

# Efficient and feasible inference for the components of financial variation using blocked multipower variation

PER A. MYKLAND

*Department of Statistics, University of Chicago  
Eckhart Hall, 5734 University Avenue, Chicago, IL 60637, USA  
mykland@pascal.uchicago.edu*

NEIL SHEPHARD

*Oxford-Man Institute, University of Oxford,  
Eagle House, Walton Well Road, Oxford OX2 6ED, UK  
£*

*Department of Economics, University of Oxford  
neil.shephard@economics.ox.ac.uk*

KEVIN SHEPPARD

*Oxford-Man Institute, University of Oxford,  
Eagle House, Walton Well Road, Oxford OX2 6ED, UK  
£*

*Department of Economics, University of Oxford  
kevin.sheppard@economics.ox.ac.uk*

30th May, 2011

## Abstract

High frequency financial data allows us to learn more about volatility and jumps. One of the key techniques developed in the literature in recent years has been bipower variation and its multipower extension, which estimates time-varying volatility robustly to jumps. We improve the scope and efficiency of bi- and multipower variation by the use of a more sophisticated exploitation of high frequency data. This suggests very significant improvements in the power of jump tests. It also yields efficient estimates of the integrated variance of the continuous part of a semimartingale. The paper also shows how to extend the theory to the case where there is microstructure in the observations. A fundamental device in the paper is a new type of result showing path-by-path (strong) approximation between multipower and the (unobserved) RV based on the continuous part of the process. Our results also have relevance for estimation of the volatility of volatility.

Keywords: bipower variation; jumps; market microstructure noise; multipower variation; quadratic variation; semimartingale; volatility; volatility of volatility.

## 1 Introduction

High frequency data has been demonstrated to improve our ability to understand and forecast financial volatility (e.g. Andersen, Bollerslev, Diebold, and Labys (2001) and Barndorff-Nielsen and Shephard (2002)). Estimators like realised volatility have also been shown, when combined with implied volatility, to powerfully explain variations in the cross sectional and temporal behaviour of risk premia (e.g. Bollerslev, Tauchen, and Zhou (2009)). Further, multipower variation and

threshold tests based on high frequency data have shown convincingly that relatively frequent jumps play an important role in the evolution of price processes for commonly held assets (e.g. Barndorff-Nielsen and Shephard (2006), Huang and Tauchen (2005), Andersen, Bollerslev, and Diebold (2007) and Patton and Sheppard (2009)) while jumps play an important role in determining extreme moves in financial markets (e.g. Bollerslev and Todorov (2011)).

All of the above methods are based on the following framework. Suppose there is an underlying efficient price process  $X$ , which will be assumed to be a semimartingale due to the absence of arbitrage (e.g. Delbaen and Schachermeyer (2006)). We observe  $n$  returns over some fixed interval of time, say time 0 to time  $T$ , typically representing a day. Then a natural ex-post measure of its variation is  $X$ 's quadratic variation (QV) (e.g. Andersen, Bollerslev, and Diebold (2009) and Barndorff-Nielsen and Shephard (2007)). Some of the literature on this topic have formally allowed for the effect of market microstructure effects on the analysis (e.g. Zhang, Mykland, and Ait-Sahalia (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) and Jacod, Li, Mykland, Podolskij, and Vetter (2009)), others use moderate frequency data such as five minute returns in order to ameliorate the effect (e.g. Andersen, Bollerslev, Diebold, and Labys (2000) and Bandi and Russell (2008)).

In this paper we focus on multipower type estimators and tests. Most of the time we will be using these statistics to estimate and make inference on the quadratic variation of  $X$ ,  $[X, X]_T$ , and the quadratic variation of the continuous part of  $X$ ,  $[X^c, X^c]_T$ . Multipower variation looks inside quadratic variation, splitting up the variation into that due to continuous evolution of prices and that due to jumps. It was introduced by Barndorff-Nielsen and Shephard (2004) and Barndorff-Nielsen and Shephard (2006). The choice of which power  $K$  determines the configuration of the statistic, with  $K = 1$  being realised variance,  $K = 2$  being bipower variation,  $K = 3$  being tripower variation, etc.

The contribution of this paper is to extend multipower variation type estimators, allowing them to be more efficient. We call this new estimator "blocked multipower variation".

The ideas here echo Mykland and Zhang (2009) who show how to justify the use of blocks of length  $M$  of high frequency returns. Econometrically we go much further however. Our results generate a path-by-path analysis of the limiting behaviour of high frequency estimators. This is new for the literature on high frequency financial econometrics and this type of new analysis may have very wide applicability outside the scope of this paper.

