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

Ole E. Barndorff-Nielsen  
_Aarhus University_

Peter Reinhard Hansen  
_Stanford University_

Asger Lunde  
_Aarhus School of Business_

Neil Shephard  
_Nuffield College, Oxford University_

1. Introduction

- In the stochastic volatility framework:
  \[ Y_t = \int_0^t a_s \, ds + \int_0^t \sigma_s \, dW_s, \]
  where \( a \) is predictable, \( \sigma \) is cadlag, and \( W \) is a standard Brownian Motion.

- Object of interest
  \[ IV \equiv \int_0^1 \sigma_s^2 \, ds, \quad \text{(Integrated Variance)}. \]
Prices observed at times: \( t = \frac{i}{n}, \quad i = 0, \ldots, n. \)

Intraday returns: \( y_i \equiv Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}, \quad \text{for } i = 1, \ldots, n. \)

Realized Variance: \( RV^{(n)} \equiv \sum_{i=1}^{n} y_i^2, \quad \text{(at frequency } n). \)
2. REALIZED VARIANCE: THE CASE WITHOUT NOISE

Consistency:

\[ RV^{(n)} \xrightarrow{p} IV. \]

- Andersen & Bollerslev (1998),
- Barndorff-Nielsen & Shephard (2002c),
- Meddahi (2002), (detailed analysis of relation between \( RV^{(n)} \) and \( IV \)).

CLT

\[ n^{1/2}(RV^{(n)} - IV) \xrightarrow{d} N(0, 2IQ), \quad \text{where} \quad IQ \equiv \int_{0}^{1} \sigma^4(s)ds. \]

- Barndorff-Nielsen & Shephard (2002a),
But there is noise!

- Suppose

\[ X_t = Y_t + U_t, \]

- \( X_t \) is the observed process.
- \( Y_t \) is the latent BSM.
- \( U_t \) is the noise/measurement error.

- \( U_t \) (Market Microstructure Noise) due to
  - Rounding error
  - Temporary demand-supply imbalance (large trades).
  - Data errors (zeros, comma, time-stamp)
Illustration.
$RV^{(n)}$ is biased and inconsistent for $[Y, Y] = IV$. 

Realized variance for Microsoft  (x-axis in logs).

Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.6)
Assumption 1 \( U \) is a zero mean, weak white noise process with

\[
\text{var}(U_t) = \omega^2, \quad \text{var}(U_t^2) = \lambda^2 \omega^4, \quad U_s \perp U_t \quad (s \neq t).
\]

- Here:

\[
E[RV^{(n)}] = E\left[ \sum_{i=1}^{n} y_i^2 \right] + 2n\omega^2
\]

- First-order Autocovariance: \( E[x_i x_{i-1}] = E[u_i u_{i-1}] = -\omega^2 \).

- So the estimator by Zhou (1996) is unbiased

\[
RV_{AC_1}^{(n)} \equiv \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} x_i x_{i-1} + \sum_{i=1}^{n} x_i x_{i+1},
\]

but unfortunately this estimator is inconsistent.

*Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.7)*
Sample Sparsely:

Pre-Whitening... and sample sparsely:

Kernel Estimators:

Subsampling: Sample sparsely many times and take average
What do we know about the Noise?

  - “No noise” results reliable at low sampling frequencies (20 min.)
  - “Independent noise” results ... moderate frequencies (1-2 min.)
  - “General noise”... we characterize the RV’s bias
  - Kernel-based estimators are useful for bias correction/study noise.

- Empirical Properties of the noise
  - Empirical evidence that noise is time-dependent.
  - Negative correlation between efficient returns and noise.
  - Properties of the Noise have changed.
  - Noise is “small” in recent years.

Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.9)
3. Contributions of this Paper

- **Kernel Estimators in Details:**
  - Derive **bias** and **variance** properties for all kernel estimators.

- **Relation to the consistent subsample estimator** of ZMA.
  - Asymmetric kernels can be made **consistent** by subsampling.

- **Symmetric Kernel Estimators:**
  - Consistent and converges at rate $n^{1/4}$.

- Estimation of $IQ$ and $\omega^2$.

- Effect of **stochastically spaced data**.

- **Simulations and Empirical Illustration**.

*Barndorf-Nielsen, Hansen, Lunde & Shephard, 2005 (p.10)*
4. **Realized Kernels**

For an arbitrary process, $Z$, and a small $\delta > 0$ write the HF returns

$$z_{j,\varepsilon} = Z(j+\varepsilon)\delta - Z(j+\varepsilon-1)\delta,$$

where $\varepsilon$ an **offset** introduced to deal with subsampling.
The key quantities are the realised autocovariation processes:

\[
\gamma_h(Z_\delta, X_\delta : \varepsilon)_t = \sum_{j=1}^{\lfloor t/\delta \rfloor} z_{j,\varepsilon} x_{j-h,\varepsilon}, \quad h = -H, \ldots, -1, 0, 1, 2, \ldots, H,
\]

and their averaged version

\[
\gamma_h(Z_\delta, X_\delta; S)_t = \frac{1}{S+1} \sum_{s=0}^{S} \gamma_h(Z_\delta, X_\delta : s/(S+1))_t.
\]

