|\sigma^2\right)$, $x|\sigma^2 \sim N(0, \sigma^2)$ we note that $\rho_\sigma(g_{2,1}) = \rho_\sigma(g_{2.2}) = 0$, which enormously simplifies the task. Inspection of their proof shows two steps use this assumption. It is used on page 67, where various features of their z_i^n are defined and studied. In our case $z_i^n = 0$ and so they follow immediately.

The only non-trivial step involves their equation (4.12) applied to the bipower case which is presented in the first equation of their Proposition 4.2. This corresponds to checking condition (7.29) in Jacod and Shiryaev (2003). The sole task then is satisfied if we can show

$$\sum_{j=1}^{n} \mathrm{E}\left(\zeta_j^n w_j | \mathcal{F}_{\frac{j-1}{n}}\right) \xrightarrow{p} 0,$$

(actually converging to a continuous process would be enough) where we define

$$\zeta_j^n = \frac{1}{\sqrt{n}} g_{2,1}(\beta_{j-1}) g_{2,2}(\beta_{j-1}'),$$

with $\beta_j = \sqrt{n}\sigma_{\frac{j-1}{n}} w_j$, $\beta'_j = \sqrt{n}\sigma_{\frac{j-1}{n}} w_{j+1}$, and $w_j = W_{\frac{j}{n}} - W_{\frac{j-1}{n}}$. Thus $\zeta^n_j = \sqrt{n}\sigma^2_{\frac{i-2}{n}} w_{j-1} w_j$. Clearly

$$\sum_{j=1}^{n} \mathrm{E}\left(\zeta^n_j \Delta^n_j W | \mathcal{F}_{\frac{j-1}{n}}\right) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \sigma^2_{\frac{j-2}{n}} \left(\Delta^n_{j-1} W\right) = \frac{1}{\sqrt{n}} \int_0^1 \sigma^2_u \mathrm{d}W_u + o_p(n^{-1/2}) \xrightarrow{p} 0.$$

Hence the result holds.

This implies then that Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006) result holds and so

$$\delta^{-1/2} \begin{pmatrix} [Y_\delta]_t - \int_0^t \sigma^2_u \mathrm{d}u \\ \gamma_1(Y_\delta) \\ \vdots \\ \gamma_H(Y_\delta) \end{pmatrix} \xrightarrow{Ls} MN\left(0, \int_0^t A(\sigma_u, g)\mathrm{d}u\right)$$

where

$$\begin{aligned} A_{l,k}(\sigma, g) &= \mathrm{Cov}\left\{g_l(x), g_k(x)|\sigma^2\right\}, \quad x|\sigma^2 \sim N(0, \sigma^2 I) \\ &= \sigma^4 \mathrm{Cov}\left\{g_l(x), g_k(x)\right\}, \quad\quad x \sim N(0, I). \end{aligned}$$

Simple calculations based on the normal distribution delivers the result immediately.
THE CROSS TERM. Be implicit about the dependence on δ, then we can express $\gamma_0(Y_\delta, U_\delta)$ and $\widetilde{\gamma}_h(Y_\delta, U_\delta)$ as

$$\sum_{j=1}^{n} y_j(U_j - U_{j-1}) = \sum_{j=1}^{n} (y_j - y_{j+1}) U_j + O_p(n^{-1/2})$$

$$\begin{aligned} &\sum_{j=1}^{n} y_j(U_{j-h} - U_{j-h-1} + U_{j+h} - U_{j+h-1}) \\ &\quad = \sum_{j=1}^{n} (y_{j+h} - y_{j+h+1} + y_{j-h} - y_{j-h-1}) U_j + O_p(n^{-1/2}). \end{aligned}$$

Then the result follows immediately as $\gamma_h(Y_\delta) = O_p(n^{-1/2})$ for $|h| > 0$.
PURE NOISE CASE. The expectation and covariance of $\gamma(U_\delta)$ and $\widetilde{\gamma}(U_\delta)$ can be computed in the following way. We write the terms into uncorrelated items

$$\begin{aligned} [U_\delta] &= \sum_{j=1}^{n} (U_j - U_{j-1})^2 = \sum_{j=1}^{n} U_j^2 + \sum_{j=1}^{n} U_{j-1}^2 - 2\sum_{j=1}^{n} U_j U_{j-1} \\ &= 2\sum_{j=1}^{n-1} U_j^2 - 2\sum_{j=1}^{n-1} U_j U_{j-1} + \left(U_n^2 + U_0^2 - 2U_n U_{n-1}\right). \\ \gamma_1(U) &= \sum_{j=1}^{n} (U_j - U_{j-1})(U_{j-1} - U_{j-2}) \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^{n} U_j U_{j-1} - \sum_{j=1}^{n} U_{j-1}^2 - \sum_{j=1}^{n} U_j U_{j-2} + \sum_{j=1}^{n} U_{j-1} U_{j-2} \\
&= 2\sum_{j=1}^{n-1} U_j U_{j-1} - \sum_{j=1}^{n-1} U_j U_{j-2} - \sum_{j=1}^{n-1} U_j^2 \\
&\qquad\qquad + \left(-U_n U_{n-2} + U_n U_{n-1} + U_0 U_{-1} - U_0^2\right).
\end{aligned}$$

$$\begin{aligned}
\gamma_2(U) &= \sum_{j=1}^{n} (U_j - U_{j-1})(U_{j-2} - U_{j-3}) \\
&= \sum_{j=1}^{n} U_j U_{j-2} - \sum_{j=1}^{n} U_{j-1} U_{j-2} - \sum_{j=1}^{n} U_j U_{j-3} + \sum_{j=1}^{n} U_{j-1} U_{j-3} \\
&= 2\sum_{j=1}^{n-1} U_j U_{j-2} - \sum_{j=1}^{n-1} U_j U_{j-1} \\
&\qquad - \sum_{j=1}^{n-1} U_j U_{j-3} + \left(-U_n U_{n-3} + U_n U_{n-2} + U_0 U_{-2} - U_0 U_{-1}\right).
\end{aligned}$$

The relevant covariance matrix can then be computed straightforwardly.