To put this in context suppose log-prices have two components

$$X_t = X_t^c + J_t, \quad t \in \mathcal{R}_{\geq 0},$$

where  $X_t^c$  is a continuous semimartingale  $X_t^c = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$  and  $J$  is a purely discontinuous finite activity jump process. Write  $RV^*$  as the infeasible realised variance of the  $X^c$  process. So  $RV^*$  has the jumps stripped out of it and so, as long as  $\sigma_t > 0$ , we have that

$$n^{1/2} \left( RV - RV^* - \sum_{0 < t \leq T} |\Delta J_t|^2 \right) \xrightarrow{\mathcal{L}^s} N \left( 0, 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2 \right), \quad (1)$$

where  $\xrightarrow{\mathcal{L}^s}$  denotes stable convergence.<sup>1</sup> We will prove that the jump variation  $JV = \sum_{0 < t \leq T} |\Delta J_t|^2$  cannot be estimated with a lower normalised variance than the scaled asymptotic variance of

---

<sup>1</sup>Let  $Z_n$  be a sequence of  $\mathcal{F}_T$ -measurable random variables. We say that  $Z_n$  converges stably in law to  $Z$  as  $n \rightarrow \infty$  if  $Z$  is measurable with respect to an extension of  $\mathcal{F}_T$  so that for all  $A \in \mathcal{F}_T$  and for all bounded continuous  $g$ ,  $E\{I_A g(Z_n)\} \rightarrow E\{I_A g(Z)\}$  as  $n \rightarrow \infty$ .  $I_A$  denotes the indicator function of  $A$ , and  $= 1$  if  $A$  and  $= 0$  otherwise. The same definition applies to triangular arrays. In the context of inference,  $Z_n = n^{1/2}(\hat{\theta}_n - \theta)$ , for example, and  $Z = N(b, a^2)$ . For further discussion of stable convergence, and for the commutation relationship to measure change, see Section 2 of Jacod and Protter (1998) Section 2.2 of Mykland and Zhang (2009), which draws on Rootzén (1980). The latter is relevant for our proofs, starting with that of Theorem 3.

$RV - RV^*$

$$4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2, \quad (2)$$

while it is well known that integrated variance  $IV = \int_0^T \sigma_t^2 dt$  cannot be estimated with a lower normalised variance than  $2 \int_0^T \sigma_t^2 dt$ , a bound which is obtained by  $RV^*$  (e.g. Barndorff-Nielsen and Shephard (2002)). Of course neither is immediately feasible as we do not know  $RV^*$ . This will be where blocked multipower variation comes in.

If the size of the block length  $M$  of our new multipower variation estimator  $MV_M^{(K)}$  increases as  $M \approx cn^{1-\beta}$  then we show (in Theorem 3) that

$$\begin{aligned} n^{1/2} \left( MV_M^{(K)} - RV^* \right) &= T^{\frac{K-1}{K}} K c^{\frac{K-1}{K}} n^{-\beta \frac{K-1}{K} + \frac{1}{2}} \sum_{0 < t \leq T} \sigma_t^{2 \frac{K-1}{K}} |\Delta J_t|^{2/K} \\ &\quad - n^{-\beta + \frac{1}{2}} \frac{K-1}{3} c^{-1} [\sigma, \sigma]_T \\ &\quad + O_p \left( n^{\frac{1}{2} - \frac{3}{2}\beta} + n^{\frac{1}{2}(\beta-1)} \right) + o_p \left( n^{\frac{1}{2} - \beta \frac{K-1}{K}} \right), \end{aligned} \quad (3)$$

where  $[\sigma, \sigma]_T$  is the volatility of volatility. Thus we quantify the path-by-path difference between the multipower and  $RV^*$  statistics and the impact of the jumps and the quadratic variation of volatility. This is a significant strengthening over a previous result in Barndorff-Nielsen, Shephard, and Winkel (2006) which showed that when  $K \geq 3$  that  $n^{1/2} \left( MV_1^{(K)} - MV_1^{(K)*} \right) = o_p(1)$ , where  $MV_1^{(K)*}$  is the infeasible multipower variation statistic applied to the  $X_t^c$  process. Of course  $MV_1^{(K)*}$  is not asymptotically equivalent to  $RV^*$ , so the results are distinct.