We call \(\gamma_h(X_\delta; S) = \gamma_h(X_\delta, X_\delta; S)\) the realised autocovariation process, and note that

\[
\gamma_h(X_\delta; S) = \gamma_h(Y_\delta; S) + \gamma_h(U_\delta; S) + \gamma_h(Y_\delta, U_\delta; S) + \gamma_h(U_\delta, Y_\delta; S).
\]  (1)

Throughout we write, for \(h > 0\),

\[
\tilde{\gamma}_h(Z_\delta, X_\delta; S) = \gamma_h(Z_\delta, X_\delta; S) + \gamma_{-h}(Z_\delta, X_\delta; S).
\]

When \(S = 0\) we write: \(\gamma_h(X_\delta)_t = \gamma_h(X_\delta; 0)_t\).
5. Defining the realised kernel

A general realised kernel takes the form

$$K_w(X_\delta; S)_t = \sum_{h=-H}^{H} w_h Y_h(X_\delta; S)$$

$$= K_w(Y_\delta; S) + K_w(U_\delta; S) + K_w(Y_\delta, U_\delta; S) + K_w(U_\delta, Y_\delta; S),$$

where the weights

$$w = (w_{-H}, \ldots, w_0, w_1, \ldots, w_H)'$$

are non-stochastic and

$$K_w(Y_\delta, U_\delta; S) = \sum_{h=-H}^{H} w_h Y_h(Y_\delta, U_\delta; S).$$
Important classes of kernels are the asymmetric and symmetric kernels, respectively,

\[
K_w(X_{\delta}) = \sum_{h=0}^{H} w_h \gamma_h(X_{\delta}; S)
\]

\[
\tilde{K}_w(X_{\delta}) = \sum_{h=0}^{H} w_h \tilde{\gamma}_h(X_{\delta}; S)
\]

Throughout we will write

\[
\gamma(X_{\delta}; S) = \{\gamma_0(X_{\delta}; S), 2\gamma_1(X_{\delta}; S), ..., 2\gamma_H(X_{\delta}; S)\}' .
\]

\[
\tilde{\gamma}(X_{\delta}; S) = \{\gamma_0(X_{\delta}; S), \tilde{\gamma}_1(X_{\delta}; S), ..., \tilde{\gamma}_H(X_{\delta}; S)\}' .
\]

\[
\tilde{\gamma}(Y_{\delta}, U_{\delta}; S) = (\gamma_0(Y_{\delta}, U_{\delta}; S), \tilde{\gamma}_1(Y_{\delta}, U_{\delta}; S), ..., \tilde{\gamma}_H(Y_{\delta}, U_{\delta}; S))' .
\]
Note that

\[ \text{var}(K_w(X)) = [\text{No Noise}] + [\text{Pure Noise}] + [\text{Cross Terms}], \]

since

\[ x_ix_{i+h} = [y_iy_{i+h}] + [u_iu_{i+h}] + [u_iy_{i+h} + y_iu_{i+h}]. \]
Proposition 1: Suppose $Y \in \mathcal{BSM}$, $U \in \mathcal{WN}$ and $Y \perp \perp U$, then as $\delta \downarrow 0$ for fixed $S$ and $h > 0$.

$\gamma_h(Y_\delta; S) - \gamma_{-h}(Y_\delta; S) = O_p(\delta), \quad \gamma_h(U_\delta; S) - \gamma_{-h}(U_\delta; S) = O_p(\delta).$

Further

$\gamma_h(U_\delta; S)_t - \gamma_{-h}(U_\delta; S)_t = O_p(1). \quad (2)$

- So, the properties of asymmetric and symmetric kernels differ due to (2).
- Hence, we will need to study the properties of

$\gamma(Y_\delta; S), \quad \gamma(U_\delta; S), \quad \tilde{\gamma}(U_\delta; S) \quad \text{and} \quad \tilde{\gamma}(Y_\delta, U_\delta; S).$
6.1. Core results: No Noise

**Theorem 1.** Suppose that $Y \in BSM$ and that $a$ and $\sigma$ are predictable. Then as $\delta \downarrow 0$ so for fixed $S$

\[
\begin{pmatrix}
[Y_\delta; S]_t - \int_0^t \sigma_u^2 du \\
Y_1(Y_\delta; S) \\
\vdots \\
Y_H(Y_\delta; S)
\end{pmatrix}
\delta^{-1/2} \xrightarrow{L} MN(0, A_S \times IQ)
\]

The convergence is in law, with

\[
A_S = (S + 1)^{-3}
\]

\[
\begin{pmatrix}
A_S^{1,1} & \bullet & \bullet & \cdots & \cdots & \bullet \\
\bullet & A_S^{2,2} & \bullet & \cdots & \cdots & \bullet \\
\bullet & \bullet & A_S^{3,2} & A_S^{2,2} & \cdots & \cdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots \\
\vdots & \vdots & \vdots & \ddots & \bullet & \bullet \\
0 & 0 & \cdots & 0 & A_S^{3,2} & A_S^{2,2} \\
0 & 0 & \cdots & 0 & \bullet & A_S^{2,2}
\end{pmatrix}
\]

Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.17)
Comments: the no noise case