Finally, we show that end effects matter in the pure noise case. For simplicity we consider solely the $h = 1$ case. Writing $n = \lfloor t/\delta \rfloor$

$$\begin{aligned}
\gamma_1(U_\delta) &= 2\sum_{j=1}^{n-1} U_{j\delta} U_{(j-1)\delta} - \sum_{j=1}^{n-1} U_{j\delta} U_{(j-2)\delta} - \sum_{j=1}^{n-1} U_{j\delta}^2 + G_1 \\
\gamma_{-1}(U_\delta) &= 2\sum_{j=1}^{n-1} U_{j\delta} U_{(j-1)\delta} - \sum_{j=1}^{n-1} U_{j\delta} U_{(j-2)\delta} - \sum_{j=1}^{n-1} U_{j\delta}^2 + G_{-1},
\end{aligned}$$

with end effects: $G_1 = -U_{n\delta} U_{(n-2)\delta} + U_{n\delta} U_{(n-1)\delta} + U_0 U_{(-1)\delta} - U_0^2$, and $G_{-1} = -U_{(n+1)\delta} U_{(n-1)\delta} + U_{(n+1)\delta} U_{n\delta} + U_{1\delta} U_0 - U_{n\delta}^2$. The terms U_0^2 and $U_{n\delta}^2$ differ and they never go away as n gets large. □

Proof of Lemma 1. Recall that $Y_n^{\#} \xrightarrow{Ls} Y^{\#}$ means that for any bounded random variable V on $(\Omega, \mathcal{F}, P)$ and any bounded continuous function f, as $n \to \infty$, $\mathrm{E}\left\{Vf\left(Y_n^{\#}\right)\right\} \to \mathrm{E}\left\{Vf\left(Y^{\#}\right)\right\}$.

To show that $X_n^{\#} \xrightarrow{Ls} Y^{\#}$ it suffices to verify that for any given V and f, as above,

$$\mathrm{E}\left\{Vf\left(X_n^{\#}\right)\right\} - \mathrm{E}\left\{Vf\left(Y_n^{\#}\right)\right\} \to 0. \tag{B.1}$$

Now, let c and c' be constants such that $|V| \le c < \infty$ and $|f| \le c' < \infty$, then

$$\begin{aligned}
\left|\mathrm{E}\left\{Vf\left(X_n^{\#}\right)\right\} - \mathrm{E}\left\{Vf\left(Y_n^{\#}\right)\right\}\right| &\le \mathrm{E}\left\{|V|\left|f\left(X_n^{\#}\right) - f\left(Y_n^{\#}\right)\right|\right\} \\
&\le c\mathrm{E}\left\{\left|f\left(X_n^{\#}\right) - f\left(Y_n^{\#}\right)\right|\right\}.
\end{aligned}$$

By the assumed properties of $Y_n^{\#}$ and $Z_n^{\#}$, to any $\varepsilon > 0$ there exists an n_0 and a finite closed interval I such that for any $n > n_0$ we have $\Pr\left\{Y_n^{\#} \notin I \text{ or } X_n^{\#} \notin I\right\} < \varepsilon$ and hence, for $n > n_0$,

$$\mathrm{E}\left\{\left|f\left(X_n^{\#}\right) - f\left(Y_n^{\#}\right)\right|\right\} \le c'\varepsilon + \mathrm{E}\left\{\left|f\left(X_n^{\#}\right) - f\left(Y_n^{\#}\right)\right| \mathbf{1}_{\left\{Y_n^{\#} \in I \text{ and } X_n^{\#} \in I\right\}}\right\}.$$

Since f is uniformly continuous on I there exists a $\delta = \delta(\varepsilon)$ such that

$$|f(x) - f(y)| < \varepsilon \quad \text{for all } x, y \in I \text{ with } |x - y| < \delta.$$

Next, take $n_0' > n_0$ and so large that $P\left\{\left|Z_n^{\#}\right| \geq \delta\right\} < \varepsilon$ provided $n > n_0'$. Then, for such n

$$\begin{aligned}
&\mathrm{E}\left\{\left|f\left(X_n^{\#}\right) - f\left(Y_n^{\#}\right)\right| \mathbf{1}_{\left\{Y_n^{\#} \in I \text{ and } X_n^{\#} \in I\right\}}\right\} \\
&\quad \leq c'\varepsilon + \mathrm{E}\left\{\left|f\left(X_n^{\#}\right) - f\left(Y_n^{\#}\right)\right| \mathbf{1}_{\left\{Y_n^{\#} \in I \text{ and } X_n^{\#} \in I\right\}} \mathbf{1}_{\left\{\left|Z_n^{\#}\right| < \delta\right\}}\right\} \\
&\quad \leq c'\varepsilon + \mathrm{E}\left\{\left|f\left(X_n^{\#}\right) - f\left(Y_n^{\#}\right)\right| \mathbf{1}_{\left\{\left|Z_n^{\#}\right| < \delta\right\}}\right\} \leq c'\varepsilon + \varepsilon.
\end{aligned}$$

All in all we therefore have that $\left|\mathrm{E}\left\{Vf\left(X_n^{\#}\right)\right\} - \mathrm{E}\left\{Vf\left(Y_n^{\#}\right)\right\}\right| \leq c\left(2c'\varepsilon + \varepsilon\right)$ for $n > n_0'$ from which (B.1) follows. $\square$

Proof of Proposition 1. The càdlàg property of υ follows by direct argument. Further, by Lebesgue's Theorem, the integral (23) is the same whether interpreted as a Riemann integral or a Lebesgue integral. With the latter interpretation we find

$$\Upsilon_t = \int_0^t \sigma_{T_u}^2 \tau_u^2 \mathrm{d}u = \int_0^t \sigma_{T_u}^2 \mathrm{d}T_u = \int_0^{T_t} \sigma_u^2 \mathrm{d}u = S \circ T_t.$$

$\square$

Proof of Theorem 2. Write $w = (\overline{w}^\intercal, v^\intercal)^\intercal$ where $\overline{w} = (1,1)^\intercal$ and v is a $(H-1) \times 1$ vector. Then the variances of $K_v(U_\delta)$ and $\widetilde{K}_v(U_\delta)$ are, respectively,

$$V = 4\omega^4 \left\{nV_C(v) + V_D(v)\right\}, \quad \widetilde{V} = 4\omega^4 \left\{nV_C(v) + V_{\widetilde{D}}(v)\right\}, \tag{B.2}$$

where, for a generic matrix E which splits as before into blocks E_{11}, E_{21}, E_{22},

$$V_E(v) = \overline{w}^\intercal E_{11} \overline{w} + 2v^\intercal E_{21} \overline{w} + v^\intercal E_{22} v.$$

Likewise, under $Y \in \mathcal{BSM}$ the conditional variance of $\widetilde{K}_v(Y_\delta, U_\delta)$ given Y is

$$\mathrm{Var}\left\{\widetilde{K}_v(Y_\delta, U_\delta)|Y\right\} = 2\omega^2 V_B(v)[Y] + O(\delta),$$

and we have that

$$\frac{n^{1/2}}{\sqrt{t\left(3 + 2\sum_{h=2}^H v_h^2\right)}} \left(\widetilde{K}_v(Y_\delta) - \int_0^t \sigma_u^2 \mathrm{d}u\right) \xrightarrow{Ls} MN\left(0, 2t\int_0^t \sigma_u^4 \mathrm{d}u\right). \tag{B.3}$$