Now note that (3) means that if the QV of volatility  $[\sigma, \sigma]_T$  is high but there are no jumps then  $MV_M^{(K)}$  will be low compared to  $RV^*$ , while jumps will push  $MV_M^{(K)}$  up a little. These effects diminish as  $n$  increases if  $\beta$  and  $K$  are appropriately chosen. The theory also shows that  $MV_M^{(K)}$  and  $RV^*$  are asymptotically equivalent under jumps if  $\beta \in (1/2, 1)$  and  $\beta \frac{K-1}{K} > 1/2$ . This implies, as  $K$  is an integer, we must have at least  $K \geq 3$ . A combination which has an attractive optimality property is  $\beta = 6/7$  (so the block size increases slowly with the sample size) when  $K = 3$ . Asymptotically the jump term has the dominant larger term, hence the rates of convergence to zero of the two terms in (3) become similar as  $K$  increases. These results mean that under the presence of jumps

$$n^{1/2} \left( MV_M^{(K)} - \int_0^T \sigma_t^2 dt \right) \xrightarrow{\mathcal{L}^s} N \left( 0, 2T \int_0^T \sigma_t^4 dt \right) \quad (4)$$

and

$$n^{1/2} \left\{ \left( RV - MV_M^{(K)} \right) - \sum_{0 < t \leq T} |\Delta J_t|^2 \right\} \xrightarrow{\mathcal{L}^s} N \left( 0, 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2 \right). \quad (5)$$

We thus demonstrate that this feasible approach to estimating  $IV$  and  $JV$  is efficient, and is helpful in generating narrow confidence intervals for the components of quadratic variation — for we can consistently estimate the “integrated quarticity”  $IQ = \int_0^T \sigma_t^4 dt$  and “squared jump times vol”  $SJV = \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2$  using block multipower statistics.

One of the interesting features of (3) is that it indicates the squared jumps appear in the asymptotics in a simple additive way, which means their effect depends roughly on the size of the quadratic variation of the jumps rather than, for example, the size of the largest jumps. This makes us speculate that this asymptotic equivalence result may also hold for infinite activity processes.

One indication in this direction are the results of Veraart (2010), who studies the estimation of  $JV$  by looking at the difference between  $RV$  and classical  $MV$  (block size 1) also for infinite activity processes, and who obtains results that are similar to ours to the extent that our two setups coincide.<sup>2</sup> However, her single block size means that her inference does not obtain the nonparametric efficiency bounds.

Alternative methods for looking at jumps include the threshold type tests instigated by Mancini (2004) and further developed by, for example, Mancini (2009), Aït-Sahalia and Jacod (2007), Jacod (2008a) and Jacod (2008b). Christensen and Podolskij (2006) have looked at bipower type objects based upon ranges, Aït-Sahalia and Jacod (2008) considers parametric methods, while Andersen, Dobrev, and Schaumburg (2009) use a nearest neighbour method. Barndorff-Nielsen, Kinnebroeck, and Shephard (2009) take the analysis further and analyse positive and negative jumps separately.

For estimation purposes, the strong approximation in Theorem 3 means that the asymptotics of realised volatility carries over directly to asymptotics of bi- and multipower variation.

One of the principle uses of multipower variation is to nonparametrically test for jumps over the interval  $[0, T]$ . Under the null of no jumps as  $M \approx cn^{1-\beta}$ , so long as  $\beta > 1/2$ , then we show that

$$J = \frac{n^{1-\beta/2} \left( RV - MV_M^{(K)} \right)}{\sqrt{T \frac{3}{4c} \int_0^T \sigma_s^4 ds}} \xrightarrow{\mathcal{L}} N(0, 1).$$

The null is rejected when we see a large value of  $J$ . This result shows the asymptotic power of the Barndorff-Nielsen and Shephard (2006) jump test can be improved by an infinite amount by using blocked multipower variation.

This paper has the following structure. In Section 2 we make our assumptions and notation clear, reviewing the structure of the standard bipower variation statistic. In Section 3 we extend bipower to allow for more efficiency and discuss the properties of the resulting statistic. In Section 3.5 and Appendix A.1 we provide the technically most challenging section which allows the block size of the new efficient multipower power statistic to go to infinity with the sample size. These sections also have bearing on the estimation of the volatility of volatility. In Section 8 we draw some conclusions and the Appendix has the proofs of various results we give in the main text of the paper.

## 2 Framework

### 2.1 Model and measures of variation

We will work with a univariate Brownian semimartingale (Ito process) model for log-prices defined on some filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ ,

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \tag{6}$$

where  $\mu$  and  $\sigma$  are predictable locally bounded drift and volatility processes, and  $W$  is a standard Brownian motion adapted to  $(\mathcal{F}_t)$ . For reviews of the econometrics of this type of process see, for example, Ghysels, Harvey, and Renault (1996).