- Key to the asymptotic distribution of the $\gamma_h(Y_\delta)$ is the $A_S$ matrix. Important special cases of this result are

\[
A_0 = \begin{bmatrix}
  2 & \cdots & \cdots \\
  0 & 1 & \cdots & \cdots \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & 1
\end{bmatrix},
\]

\[
A_1 = \begin{bmatrix}
  3/2 & \cdots & \cdots & \cdots & \cdots \\
  1/4 & 3/4 & \cdots & \cdots & \cdots \\
  0 & 1/8 & 3/4 & \cdots & \cdots \\
  \vdots & \vdots & \ddots & \ddots & \vdots \\
  \vdots & \vdots & \vdots & \ddots & \cdots \\
  0 & 0 & \cdots & 0 & 1/8 & 3/4 \\
  0 & 0 & \cdots & 0 & 1/8 & 3/4
\end{bmatrix}, \quad \lim_{S \to \infty} A_S = \begin{bmatrix}
  4/3 & \cdots & \cdots & \cdots & \cdots \\
  1/3 & 2/3 & \cdots & \cdots & \cdots \\
  0 & 1/6 & 2/3 & \cdots & \cdots \\
  \vdots & \vdots & \ddots & \ddots & \vdots \\
  \vdots & \vdots & \vdots & \ddots & \cdots \\
  0 & 0 & \cdots & 0 & 1/6 & 2/3 \\
  0 & 0 & \cdots & 0 & 1/6 & 2/3
\end{bmatrix}.
\]

- Clearly the limiting result is a good approximation even for quite small $S$.  

Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.18)
Special cases:

- The asymptotic distribution

\[ \delta^{-1/2} ([Y_\delta]_t - [Y]_t) \xrightarrow{L} MN \left( 0, 2 \int_0^t \sigma_u^4 du \right) \]  

(3)


- The extension of (3) to the subsampled case

\[ \delta^{-1/2} ([Y_\delta; S]_t - [Y]_t) \xrightarrow{L} MN \left( 0, 2 \vartheta_S \int_0^t \sigma_u^4 du \right), \quad \vartheta_S = \frac{A_S^{1,1}}{2(S + 1)^3}, \]  

(4)

is in Zhang, Mykland & Aït-Sahalia (2005b).

→ Note that \( \vartheta_S \) falls from 1 to 2/3 as \( S \) rises from 0 to infinity.
Some simple kernels:

- The asymptotic distribution of the Zhou (1996b) kernel estimator can be found from Theorem 1 as

  \[
  \delta^{-1/2} \left( [Y_\delta; S]_t + \tilde{\gamma}_1(Y_\delta; S) - [Y]_t \right) \xrightarrow{L} MN \left( 0, \frac{A^{1,1}_s + 4A^{2,1}_s + 4A^{2,2}_s}{(S+1)^2} \int_0^t \sigma_u^4 \, du \right). 
  \]

- The most interesting special cases are

  \[
  \delta^{-1/2} \left( [Y_\delta]_t + \tilde{\gamma}_1(Y_\delta) - [Y]_t \right) \xrightarrow{L} MN \left( 0, 6 \int_0^t \sigma_u^4 \, du \right), \tag{5}
  \]

  \[
  \delta^{-1/2} \left( [Y_\delta; \infty]_t + \tilde{\gamma}_1(Y_\delta; \infty) - [Y]_t \right) \xrightarrow{L} MN \left( 0, \frac{16}{3} \int_0^t \sigma_u^4 \, du \right). \tag{6}
  \]

- So using the entire sample path of \(Y\) in this way reduces the asymptotic variance by only around 10%.
6.2. Core results: Cross Terms

**Theorem 1, Cont.** By conditioning on \( Y \), if \( U \in \mathcal{WN} \) and \( Y \perp U \) then

\[
\tilde{\gamma}(Y_\delta, U_\delta; S) \overset{L}{\to} MN \left( 0, \frac{2\omega^2}{S + 1} [Y]B \right),
\]

where \( B \) is a \((H + 1) \times (H + 1)\) symmetric matrix with block structure

\[
B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},
\]

\[
B_{22} = \begin{pmatrix} 2 & \ddots & \ddots & \ddots \\ -1 & 2 & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ \cdots & \cdots & 0 & -1 & 2 \end{pmatrix}, \quad B_{11} = \begin{pmatrix} 1 \\ \vdots \end{pmatrix}, \quad B_{12} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 \end{pmatrix},
\]

\( B_{21} = B_{12}' \). Here \( B_{22} \) is a \((H - 1) \times (H - 1)\) symmetric matrix.
6.3. Core results: Pure noise

Theorem 1, Cont. Finally, when $U \in \mathcal{WN}$ and writing $n = \lfloor t/\delta \rfloor$

$$E \{\gamma(U_\delta; S)\} = E \{\tilde{\gamma}(U_\delta; S)\} = 2\omega^2 n (1, -1, 0, 0, \ldots, 0)^t,$$

while for $n \geq H$

$$\text{Cov} \{\gamma(U_\delta; S)\} = \frac{4\omega^4}{S + 1} (nC + D) \quad (7)$$

$$\text{Cov} \{\tilde{\gamma}(U_\delta; S)\} = \frac{4\omega^4}{S + 1} \left(nC + \tilde{D}\right).$$

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

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}, \quad D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}, \quad \tilde{D} = \begin{pmatrix} \tilde{D}_{11} & \tilde{D}_{12} \\ \tilde{D}_{21} & \tilde{D}_{22} \end{pmatrix}$$
6.4. Putting it together

Putting it together (Mean)

\[ IVw_0 + 2n\omega^2(w_0 - w_1). \]

Unbiased kernels: \( w = (u', v')' \) where \( u = (1, 1)' \).