We study the large sample behaviour of $v^\intercal A_{22} v$, $v^\intercal B_{22} v$, $v^\intercal C_{22} v$, $v^\intercal D_{22} v$.

$$\begin{aligned}
v^\intercal A_{22} v &= \sum_{i=1}^H v_i^2, \\
v^\intercal B_{22} v &= v_1(2v_1 - v_2) + \sum_{i=2}^{H-1} v_i(-v_{i-1} + 2v_i - v_{i+1}) + v_H(-v_{i-1} + 2v_i),
\end{aligned}$$

$$
\begin{aligned}
v^\intercal C_{22} v = \ & v_1(6v_1 - 4v_2 + v_3) + v_2(-4v_1 + 6v_2 - 4v_3 + v_4) \\
& + \sum_{i=3}^{H-2} v_i(v_{i-2} - 4v_{i-1} + 6v_i - 4v_{i+1} + v_{i+2}) \\
& + v_{H-1}(v_{H-3} - 4v_{H-2} + 6v_{H-1} - 4v_H) + v_H(v_{H-2} - 4v_{H-1} + 6v_H), \\
-v^\intercal D_{22} v = \ & v_1(2v_1 - 2v_2 + v_3) + v_2(-2v_1 + 2v_2 - 2v_3 + v_4) \\
& + \sum_{i=3}^{H-2} v_i(v_{i-2} - 2v_{i-1} + 2v_i - 2v_{i+1} + v_{i+2}) \\
& + v_{H-1}(v_{H-3} - 2v_{H-2} + 2v_{H-1} - 2v_H) + v_H(v_{H-2} + 2v_{H-1} - 2v_H). \\
v^\intercal \tilde{D}_{22} v = \ & v_1 \{-7v_1 + 6v_2 - 2v_3\} + v_2 \left\{6v_1 - 10v_2 + 8v_3 - \frac{5}{2}v_4\right\} \\
& + \sum_{h=3}^{H-2} v_h \Bigg\{ -\frac{h+1}{2} v_{h-2} + 2(h+1)v_{h-1} - (3h+4)v_h + 2(h+2)v_{h+1} \\
& \qquad - \frac{h+3}{2} v_{h+2} \Bigg\} \\
& + v_{H-1} \left\{ -\frac{H}{2} v_{H-3} + 2H v_{H-2} - (3H+1)v_{H-1} + 2(H+1)v_H \right\} \\
& + v_H \left\{ -\frac{H+1}{2} v_{H-2} + 2(H+1)v_{H-1} - (3H+4)v_H \right\}.
\end{aligned}
$$

End-Effects: First we study the end-effects. We have the following table of coefficients

	$k(0)$	$k'(0)\delta$	$k''(0)\delta^2$	$k'''(0)\delta^3$	$k''''(0)\delta^4$
$v_1 = k(\delta)$	1	1	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{24}$
$v_2 = k(2\delta)$	1	2	$\frac{4}{2}$	$\frac{8}{6}$	$\frac{16}{24}$
$v_3 = k(3\delta)$	1	3	$\frac{9}{2}$	$\frac{27}{6}$	$\frac{81}{24}$
$v_4 = k(4\delta)$	1	4	$\frac{16}{2}$	$\frac{32}{6}$	$\frac{128}{24}$

from which it follows that

$$
\begin{aligned}
2v_1 - v_2 &= k(0) - k''(0)\delta^2 - k'''(0)\delta^3 + O(\delta^4) = 1 + O(H^{-2}) \\
2v_1 - 2v_2 + v_3 &= k(0) + k'(0)\delta + \tfrac{3}{2}k''(0)\delta^2 + O(\delta^3) \\
&= 1 + k'(0)H^{-1} + O(H^{-2}) \\
-2v_1 + 2v_2 - 2v_3 + v_4 &= -1 + O(H^{-2}) \\
6v_1 - 4v_2 + v_3 &= 3 + k'(0)H^{-1} - \frac{1}{2}k''(0)H^{-2} + \frac{1}{6}k'''(0)H^{-3} + O(H^{-4}) \\
-4v_1 + 6v_2 - 4v_3 + v_4 &= -1 + O(H^{-4}), \\
-7v_1 + 6v_2 - 2v_3 &= -3k(0) - k'(0)\delta - \frac{1}{2}k''(0)\delta^2 + O(\delta^3), \\
6v_1 - 10v_2 + 8v_3 - \frac{5}{2}v_4 &= 3/2k(0) - k''(0)\delta^2 + O(\delta^3).
\end{aligned}
$$

Similarly,

	$k(1)$	$k'(1)\delta$	$k''(1)\delta^2$	$k'''(1)\delta^3$	$k''''(1)\delta^4$
v_H	1	0	0	0	0
$v_{H-1} = k(1-\delta)$	1	-1	$\frac{1}{2}$	$-\frac{1}{6}$	$\frac{1}{24}$
$v_{H-2} = k(1-2\delta)$	1	-2	$\frac{4}{2}$	$-\frac{8}{6}$	$\frac{16}{24}$
$v_{H-3} = k(3\delta)$	1	-3	$\frac{9}{2}$	$-\frac{27}{6}$	$\frac{81}{24}$

such that $(k(1) = 0)$

$$\begin{aligned}
v_{H-1} &= 7 - k'(1)\delta + \tfrac{1}{2}k''(1)\delta^2 - \tfrac{1}{6}k'''(1)\delta^3 + O(\delta^4) \\
v_{H-3} - 4v_{H-2} + 6v_{H-1} &= -k'(1)\delta - \tfrac{1}{2}k''(1)\delta^2 - \tfrac{1}{6}k'''(1)\delta^3 + O(\delta^4) \\
v_{H-3} - 2v_{H-2} + 2v_{H-1} &= -k'(1)\delta - \tfrac{3}{2}k''(1)\delta^2 + O(\delta^3) \\
-\tfrac{1}{2}v_{H-3} + 2v_{H-2} - 3v_{H-1} &= \tfrac{1}{2}k'(1)\delta + \tfrac{1}{4}k''(1)\delta^2 + O(\delta^3)
\end{aligned}$$

and

$$\begin{aligned}
v_{H-1}(v_{H-3} - 2v_{H-2} + 2v_{H-1}) &= \delta^2 k'(1)^2 + \delta^3 k'(1)k''(1) + O(\delta^4) \\
v_{H-1}(v_{H-3} - 4v_{H-2} + 6v_{H-1}) &= -\delta k'(1)\left\{-k'(1)\delta - \tfrac{1}{2}k''(1)\delta^2 + O(\delta^3)\right\} \\
&\quad + \tfrac{1}{2}k''(1)\delta^2\left\{-k'(1)\delta + O(\delta^2)\right\} \\
&= k'(1)^2\delta^2 + O(\delta^4)
\end{aligned}$$

B **End-Effects** $v_1(2v_1 - v_2) = 1 + k'(0)\delta + O(\delta^2)$.