Putting it together (Variance)

\[ w' \left( 2IQA \frac{A}{n} + 8\omega^2 IVB + 4\omega^2(nC + D) \right) w, \]

which we seek to minimize with respect to \( w \) (or \( v \)).
We show that

\[ \hat{v}_h = 1 + \frac{2h^3 - h(3H + 2) - 3h^2 H}{H(H + 1)(H + 2)} \]

such that: \[ \hat{v}_{[sH]} \rightarrow 1 + 2s^3 - 3s^2, \quad H \rightarrow \infty. \]
How to choose $H$ (Part I):

**Symmetric Kernel:**

With

$$\hat{v}_h = \frac{(H - h)(H - h + 1)(H + 2h + 2)}{H(H + 1)(H + 2)}$$

we have

$$w' \frac{1}{n} Aw + w' Bw + w' (Cn)w + w' \tilde{D}w,$$

which is $O(n^{-1/2})$ when

$$H \propto n^{1/2}.$$
How to choose $H$ (Part II):

Let $H = c \sqrt{n}$.

$$n^{1/4} \left( \tilde{K}_w(X) - IV \right) \xrightarrow{L} MN(0, \frac{52}{35} IQc + \frac{48}{6} IV \omega^2 c^{-1} + 4 \omega^4 (12c^{-3} + \frac{6}{5} c^{-1}))$$.

So in general the optimal $H$ is given by

$$H^*_{IV, IQ, \omega^2, n} = \sqrt{n} \omega \left( \frac{(2IV + \omega^2) + \sqrt{(2IV + \omega^2)^2 + \frac{260}{7} IQ}}{13IQ/21} \right)^{1/2}$$.
Proposition 6: Suppose $\sigma_i^2 = \sigma^2$ (constant volatility).

$$V(c) = 4\sigma^4 \left\{ \frac{13}{35} c + \frac{12}{6} \frac{\omega^2}{\sigma^2} c^{-1} + \frac{\omega^4}{\sigma^4} \left( 12c^{-3} + \frac{6}{5} c^{-1} \right) \right\}.$$ 

As $\omega^2 / \sigma^2 \to 0$ we have

$$\hat{c} = 3.6867 \frac{\omega}{\sigma} + O \left( \frac{\omega^2}{\sigma^2} \right),$$

while

$$V(\hat{c}) = \sigma^3 \omega \left\{ 9.0387 + O \left( \frac{\omega^2}{\sigma^2} \right) \right\}.$$
How (in)efficient?

Suppose $U_t \sim iid \, N(0, \omega^2)$ and $\sigma_t^2 = \sigma^2$. Then

$$
\begin{pmatrix}
X_1 \\
X_2 \\
\vdots \\
X_n
\end{pmatrix}
\sim N_n(0, \begin{pmatrix}
\frac{\sigma^2}{n} + 2\omega^2 & \cdot & \cdot & \cdot \\
-\omega^2 & \frac{\sigma^2}{n} + 2\omega^2 & \cdot & \cdot \\
0 & -\omega^2 & \frac{\sigma^2}{n} + 2\omega^2 & \cdot \\
\cdot & \cdot & \cdot & \cdot
\end{pmatrix})
$$

and

$$
n^{1/4}(\sigma_{ML}^2 - \sigma^2) \xrightarrow{L} N(0, 8\omega \sigma^3).
$$
MSE plots $\omega^2 = 0.001$ (vertical axis in logs).

Barndorf-Nielsen, Hansen, Lunde & Shephard, 2005 (p.29)
We simulate over the interval in time $[0, 1]$.

We normalize one second to be $1/23,400$, so that the interval $[0, 1]$ contains 6.5 hours.

In generating the observed price, we discretize $[0, 1]$ into a number $N = 23,400$ of intervals.

When assessing the performance of the estimators under sparse sampling, we look at values of $n$ which are factors of $N$. 

Barndorf-Nielsen, Hansen, Lunde & Shephard, 2005 (p.30)
7.1. Stochastic Volatility

We consider the following SV model (see e.g. Huang & Tauchen (2005))

\[ dY_t = \mu dt + \sigma_t dW_t, \]

\[ \sigma_t = \exp(\beta_0 + \beta_1 \tau_t), \quad \text{with} \quad \text{corr}(dW_t, dB_t) = \rho \]

\[ d\tau_t = \alpha \tau_t dt + dB_t. \]

Using the fact that

\[ \tau_t | \tau_0 \sim N \left( e^{-\alpha t} \tau(0), \frac{1}{-2\alpha} (1 - e^{-\alpha t}) \right) \xrightarrow{L} N \left( 0, \frac{1}{-2\alpha} \right), \]

we restart the volatility process each day at \( \tau_0 \sim N(0, (-2\alpha)^{-1}) \).

For the noise we use

\[ U_{j/N} \overset{i.i.d.}{\sim} N(0, \omega^2), \quad j = 0, \ldots, N. \]
7.2. Implementation of realised kernel

- We ignore data at the start and end of the sample and calculate

\[
\gamma_h^H(\delta) = \sum_{j=H+1}^{n-H} (X_{\delta j} - X_{\delta(j-1)})(X_{\delta(j-h)} - X_{\delta(j-h-1)})
\]

for \(h = -H, \ldots, -1, 0, 1, \ldots, H\).

- From this we compute our estimator of \([Y]_1\) as

\[
\tilde{K}_H(\delta) = (1 - H\delta)^{-1} \sum_{h=0}^{H} w_h \{\gamma_h^H(\delta) + \gamma_{-h}^H(\delta)\},
\]

(8)

- The factor \((1 - H\delta)^{-1}\) is a sensible small sample correction which makes no difference to the asymptotic analysis.
The asymptotic variance of the realised kernels we will use are

$$\text{Var}\left\{ \tilde{K}_H(X_\delta) \right\} \approx \varpi(IV, IQ, \omega^2, n, H)$$

$$= 4\omega^4 \left\{ \frac{12n}{H(H+1)(H+2)} + \frac{6}{5H} \right\} + \frac{9.6\omega^2}{H} IV + \frac{2}{n} \left( 3 + 2 \sum_{h=2}^{H} \nu_h^2 \right) tIQ.$$

Our simple rule-of-thumb is in the infeasible case

$$H_{\text{simple}}^* = 3.6867 \frac{\omega}{\sqrt{\int_0^1 \sigma_u^2 du}} \sqrt{n}$$

The feasible version of this is

$$\hat{H}_{\text{simple}}^* = 3.6867 \frac{\tilde{\omega}_\delta}{\sqrt{[X_\delta^*]_1}} \sqrt{n},$$

where $[X_\delta^*]_1$ is a RV estimator based on low frequency data.
Our focus is on the asymptotically pivotal t-statistic

\[ T(IV, IQ, \omega^2, n, H) = \frac{\left( \tilde{K}_H(X_{\delta})_1 - IV \right)}{\sqrt{\omega(IV, IQ, \omega^2, n, H)/\sqrt{n}}} \xrightarrow{L} N(0, 1), \]

We do this in two ways:

(i) [infeasible case] when \( \omega \) and \( \sigma \) are known, imply \( \hat{H} \) and \( \omega(\bullet) \),

Barndoff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.34)
Summary statistics for the symmetric kernel:

\[ \omega^2 = 0.001, \text{ number of reps.} = 50000 \]

<table>
<thead>
<tr>
<th>( n )</th>
<th>( H_{\text{simple}}^* )</th>
<th>( \tilde{K}<em>H(X</em>\delta) )</th>
<th>RMSE(( \tilde{K}<em>H(X</em>\delta) ))</th>
<th>%neg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>1.0</td>
<td>1.003</td>
<td>0.507</td>
<td>0.504</td>
</tr>
<tr>
<td>39</td>
<td>1.0</td>
<td>1.002</td>
<td>0.413</td>
<td>0.092</td>
</tr>
<tr>
<td>78</td>
<td>2.0</td>
<td>0.976</td>
<td>0.307</td>
<td>0</td>
</tr>
<tr>
<td>195</td>
<td>2.0</td>
<td>0.991</td>
<td>0.200</td>
<td>0</td>
</tr>
<tr>
<td>390</td>
<td>3.0</td>
<td>0.991</td>
<td>0.156</td>
<td>0</td>
</tr>
<tr>
<td>780</td>
<td>4.0</td>
<td>0.993</td>
<td>0.122</td>
<td>0</td>
</tr>
<tr>
<td>1560</td>
<td>5.0</td>
<td>0.995</td>
<td>0.096</td>
<td>0</td>
</tr>
<tr>
<td>4680</td>
<td>8.0</td>
<td>0.997</td>
<td>0.069</td>
<td>0</td>
</tr>
<tr>
<td>23400</td>
<td>18.0</td>
<td>0.999</td>
<td>0.044</td>
<td>0</td>
</tr>
</tbody>
</table>
Summary Statistics - Infeasible case - Stochastic Volatility

$\omega^2 = 0.001$, number of reps. = 200,000

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\bar{H}_{simple}^*$</th>
<th>Mean</th>
<th>Stdv.</th>
<th>0.5%</th>
<th>2.5%</th>
<th>97.5%</th>
<th>99.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>195</td>
<td>3.09</td>
<td>-0.002</td>
<td>1.063</td>
<td>0.21</td>
<td>1.99</td>
<td>95.77</td>
<td>98.52</td>
</tr>
<tr>
<td>390</td>
<td>4.17</td>
<td>-0.002</td>
<td>1.060</td>
<td>0.31</td>
<td>2.26</td>
<td>96.00</td>
<td>98.73</td>
</tr>
<tr>
<td>780</td>
<td>5.69</td>
<td>0.000</td>
<td>1.052</td>
<td>0.36</td>
<td>2.43</td>
<td>96.22</td>
<td>98.90</td>
</tr>
<tr>
<td>1560</td>
<td>7.84</td>
<td>0.001</td>
<td>1.047</td>
<td>0.40</td>
<td>2.54</td>
<td>96.38</td>
<td>98.99</td>
</tr>
<tr>
<td>4680</td>
<td>13.2</td>
<td>0.001</td>
<td>1.033</td>
<td>0.43</td>
<td>2.49</td>
<td>96.71</td>
<td>99.14</td>
</tr>
<tr>
<td>5850</td>
<td>14.7</td>
<td>0.002</td>
<td>1.029</td>
<td>0.44</td>
<td>2.51</td>
<td>96.82</td>
<td>99.18</td>
</tr>
<tr>
<td>7800</td>
<td>16.9</td>
<td>0.001</td>
<td>1.028</td>
<td>0.41</td>
<td>2.51</td>
<td>96.85</td>
<td>99.21</td>
</tr>
<tr>
<td>11700</td>
<td>20.6</td>
<td>-0.000</td>
<td>1.021</td>
<td>0.43</td>
<td>2.49</td>
<td>97.01</td>
<td>99.27</td>
</tr>
<tr>
<td>23400</td>
<td>28.9</td>
<td>-0.001</td>
<td>1.018</td>
<td>0.45</td>
<td>2.42</td>
<td>97.02</td>
<td>99.31</td>
</tr>
</tbody>
</table>