C **End-Effects**

$$\begin{aligned}
v_1(6v_1 - 4v_2 + v_3) &= 1\left\{3 + k'(0)\delta - \tfrac{1}{2}k''(0)\delta^2 + \tfrac{1}{6}k'''(0)\delta^3 + O(\delta^4)\right\} \\
&\quad + k'(0)\delta\left\{3 + k'(0)\delta - \tfrac{1}{2}k''(0)\delta^2 + O(\delta^3)\right\} \\
&\quad + \tfrac{1}{2}k''(0)\delta^2\left\{3 + k'(0)\delta + O(\delta^2)\right\} + \tfrac{1}{6}k'''(0)\delta^3\left\{3 + O(\delta)\right\} \\
&= 3 + 4k'(0)\delta + \left\{k'(0)^2 + k''(0)\right\}\delta^2 + \tfrac{2}{3}k'''(0)\delta^3 + O(\delta^4),
\end{aligned}$$

and

$$\begin{aligned}
v_2(-4v_1 + 6v_2 - 4v_3 + v_4) &= \left\{1 + 2k'(0)\delta + \tfrac{4}{2}k''(0)\delta^2 + \tfrac{8}{6}k'''(0)\delta^3 + O(\delta^4)\right\} \\
&\quad \times\left\{-1 + O(\delta^4)\right\} \\
&= -1 - 2k'(0)\delta - 2k''(0)\delta^2 - \tfrac{4}{3}k'''(0)\delta^3 + O(\delta^4),
\end{aligned}$$

which add up to

$$\begin{aligned}
&v_1(6v_1 - 4v_2 + v_3) + v_2(-4v_1 + 6v_2 - 4v_3 + v_4) \\
&\quad = 2 + 2k'(0)\delta + \left\{k'(0)^2 - k''(0)\right\}\delta^2 - \frac{4}{6}k'''(0)\delta^3 + O(\delta^4).
\end{aligned}$$

The last end-term is $v_{H-1}(v_{H-3} - 4v_{H-2} + 6v_{H-1} - 4v_H) = k'(1)^2\delta^2 + O(\delta^4)$.

D **End-Effects**

$$\begin{aligned} v_1(2v_1 - 2v_2 + v_3) + v_2(-2v_1 + 2v_2 - 2v_3 + v_4) &= 1 + 2k'(0)\delta \\ &\quad - 1 - 2k'(0)\delta + O(\delta^2) = O(\delta^2) \\ v_{H-1}(v_{H-3} - 2v_{H-2} + 2v_{H-1} - 2v_H) &= k'(1)^2\delta^2 + O(\delta^4). \end{aligned}$$

$\tilde{D}$ **End-Effects:**

$$\begin{aligned} v_1\left\{-7v_1 + 6v_2 - 2v_3\right\} &= -\left\{k(0) + k'(0)\delta + O(\delta^2)\right\}\left\{3k(0) + k'(0)\delta + O(\delta^2)\right\} \\ &= -3 - 4k'(0)\delta + O(\delta^2), \end{aligned}$$

$$\begin{aligned} v_2\left\{6v_1 - 10v_2 + 8v_3 - \tfrac{5}{2}v_4\right\} = &\left\{k(0) + 2k'(0)\delta + O(\delta^2)\right\} \\ &\times\left\{\tfrac{3}{2}k(0) - k''(0)\delta^2 + O(\delta^3)\right\} \\ = &\ \tfrac{3}{2} + 3k'(0)\delta + O(\delta^2), \end{aligned}$$

$$\begin{aligned} &v_{H-1}\left\{-\tfrac{H}{2}v_{H-3} + 2Hv_{H-2} - (3H+1)v_{H-1}\right\} \\ &\quad = v_{H-1}\delta^{-1}\left\{-\tfrac{1}{2}v_{H-3} + 2v_{H-2} - 3v_{H-1}\right\} - v_{H-1}^2, \end{aligned}$$

$$\begin{aligned} v_{H-1}\delta^{-1}\left\{-\tfrac{1}{2}v_{H-3} + 2v_{H-2} - 3v_{H-1}\right\} = &\left(-k'(1)\delta + \tfrac{1}{2}k''(1)\delta^2 + O(\delta^3)\right) \\ &\times\delta^{-1}\left\{\tfrac{1}{2}k'(1)\delta + \tfrac{1}{4}k''(1)\delta^2 + O(\delta^3)\right\} \\ = &-\tfrac{1}{2}k'(1)^2\delta + 0 \times k'(1)k''(1)\delta^2 + O(\delta^3), \end{aligned}$$

$$-v_{H-1}^2 = -\left\{-k'(1)\delta + O(\delta^2)\right\}^2 = -k'(1)^2\delta^2 + O(\delta^3).$$

So the total contribution from end-effects is $-\frac{3}{2} - \left\{k'(0) + \frac{1}{2}k'(1)^2\right\}\delta + O(\delta^2)$.

Summations: From the following table of coefficients,

	$k(x)$	$k'(x)\delta$	$k''(x)\delta^2$	$k'''(x)\delta^3$	$k''''(x)\delta^4$
$k(x-2\delta)$	1	-2	$\frac{4}{2}$	$-\frac{8}{6}$	$\frac{16}{24}$
$k(x-\delta)$	1	-1	$\frac{1}{2}$	$-\frac{1}{6}$	$\frac{1}{24}$
$k(x)$	1	0	0	0	0
$k(x+\delta)$	1	1	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{24}$
$k(x+2\delta)$	1	2	$\frac{4}{2}$	$\frac{8}{6}$	$\frac{16}{24}$

we find that the terms in the summation expressions are given by

$$\begin{aligned} k(x+\delta) - 2k(x) + k(x-\delta) &= \delta^2 k''(x) + O(\delta^4) \\ k(x+2\delta) - 4k(x+\delta) + 6k(x) - 4k(x-\delta) + k(x-2\delta) &= \delta^4 k''''(x) + O(\delta^6) \\ k(x+2\delta) - 2k(x+\delta) + 2k(x) - 2k(x-\delta) + k(x-2\delta) &= 2\delta^2 k''(x) + O(\delta^4). \end{aligned}$$