Barndorf-Nielsen, Hansen, Lunde & Shephard, 2005 (p.36)
(ii) [feasible case] estimating $\omega$ and $IV$ from the data, implying plugged in values of $\hat{H}$ and $\varpi(\bullet)$.

- Two-step approach
  1. choose $H$ use a low frequency RV to estimate $[Y]$
  2. then compute the estimator
  3. then compute the standard error

Barndoff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.37)
Use an unbiased kernel estimator of $\omega^2$

$$N_w(X_\delta) = \left\{ w_0 \tilde{V}_0(X_\delta) + \sum_{h=1}^{H} w_h \tilde{V}_h(X_\delta) \right\} / (-2n)$$

has $w = (\tilde{w}', v')'$, where $v$ is freely chosen and $\tilde{w} = (0, 1)'$ and

$$\hat{V}_h^\omega = (h + 1) \frac{(H - h)(H - h + 1)}{H(H + 1)}.$$ 

We use $H = c_\omega n^{1/3}$ with $c_\omega \approx 2.8377\omega / \sqrt{\int_0^t \sigma_u^2 du}$. 

Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.38)
Estimating integrated quarticity is a tougher problem than estimating QV as the effect of noise is magnified up.

Define the subsampled squared returns

\[ x_{j,.}^2 = \frac{1}{S + 1} \sum_{s=0}^{S} x_{j,s/(S+1)}^2, \quad j = 1, 2, \ldots, n. \]

This allows us to define a bipower variation estimator of integrated quarticity

\[ \{ X_\delta, \omega^2; S \}^{[2,2]} = \delta^{-1} \sum_{j=1}^{\lfloor t/\delta \rfloor} \left( x_{j,.}^2 - 2\omega^2 \right) \left( x_{j-2,.}^2 - 2\omega^2 \right). \]
We can improve the finite sample performance of our estimator of integrated quarticity by using the inequality

\[ \int_0^t \sigma_u^4 \, du \geq \frac{1}{t} \left( \int_0^t \sigma_u^2 \, du \right)^2. \]

Thus our preferred way of estimating integrated quarticity is

\[ \hat{IQ}_{\delta, S} = \max \left[ \frac{1}{t} \left( \tilde{K}_v \left( X_\delta ; S \right) \right)^2 , \{X_\delta, N_{\hat{v}_\omega} \left( X_\delta \right); S}\right]^{[2,2]}. \]
## Summary Statistics - Feasible case - Stochastic Volatility

\( \omega^2 = 0.001, \text{ number of reps.} = 200,000 \)

<table>
<thead>
<tr>
<th>n</th>
<th>( \bar{H}_{\text{simple}} )</th>
<th>Mean</th>
<th>Stdv.</th>
<th>0.5%</th>
<th>2.5%</th>
<th>97.5%</th>
<th>99.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>195</td>
<td>2.95</td>
<td>-0.249</td>
<td>1.131</td>
<td>3.62</td>
<td>7.50</td>
<td>99.57</td>
<td>99.98</td>
</tr>
<tr>
<td>390</td>
<td>3.90</td>
<td>-0.188</td>
<td>1.083</td>
<td>2.64</td>
<td>6.23</td>
<td>99.17</td>
<td>99.95</td>
</tr>
<tr>
<td>780</td>
<td>5.28</td>
<td>-0.145</td>
<td>1.052</td>
<td>2.00</td>
<td>5.31</td>
<td>98.92</td>
<td>99.90</td>
</tr>
<tr>
<td>1560</td>
<td>7.24</td>
<td>-0.116</td>
<td>1.036</td>
<td>1.55</td>
<td>4.64</td>
<td>98.63</td>
<td>99.86</td>
</tr>
<tr>
<td>4680</td>
<td>12.2</td>
<td>-0.085</td>
<td>1.018</td>
<td>1.14</td>
<td>3.93</td>
<td>98.36</td>
<td>99.78</td>
</tr>
<tr>
<td>5850</td>
<td>13.5</td>
<td>-0.077</td>
<td>1.012</td>
<td>1.10</td>
<td>3.79</td>
<td>98.35</td>
<td>99.78</td>
</tr>
<tr>
<td>7800</td>
<td>15.6</td>
<td>-0.072</td>
<td>1.012</td>
<td>1.01</td>
<td>3.67</td>
<td>98.30</td>
<td>99.75</td>
</tr>
<tr>
<td>11700</td>
<td>18.9</td>
<td>-0.065</td>
<td>1.005</td>
<td>0.93</td>
<td>3.48</td>
<td>98.28</td>
<td>99.74</td>
</tr>
<tr>
<td>23400</td>
<td>26.6</td>
<td>-0.056</td>
<td>1.004</td>
<td>0.82</td>
<td>3.28</td>
<td>98.13</td>
<td>99.72</td>
</tr>
</tbody>
</table>

*Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.41)*
An alternative to this is to use the delta method and base the asymptotic analysis on a log transformation

\[ T_{\text{log}} = \frac{\log \left\{ \tilde{K}_v(X_\delta) \delta + d \right\} - \log \left\{ \int_0^1 \sigma^2 u \, du + d \right\}}{\sqrt{\omega} / \left\{ \tilde{K}_v(X_\delta) \delta + d \right\}} \overset{L}{\rightarrow} N(0, 1). \]

The presence of \( d \geq 0 \) allows for the possibility that \( \tilde{K}_v(X_\delta) \delta \) may be truncated to be exactly zero.

In our simulations we have taken \( d = 0.2 \).
### Summary Statistics - Feasible case - Log version - Stochastic Volatility

\[ \omega^2 = 0.001, \text{ number of reps.} = 200,000 \]

<table>
<thead>
<tr>
<th>n</th>
<th>( \bar{H}_{\text{simple}} )</th>
<th>Mean</th>
<th>Stdv.</th>
<th>0.5%</th>
<th>2.5%</th>
<th>97.5%</th>
<th>99.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>195</td>
<td>2.95</td>
<td>-0.163</td>
<td>1.061</td>
<td>1.98</td>
<td>5.32</td>
<td>98.60</td>
<td>99.82</td>
</tr>
<tr>
<td>390</td>
<td>3.90</td>
<td>-0.125</td>
<td>1.045</td>
<td>1.48</td>
<td>4.58</td>
<td>98.29</td>
<td>99.72</td>
</tr>
<tr>
<td>780</td>
<td>5.28</td>
<td>-0.097</td>
<td>1.029</td>
<td>1.19</td>
<td>4.04</td>
<td>98.17</td>
<td>99.69</td>
</tr>
<tr>
<td>1560</td>
<td>7.24</td>
<td>-0.078</td>
<td>1.022</td>
<td>0.97</td>
<td>3.65</td>
<td>97.98</td>
<td>99.65</td>
</tr>
<tr>
<td>4680</td>
<td>12.2</td>
<td>-0.057</td>
<td>1.011</td>
<td>0.77</td>
<td>3.24</td>
<td>97.84</td>
<td>99.62</td>
</tr>
<tr>
<td>5850</td>
<td>13.5</td>
<td>-0.052</td>
<td>1.007</td>
<td>0.74</td>
<td>3.13</td>
<td>97.87</td>
<td>99.60</td>
</tr>
<tr>
<td>7800</td>
<td>15.6</td>
<td>-0.048</td>
<td>1.007</td>
<td>0.70</td>
<td>3.07</td>
<td>97.84</td>
<td>99.60</td>
</tr>
<tr>
<td>11700</td>
<td>18.9</td>
<td>-0.044</td>
<td>1.001</td>
<td>0.65</td>
<td>2.98</td>
<td>97.86</td>
<td>99.60</td>
</tr>
<tr>
<td>23400</td>
<td>26.6</td>
<td>-0.039</td>
<td>1.001</td>
<td>0.61</td>
<td>2.84</td>
<td>97.79</td>
<td>99.60</td>
</tr>
</tbody>
</table>

Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.43)
Inference for General Electric volatility; Five days in November 2004.

<table>
<thead>
<tr>
<th>Day</th>
<th>Trans</th>
<th>Lower</th>
<th>RV20m</th>
<th>Upper</th>
<th>n</th>
<th>Lower</th>
<th>KV60s</th>
<th>Upper</th>
<th>n</th>
<th>H</th>
<th>$\bar{\omega}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4631</td>
<td>0.48</td>
<td>0.83</td>
<td>1.46</td>
<td>20</td>
<td>0.50</td>
<td>0.70</td>
<td>0.99</td>
<td>357</td>
<td>4</td>
<td>.0016</td>
</tr>
<tr>
<td>2</td>
<td>4974</td>
<td>0.62</td>
<td>1.19</td>
<td>2.28</td>
<td>20</td>
<td>0.79</td>
<td>1.15</td>
<td>1.67</td>
<td>356</td>
<td>4</td>
<td>.0025</td>
</tr>
<tr>
<td>3</td>
<td>4918</td>
<td>0.51</td>
<td>0.92</td>
<td>1.63</td>
<td>20</td>
<td>0.73</td>
<td>1.02</td>
<td>1.42</td>
<td>352</td>
<td>4</td>
<td>.0021</td>
</tr>
<tr>
<td>4</td>
<td>5493</td>
<td>0.26</td>
<td>0.52</td>
<td>1.03</td>
<td>20</td>
<td>0.42</td>
<td>0.59</td>
<td>0.83</td>
<td>344</td>
<td>4</td>
<td>.0013</td>
</tr>
<tr>
<td>5</td>
<td>5504</td>
<td>0.65</td>
<td>1.26</td>
<td>2.44</td>
<td>20</td>
<td>1.12</td>
<td>1.56</td>
<td>2.17</td>
<td>344</td>
<td>3</td>
<td>.0028</td>
</tr>
</tbody>
</table>