Thus

$$\frac{1}{H}v^{\mathsf{T}}A_{22}v = \frac{1}{H}\sum_{i=1}^{H} v_i^2 \to \int_0^1 k^2(x)\mathrm{d}x$$

$$
\begin{aligned}
v^{\intercal}B_{22}v &= 1 + k'(0)H^{-1} - H^{-1}\int_0^1 k(x)k''(x)\mathrm{d}x + O(H^{-2})\\
&= 1 + H^{-1}6/5 + O(H^{-2})\\
v^{\intercal}C_{22}v &= 2 + 2k'(0)H^{-1} + H^{-2}\left\{k'(0)^2 + k'(1)^2 - k''(0)\right\} - H^{-3}\frac{4}{6}k'''(0)\\
&\quad + H^{-3}\int_0^1 k(x)k''''(x)\mathrm{d}x + O(H^{-4})\\
v^{\intercal}D_{22}v &= \; -2H^{-1}\int_0^1 k(x)k''(x)\mathrm{d}x + O(H^{-2}) = H^{-1}12/5 + O(H^{-2}).
\end{aligned}
$$

The term for $\tilde{D}_{22}$ is slightly more involved.

$$
\begin{aligned}
&-\frac{h+1}{2}v_{h-2} + 2(h+1)v_{h-1} - (3h+4)v_h + 2(h+2)v_{h+1} - \frac{h+3}{2}v_{h+2}\\
&\quad = -\frac{h}{2}\left\{v_{h-2} - 4v_{h-1} + 6v_h - 4v_{h+1} + v_{h+2}\right\}\\
&\qquad - \frac{1}{2}\left\{v_{h-2} - 4v_{h-1} + 8v_h - 8v_{h+1} + 3v_{h+2}\right\}\\
&\quad = -\frac{h}{2}\left\{v_{h-2} - 4v_{h-1} + 6v_h - 4v_{h+1} + v_{h+2}\right\}\\
&\qquad - \frac{1}{2}\left\{v_{h-2} - 4v_{h-1} + 6v_h - 4v_{h+1} + v_{h+2}\right\}\\
&\qquad - \left\{v_h - 2v_{h+1} + v_{h+2}\right\}\\
&\quad = -\frac{h}{2}\delta^4 k''''(x) + O(\delta^4) - \delta^2 k''(x) + O(\delta^4).
\end{aligned}
$$

So

$$
\begin{aligned}
&\sum_{h=3}^{H-2} v_h \left\{-\frac{h+1}{2}v_{h-2} + 2(h+1)v_{h-1} - (3h+4)v_h + 2(h+2)v_{h+1} - \frac{h+3}{2}v_{h+2}\right\}\\
&\quad = -\delta\int k(x)k''(x)\mathrm{d}x - \frac{1}{2}\delta^2\int xk(x)k''''(x)\mathrm{d}x + O(\delta^3).
\end{aligned}
$$

Other Terms: Finally we obtain

$$
\begin{aligned}
&\overline{w}^{\intercal}C_{11}\overline{w} = 1, \quad \overline{w}^{\intercal}B_{11}\overline{w} = 1, \quad \overline{w}^{\intercal}D_{11}\overline{w} = \lambda^2/2, \quad \overline{w}^{\intercal}\tilde{D}_{11}\overline{w} = -3/2,\\
&\overline{w}^{\intercal}C_{12}v = -3v_1 + v_2 = -2 + H^{-1}k'(0) + \tfrac{1}{2}H^{-2}k''(0) + \tfrac{5}{6}H^{-3}k'''(0) + O(H^{-4})\\
&\overline{w}^{\intercal}B_{12}v = -v_1 = -1 - H^{-1}k'(0) + O(H^{-2})\\
&\overline{w}^{\intercal}D_{12}v = v_1 - v_2 = -H^{-1}k'(0) + O(H^{-2}).\\
&\overline{w}^{\intercal}\tilde{D}_{12}v = 3v_1 - 3/2v_2 = 3/2 - 3/2k''(0)\delta^2 + O(\delta^3).
\end{aligned}
$$

Combining the results yields the formula given in the Theorem.

Proof of Proposition 2. Ignore the notation δ. Now

$$
\sum_{i=1}^{n}(U_i - U_{i-1})(U_{i-h} - U_{i-h-1}) = 2\sum_{i=1}^{n}U_iU_{i-h} - \sum_{i=1}^{n}U_iU_{i-h-1} - \sum_{i=1}^{n}U_iU_{i-h+1}
$$

$$+ U_0(U_{-h} - U_{-h+1}) - U_n(U_{n-h} - U_{n-h+1}).$$

Hence

$$\begin{aligned}
\sum_{h=-H}^{H} k\left(\tfrac{h}{H}\right)\gamma_h(U_\delta) = & \sum_{h=-H}^{H} k\left(\tfrac{h}{H}\right)\left\{2\bar{\gamma}_h(U_\delta) - \bar{\gamma}_{h-1}(U_\delta) - \bar{\gamma}_{h+1}(U_\delta)\right\} \\
& + U_0 \sum_{h=-H}^{H} k\left(\tfrac{h}{H}\right)(U_{-h} - U_{-h+1}) \\
& - U_n \sum_{h=-H}^{H} k\left(\tfrac{h}{H}\right)(U_{n-h} - U_{n-h+1}).
\end{aligned}$$

Next we show that the last two terms are $O_p(H^{-1/2})$. We have

$$\begin{aligned}
\sum_{h=-H}^{H} k\left(\tfrac{h}{H}\right)(U_{-h} - U_{-h+1}) &= \sum_{h=-H}^{H} k\left(\tfrac{h}{H}\right)U_{-h} - \sum_{h=-H+1}^{H+1} k\left(\tfrac{h-1}{H}\right)U_{-h} \\
&= \sum_{h=-H+1}^{H} \left\{k\left(\tfrac{h}{H}\right) - k\left(\tfrac{h-1}{H}\right)\right\}U_{-h},
\end{aligned}$$

since $k(-1)\,U_H - k(1)\,U_{-H-1} = 0$ by the properties of $k(x)$. Defining $k(x) = 0$ for $|x| \geq 1$ allows us to sum from $-H$, i.e. $\sum_{h=-H}^{H}\left\{k\left(\frac{h}{H}\right) - k\left(\frac{h-1}{H}\right)\right\}U_{-h}$. So

$$\sum_{h=-H}^{H} k\left(\tfrac{h}{H}\right)U_0(U_{-h} - U_{-h+1}) = H^{-1/2} \times U_0 \times \left\{H^{-1/2}\sum_{h=-H}^{H} \frac{k\left(\frac{h}{H}\right) - k\left(\frac{h-1}{H}\right)}{1/H}U_{-h}\right\},$$

which is $O_p(H^{-1/2})$ since $a_{h,H} = H^{1/2}\left\{k\left(\frac{h}{H}\right) - k\left(\frac{h-1}{H}\right)\right\}$ is such that

$$\sum_{h=-H}^{H} a_{h,H}^2 = H^{-1}\sum_{h=-H}^{H}\left\{\frac{k\left(\frac{h}{H}\right) - k\left(\frac{h-1}{H}\right)}{1/H}\right\}^2 \to \int_{-1}^{1}\{k'(x)\}^2\,\mathrm{d}x.$$

The proof for the other term is very similar.

Now we move to showing (24). We have

$$\begin{aligned}
\sum_{h=-H}^{H} k\left(\frac{h}{H}\right)\gamma_h = & \sum_{h=-H}^{H} k\left(\frac{h}{H}\right)\left(2\overline{\gamma}_h - \overline{\gamma}_{h-1} - \overline{\gamma}_{h+1}\right) + O_p(H^{-1/2}) \\
= & \sum_{h=-H}^{H}\overline{\gamma}_h\left\{2k\left(\frac{h}{H}\right) - k\left(\frac{(h-1)}{H}\right) - k\left(\frac{(h+1)}{H}\right)\right\} \\
& + O_p(H^{-1/2}) \\
= & -\frac{n}{H^2}\sum_{h=-H}^{H} k''\left(\frac{h}{H}\right)\left(\frac{1}{n}\overline{\gamma}_h\right) + O_p(nH^{-3}) + O_p(H^{-1/2}).
\end{aligned}$$

□

Proof of Proposition 3. Now we assume $k''(0) = 0$ then

$$-\frac{n}{H^2}\sum_{|h|\leq \overline{H}} k''\left(\frac{h}{H}\right)\left(\frac{1}{n}\overline{\gamma}_h\right) = -\frac{n}{H^3}\sum_{|h|\leq \overline{H}} k'''(0)\,|h|\left(\frac{1}{n}\overline{\gamma}_h\right) + O_p\left(\frac{n}{H^3}\right) = O_p\left(\frac{n}{H^3}\right).$$

This leaves us thinking of

$$\sum_{H\geq |h|>\overline{H}} k''\left(\frac{h}{H}\right)\left(\frac{1}{n}\overline{\gamma}_h\right).$$

From Bartlett (1946) we know that for $k > h$

$$\sqrt{n}\begin{pmatrix} \frac{1}{n}\overline{\gamma}_h \\ \frac{1}{n}\overline{\gamma}_k \end{pmatrix} \xrightarrow{L} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \omega^4 \sum_{j=-\infty}^{\infty}\begin{pmatrix} \rho_j^2 + \rho_{j+h}\rho_{j-h} & \rho_j\rho_{j+(k-h)} + \rho_{j+k}\rho_{j-h} \\ \rho_j\rho_{j+(k-h)} + \rho_{j+k}\rho_{j-h} & \rho_j^2 + \rho_{j+k}\rho_{j-k} \end{pmatrix}\right)$$

where ρ_j denotes the population autocorrelation. In the AR(1) case, with persistence parameter $|\rho| < 1$ then it is well known that this simplifies to

$$\sqrt{n}\begin{pmatrix} \frac{1}{n}\overline{\gamma}_h \\ \frac{1}{n}\overline{\gamma}_k \end{pmatrix} \xrightarrow{L} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 2\omega^4 \begin{pmatrix} 1 & \rho^{k-h} \\ \rho^{k-h} & 1 \end{pmatrix}\frac{1+\rho^2}{1-\rho^2}\right),$$

noting $\sum_{j=-\infty}^{\infty}\phi^{2j} = (1+\rho^2)/(1-\rho^2)$.The impact of the serial dependence is that

$$\sqrt{\frac{n}{H}}\left(\sum_{H\geq h>\overline{H}} 2k''\left(\frac{h}{H}\right)\left(\frac{1}{n}\overline{\gamma}_h\right)\right) \xrightarrow{L} N\left\{0, 4\omega^4\frac{1+\rho^2}{1-\rho^2}\int_0^1 k''(x)^2 \mathrm{d}x\right\}.$$

This implies

$$-\frac{n}{H^2}\sum_{|h|\leq \overline{H}} k''\left(\frac{h}{H}\right)\left(\frac{1}{n}\overline{\gamma}_h\right) = O_p\left(\frac{n^{1/2}}{H^{3/2}}\right).$$

Overall we have $O_p\left(n^{1/2}H^{-3/2}\right) + O_p(nH^{-3}) + O_p\left(H^{-1/2}\right)$. Placing $H = cn^{1/2}$ delivers a term which is $O_p\left(n^{-1/4}\right)$. □

Appendix B: Tables with simulation results

Table B.1: Summary Statistics for infeasible T_{raw} where ϖ is known.

ω^2 = 0.01, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	13.3	0.001	1.035	0.06	1.13	3.50	93.35	95.86	98.50
390	18.6	−0.001	1.057	0.14	1.69	4.27	93.26	95.79	98.41
780	26.0	−0.001	1.077	0.25	2.09	4.84	93.16	95.77	98.43
1560	36.6	−0.002	1.103	0.41	2.56	5.43	92.99	95.67	98.38
4680	63.1	−0.000	1.163	0.68	3.19	6.28	92.59	95.35	98.10
5850	70.5	−0.003	1.180	0.76	3.31	6.44	92.47	95.21	98.04
7800	81.3	−0.001	1.204	0.87	3.58	6.69	92.37	95.11	97.88
11700	99.5	−0.004	1.242	1.01	3.92	7.19	92.12	94.83	97.68
23400	140.4	−0.003	1.329	1.41	4.67	8.05	91.44	94.26	97.23

ω^2 = 0.001, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	4.54	−0.001	0.999	0.12	1.40	3.67	94.12	96.61	98.94
390	6.21	−0.000	1.001	0.18	1.70	4.05	94.28	96.75	99.09
780	8.58	0.000	0.999	0.22	1.87	4.27	94.42	96.93	99.19
1560	11.9	0.000	1.003	0.26	2.04	4.49	94.45	96.99	99.23
4680	20.3	−0.000	1.007	0.36	2.21	4.71	94.55	97.01	99.27
5850	22.6	0.001	1.007	0.35	2.24	4.75	94.52	97.04	99.29
7800	26.0	−0.001	1.010	0.33	2.28	4.84	94.51	97.06	99.24
11700	31.8	−0.000	1.009	0.37	2.38	4.85	94.57	97.12	99.31
23400	44.8	−0.001	1.017	0.43	2.37	5.01	94.46	97.06	99.29

ω^2 = 0.0001, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	1.75	−0.000	0.998	0.17	1.63	3.97	94.32	96.75	99.07
390	2.30	−0.000	0.999	0.25	1.86	4.22	94.47	96.95	99.20
780	3.06	0.000	0.998	0.28	1.98	4.49	94.61	97.10	99.26
1560	4.11	−0.001	1.000	0.31	2.12	4.59	94.65	97.11	99.30
4680	6.76	−0.001	1.002	0.39	2.23	4.79	94.71	97.24	99.36
5850	7.50	0.001	1.000	0.40	2.22	4.74	94.72	97.30	99.41
7800	8.58	−0.001	1.005	0.40	2.34	4.84	94.69	97.21	99.37
11700	10.4	−0.002	1.001	0.42	2.30	4.81	94.81	97.29	99.36
23400	14.5	−0.003	1.004	0.46	2.45	4.92	94.78	97.28	99.43

Summary Statistics for the infeasible T_{raw} under stochastic volatility. The first column defines the sampling frequency. The second gives the average values of H^*. Columns 3 and 5 present the mean and the standard deviation of the statistics. The remaining 6 columns give the simulated quantiles of distribution.