Barndoff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.44)
Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.45)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. (HAC)</th>
<th>$\rho([\hat{Y}], \tilde{K})$</th>
<th>acf(1)</th>
<th>acf(2)</th>
<th>acf(5)</th>
<th>acf(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Optimal kernel ($H = cn^{1/2}$)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tilde{K}<em>w(X</em>{\text{ap. 1 min}})$</td>
<td>0.915</td>
<td>0.542 (1.172)</td>
<td>1.000</td>
<td>0.35</td>
<td>0.37</td>
<td>0.28</td>
<td>0.09</td>
</tr>
<tr>
<td><strong>Simple RV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_{20 \text{ minutes}}$</td>
<td>0.879</td>
<td>0.524 (1.008)</td>
<td>0.793</td>
<td>0.28</td>
<td>0.24</td>
<td>0.26</td>
<td>0.06</td>
</tr>
<tr>
<td>$X_{5 \text{ minutes}}$</td>
<td>0.948</td>
<td>0.518 (1.100)</td>
<td>0.935</td>
<td>0.36</td>
<td>0.34</td>
<td>0.26</td>
<td>0.10</td>
</tr>
<tr>
<td>$X_{1 \text{ minutes}}$</td>
<td>0.941</td>
<td>0.382 (0.919)</td>
<td>0.878</td>
<td>0.44</td>
<td>0.40</td>
<td>0.38</td>
<td>0.11</td>
</tr>
<tr>
<td>$X_{10 \text{ seconds}}$</td>
<td>1.330</td>
<td>0.389 (1.142)</td>
<td>0.801</td>
<td>0.60</td>
<td>0.56</td>
<td>0.51</td>
<td>0.32</td>
</tr>
<tr>
<td>$X_{1 \text{ second}}$</td>
<td>2.183</td>
<td>0.569 (1.828)</td>
<td>0.739</td>
<td>0.69</td>
<td>0.66</td>
<td>0.57</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Barndorf-Nielsen, Hansen, Lunde & Shephard, 2005 (p.46)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. (HAC)</th>
<th>( \rho(\hat{Y}, \tilde{K}) )</th>
<th>acf(1)</th>
<th>acf(2)</th>
<th>acf(5)</th>
<th>acf(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Optimal kernel</strong> ((H = cn^{1/2}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\tilde{K}<em>w(X</em>{ap. 1 \min}))</td>
<td>0.915</td>
<td>0.542 (1.172)</td>
<td>1.000</td>
<td>0.35</td>
<td>0.37</td>
<td>0.28</td>
<td>0.09</td>
</tr>
<tr>
<td>(\tilde{K}<em>w(X</em>{ap. 1 \min}; \lambda))</td>
<td>0.919</td>
<td>0.544 (1.181)</td>
<td>0.996</td>
<td>0.36</td>
<td>0.35</td>
<td>0.28</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Inefficient kernel</strong> ((H = cn^{2/3}))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\tilde{K}<em>w(X</em>{1 \text{ tick}}))</td>
<td>0.907</td>
<td>0.505 (1.115)</td>
<td>0.990</td>
<td>0.37</td>
<td>0.37</td>
<td>0.30</td>
<td>0.09</td>
</tr>
<tr>
<td><strong>Simple RV subsampled</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>([X_{20 \text{ minutes}}; 1200])</td>
<td>0.885</td>
<td>0.516 (1.036)</td>
<td>0.909</td>
<td>0.27</td>
<td>0.27</td>
<td>0.27</td>
<td>0.08</td>
</tr>
<tr>
<td>([X_{5 \text{ minutes}}; 300])</td>
<td>0.943</td>
<td>0.503 (1.088)</td>
<td>0.966</td>
<td>0.37</td>
<td>0.32</td>
<td>0.30</td>
<td>0.08</td>
</tr>
<tr>
<td>([X_{1 \text{ minutes}}; 60])</td>
<td>0.942</td>
<td>0.376 (0.921)</td>
<td>0.894</td>
<td>0.46</td>
<td>0.43</td>
<td>0.38</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>ZMA (2005)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSRV((X_{1 \text{ tick}}))</td>
<td>0.786</td>
<td>0.368 (0.832)</td>
<td>0.884</td>
<td>0.41</td>
<td>0.35</td>
<td>0.32</td>
<td>0.05</td>
</tr>
<tr>
<td>TSRV2((X_{1 \text{ tick}}))</td>
<td>0.861</td>
<td>0.408 (0.921)</td>
<td>0.893</td>
<td>0.41</td>
<td>0.35</td>
<td>0.32</td>
<td>0.05</td>
</tr>
</tbody>
</table>

*Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.47)*
Illustration.

- $\tilde{K}_w(X_{\text{ap. 1 min}}; K)$
- $\tilde{K}_w(X_{1 \text{ tick}})$
- $TSRV2(X_{1 \text{ tick}})$
- $[X_{5 \text{ minutes}}; 300]$

Days in November 2004 (GE)
8. SUMMARY

- **Kernel Estimators in Details:**
  - Derive bias and variance properties for all kernel estimators.
  - Relation to the consistent subsample estimator of ZMA.
    - Asymmetric kernels can be made consistent by subsampling.

- **Symmetric Kernel Estimators:**
  - Consistent and converges at rate $n^{1/4}$.
  - Estimation of $IQ$ and $\omega^2$.
  - Effect of stochastically spaced data.
  - Simulations and Empirical Illustration.

*Barndorff-Nielsen, Hansen, Lunde & Shephard, 2005 (p.49)*
Designing realised kernels

References