Table B.2: Summary Statistics for feasible T_{raw}, estimating ϖ from the data.

ω^2 = 0.01, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	8.32	−0.220	1.094	3.12	6.89	10.1	98.50	99.64	99.99
390	11.6	−0.188	1.066	2.53	6.05	9.12	97.90	99.38	99.98
780	16.2	−0.155	1.047	2.01	5.29	8.36	97.39	99.11	99.94
1560	22.6	−0.131	1.044	1.78	4.94	8.02	96.83	98.75	99.88
4680	38.8	−0.102	1.059	1.47	4.63	7.71	95.89	98.08	99.68
5850	43.4	−0.099	1.059	1.38	4.53	7.64	95.77	97.96	99.61
7800	50.0	−0.094	1.068	1.40	4.49	7.63	95.45	97.76	99.54
11700	61.1	−0.087	1.083	1.38	4.63	7.85	95.12	97.42	99.37
23400	86.2	−0.079	1.116	1.44	4.71	8.05	94.26	96.83	99.03

ω^2 = 0.001, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	4.31	−0.232	1.065	2.88	6.45	9.59	98.67	99.69	99.99
390	5.79	−0.177	1.027	2.11	5.25	8.24	98.08	99.40	99.97
780	7.95	−0.138	1.003	1.60	4.52	7.49	97.64	99.18	99.95
1560	11.0	−0.112	0.995	1.29	4.08	6.89	97.19	98.93	99.91
4680	18.7	−0.083	0.992	0.97	3.54	6.29	96.58	98.56	99.83
5850	20.8	−0.077	0.991	0.94	3.47	6.24	96.52	98.53	99.82
7800	23.9	−0.073	0.994	0.90	3.41	6.24	96.36	98.44	99.78
11700	29.2	−0.065	0.993	0.83	3.34	6.07	96.27	98.38	99.77
23400	41.1	−0.056	1.001	0.79	3.20	5.92	95.93	98.14	99.70

ω^2 = 0.0001, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	2.78	−0.210	1.031	2.54	5.95	9.00	99.19	99.88	100.00
390	2.77	−0.160	0.994	1.70	4.68	7.62	98.34	99.55	99.99
780	3.23	−0.118	0.975	1.22	3.87	6.69	97.62	99.14	99.93
1560	4.20	−0.086	0.974	0.96	3.44	6.14	96.94	98.79	99.87
4680	6.86	−0.059	0.977	0.74	3.01	5.63	96.45	98.48	99.78
5850	7.61	−0.052	0.976	0.71	2.91	5.54	96.35	98.42	99.76
7800	8.71	−0.049	0.982	0.72	2.93	5.55	96.18	98.28	99.72
11700	10.5	−0.045	0.979	0.67	2.81	5.35	96.16	98.22	99.72
23400	14.7	−0.037	0.985	0.66	2.81	5.38	95.87	98.08	99.71

Summary Statistics for feasible T_{raw} under stochastic volatility. The first column defines the sampling frequency. The second gives the average values of H^*. Columns 3 and 5 present the mean and the standard deviation of the statistics. The remaining 6 columns give the simulated quantiles of distribution.

Table B.3: Summary Statistics for feasible $T_{\log}$, estimating ϖ from the data.

ω^2 = 0.01, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	8.32	−0.094	1.003	0.84	3.38	6.35	96.10	98.17	99.67
390	11.6	−0.085	1.007	0.80	3.24	6.11	95.76	97.91	99.61
780	16.2	−0.070	1.008	0.70	3.02	5.86	95.58	97.73	99.54
1560	22.6	−0.060	1.018	0.64	3.03	5.87	95.19	97.54	99.46
4680	38.8	−0.046	1.046	0.71	3.14	6.07	94.55	97.05	99.23
5850	43.4	−0.046	1.049	0.66	3.12	6.11	94.50	96.98	99.14
7800	50.0	−0.043	1.061	0.68	3.25	6.16	94.33	96.79	99.07
11700	61.1	−0.041	1.078	0.75	3.39	6.46	94.08	96.54	98.85
23400	86.2	−0.037	1.117	0.83	3.67	6.87	93.32	96.01	98.52

ω^2 = 0.001, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	4.31	−0.153	1.004	1.49	4.36	7.33	97.31	98.89	99.87
390	5.79	−0.118	0.992	1.16	3.76	6.54	96.90	98.68	99.81
780	7.95	−0.093	0.982	0.92	3.34	6.12	96.67	98.54	99.78
1560	11.0	−0.076	0.983	0.77	3.13	5.81	96.37	98.36	99.75
4680	18.7	−0.057	0.986	0.68	2.90	5.53	95.98	98.08	99.67
5850	20.8	−0.053	0.985	0.65	2.91	5.49	95.90	98.07	99.67
7800	23.9	−0.051	0.990	0.63	2.86	5.55	95.81	98.01	99.63
11700	29.2	−0.045	0.990	0.59	2.85	5.47	95.76	98.00	99.62
23400	41.1	−0.039	0.999	0.61	2.79	5.43	95.48	97.78	99.59

ω^2 = 0.0001, number of reps. = 150000

No. obs	$\overline{H}^*_{simple}$	Mean	Stdv.	0.5%	2.5%	5%	95%	97.5%	99.5%
195	2.78	−0.142	0.981	1.38	4.20	7.05	98.01	99.35	99.97
390	2.77	−0.114	0.970	1.01	3.51	6.23	97.33	99.01	99.91
780	3.23	−0.085	0.963	0.78	3.05	5.67	96.85	98.67	99.82
1560	4.20	−0.061	0.967	0.64	2.86	5.36	96.36	98.36	99.74
4680	6.86	−0.043	0.974	0.57	2.58	5.16	96.03	98.17	99.68
5850	7.61	−0.037	0.973	0.55	2.56	5.06	95.96	98.12	99.67
7800	8.71	−0.035	0.980	0.56	2.61	5.12	95.82	97.99	99.63
11700	10.5	−0.033	0.978	0.54	2.53	5.02	95.84	97.98	99.62
23400	14.7	−0.027	0.984	0.55	2.59	5.08	95.61	97.86	99.64

Summary Statistics for feasible $T_{\log}$ under stochastic volatility, using $d = 0.2$. The first column defines the sampling frequency. The second gives the average values of H^*. Columns 3 and 5 present the mean and the standard deviation of the statistics. The remaining 6 columns give the simulated quantiles of distribution.

	Mean	Std. (HAC)	$\rho(\widehat{[Y]}, \widetilde{K})$	acf(1)	acf(2)	acf(5)	acf(10)
Modified Tukey-Hanning kernel ($H = cn^{1/2}$)							
$\widetilde{K}_w^{\text{TH2}}(X_{\text{ap. 1 min}})$	0.962	0.568 (1.195)	1.000	0.34	0.32	0.28	0.08
Parzen kernel ($H = cn^{1/2}$)							
$\widetilde{K}_w^{\text{PAR}}(X_{\text{ap. 1 min}})$	0.962	0.570 (1.197)	1.000	0.34	0.32	0.27	0.08
Cubic kernel ($H = cn^{1/2}$)							
$\widetilde{K}_w^{\text{CUB}}(X_{\text{ap. 1 min}})$	0.959	0.568 (1.192)	1.000	0.34	0.32	0.27	0.08
5th order kernel ($H = cn^{1/2}$)							
$\widetilde{K}_w^{\text{5TH}}(X_{\text{ap. 1 min}})$	0.971	0.558 (1.186)	0.999	0.35	0.32	0.28	0.08
8th order kernel ($H = cn^{1/2}$)							
$\widetilde{K}_w^{\text{8TH}}(X_{\text{ap. 1 min}})$	0.965	0.578 (1.212)	0.995	0.34	0.32	0.27	0.09
Top-Flat Bartlett kernel ($H = cn^{2/3}$)							
$\widetilde{K}_w^{\text{BART}}(X_{\text{ap. 1 min}})$	0.963	0.562 (1.184)	0.997	0.34	0.31	0.27	0.07
Simple RV							
$[X_{20 \text{ minutes}}]$	0.879	0.524 (1.008)	0.832	0.28	0.24	0.26	0.06
$[X_{5 \text{ minutes}}]$	0.948	0.518 (1.100)	0.954	0.36	0.34	0.26	0.10
$[X_{1 \text{ minutes}}]$	0.941	0.382 (0.919)	0.887	0.44	0.40	0.38	0.11
$[X_{10 \text{ seconds}}]$	1.330	0.389 (1.142)	0.803	0.60	0.56	0.51	0.32
$[X_{1 \text{ second}}]$	2.183	0.569 (1.828)	0.733	0.69	0.66	0.57	0.48

Table 6: *Summary statistics for six realised kernels based on returns measured every K trades, where K is selected such that over the day returns on average roughly spans 60 seconds. Also given are the RV, computed using 20, 10, 1 minute, 10 and 1 second returns. Note that RV statistics based on 1 second returns is the same as RV statistics based on all trades. acf denotes serial correlation. The correlation is between the various RV statistics and our Modified Tukey-Hanning kernel.*

	Mean	Std. (HAC)	$\rho(\widehat{[Y]}, \widetilde{K})$	acf(1)	acf(2)	acf(5)	acf(10)
Modified Tukey-Hanning kernel ($H = cn^{1/2}$)							
$\widetilde{K}_w^{\text{TH2}}(X_{\text{ap. 1 min}})$	0.962	0.568 (1.195)	1.000	0.34	0.32	0.28	0.08
Modified Tukey-Hanning kernel (inefficient rate $H = cn^{2/3}$*)*							
$\widetilde{K}_w^{\text{TH2}}(X_{\text{1 tick}})$	0.945	0.521 (1.127)	0.990	0.37	0.31	0.30	0.08
Parzen kernel (inefficient rate $H = cn^{2/3}$*)*							
$\widetilde{K}_w^{\text{PAR}}(X_{\text{1 tick}})$	0.947	0.524 (1.133)	0.990	0.37	0.31	0.30	0.08
Cubic kernel (inefficient rate $H = cn^{2/3}$*)*							
$\widetilde{K}_w^{\text{CUB}}(X_{\text{1 tick}})$	0.948	0.528 (1.142)	0.991	0.37	0.32	0.30	0.08
5th order kernel (inefficient rate $H = cn^{2/3}$*)*							
$\widetilde{K}_w^{\text{5TH}}(X_{\text{1 tick}})$	0.951	0.531 (1.148)	0.989	0.37	0.31	0.30	0.08
8th order kernel (inefficient rate $H = cn^{2/3}$*)*							
$\widetilde{K}_w^{\text{8TH}}(X_{\text{1 tick}})$	0.954	0.573 (1.207)	0.998	0.34	0.31	0.27	0.09
Simple RV subsampled							
$[X_{\text{20 minutes}}; 1200]$	0.885	0.516 (1.036)	0.933	0.27	0.27	0.27	0.08
$[X_{\text{5 minutes}}; 300]$	0.943	0.503 (1.088)	0.984	0.37	0.32	0.30	0.08
$[X_{\text{1 minutes}}; 60]$	0.942	0.376 (0.921)	0.899	0.46	0.43	0.38	0.12
ZMA (2005)							
TSRV$(K, 1)$	0.544	0.321 (0.711)	0.842	0.40	0.34	0.29	0.05
TSRV$(K, 1)$ - *adj*	0.596	0.353 (0.784)	0.854	0.40	0.34	0.29	0.04
AMZ (2006)							
TSRV(K, J)	0.736	0.436 (0.929)	0.944	0.33	0.35	0.28	0.11
TSRV(K, J) - *aa*	0.946	0.560 (1.194)	0.944	0.33	0.35	0.28	0.11

Table 7: *Summary statistics: First the realised Modified Tukey-Hanning kernel using approximate 1 minute returns. Then, five realised kernels which also appear in Table 6, but now they are computed using the inefficient rate and based on all available trades. Next, subsampled versions of simple RV statistics based on 20, 5 and 1 minute returns. For instance, the subsampled* $[X_{5\ minutes}; 300]$ *calculates RV over 5 minutes, averaged over 300 times, just changing the initial place prices are recorded. Under ZMA (2005) the two scale RV estimators suggested in Zhang, Mykland, and Aït-Sahalia (2005, eq. (55) and (64)) are listed. The AMZ (2006) are two scale estimators (see Aït-Sahalia, Mykland, and Zhang (2006, eq. (4.4) and (4.22))) designed to be robust to deviations from i.i.d. noise. These authors also proposed the second estimator (their equations (64) and (4.22)), which scales the first estimator, to overcome the finite sample bias of TSRV.*