Our focus will be on econometric estimators based on returns. We assume prices are recorded at times  $0 = t_0, t_1, \dots, t_n = 1$ , which we will assume are equally spaced

$$t_i = \frac{i}{n}T, \quad i = 0, 1, \dots, n.$$

---

<sup>2</sup>Note that there are issues of concerning the mode of convergence for infinite activity processes, Jacod and Protter (1998, Section 6) on Euler discretisation of Lévy processes.

These returns will be written as

$$\Delta X_{t_i} = X_{t_i} - X_{t_{i-1}}.$$

The standard realised variance (RV) is

$$RV = \sum_{0 < t_i \leq T} (\Delta X_{t_i})^2,$$

while its square root is usually labelled the realised volatility (RVol). The realised bipower variation statistic of Barndorff-Nielsen and Shephard (2004) is based on

$$BV = \frac{1}{k_1} \sum_{0 < t_{i-1}, t_i \leq T} |\Delta X_{t_i}| |\Delta X_{t_{i-1}}|,$$

where  $k_1 = \frac{2}{\pi} \simeq 0.63661$ , which is a special case of

$$k_M = \{E\chi_M\}^2 = 2 \left\{ \frac{\Gamma((M+1)/2)}{\Gamma(M/2)} \right\}^2, \quad \text{where } \chi_M \sim |\chi_M^2|^{1/2}. \quad (7)$$

Remark 2 in Section 3.2 will note some properties of  $k_M$ .

This bipower statistic has been extensively used in the last few years as it is robust to jumps so that if the  $X_t$  process also has jumps (in addition to the dynamics given in (6)),

$$BV \xrightarrow{p} \int_0^T \sigma_s^2 ds$$

under very weak conditions. This compares with the celebrated result

$$RV \xrightarrow{p} \int_0^T \sigma_s^2 ds + \sum_{0 \leq t \leq T} (X_t - X_{t-})^2 = IV + JV,$$

the sum of integrated variance and jump variation. See the discussion in Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen, Shephard, and Winkel (2006), Jacod and Shiryaev (2003) and Jacod (2008a).

The idea of bipower variation has been generalised in a number of directions, which are collectively now called multipower variation (e.g. Barndorff-Nielsen and Shephard (2006) and Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006)). Its target is the more general object  $\int_0^T \sigma_t^r dt$ ,  $r > 0$ . Typically focus is on  $r = 2$ , which we have discussed before, and  $r = 4$  which is useful for carrying out inference as integrated quarticity appears in the standard error of a large number of econometrically interesting quantities (e.g. Andersen, Dobrev, and Schaumburg (2011)). The multipower extension of bipower variation is

$$MV^{(r,K)} = \frac{1}{(k_{1,r/K})^K} \sum_{0 < t_{i-k} \leq 1} \prod_{k=0}^{K-1} |\Delta X_{t_{i-k}}|^{r/K}, \quad k_{M,s} = E\{(\chi_M)^s\},$$

noting that when  $r = 2$  we have that  $K = 1$  delivers RV,  $K = 2$  being BV and  $K = 3$  being tripower variation. So long as  $K > 1$  then this statistic is robust to jumps and  $MV^{(2,K)} \xrightarrow{p} \int_0^T \sigma_s^2 ds$ . In the rest of this section we will focus on RV and BV variation, returning to the more general case later.

## 2.2 Testing for jumps using bipower variation

It is now well known (Barndorff-Nielsen and Shephard (2006) and, e.g., Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006)) that under (6)

$$\sqrt{n} \begin{pmatrix} RV - \int_0^T \sigma_s^2 ds \\ BV - \int_0^T \sigma_s^2 ds \end{pmatrix} \xrightarrow{\mathcal{L}^s} N \left( 0, \begin{pmatrix} 2 & 2 \\ 2 & \theta_1 \end{pmatrix} T \int_0^T \sigma_s^4 ds \right),$$

where  $\mathcal{L}^s$  denotes convergence in law stably. Here  $\theta_1 \simeq 2.6090$  and is given as a special case of

$$\begin{aligned} \theta_M &= \frac{1}{k_M^2} [\text{Var}(\chi_{M,1}\chi_{M,2}) + 2\text{Cov}(\chi_{M,1}\chi_{M,2}, \chi_{M,3}\chi_{M,2})] \\ &= \frac{1}{k_M^2} [\{\text{E}(\chi_M^2)\}^2 - k_M^2 + 2(k_M \text{E}(\chi_M^2) - k_M^2)] \\ &= \frac{1}{k_M^2} (M^2 + 2k_M M - 3k_M^2), \end{aligned} \tag{8}$$

where  $\{\chi_{M,i}\}$  are independent with law  $\chi_M$  and  $k_M = k_{M,1}$ .

Barndorff-Nielsen and Shephard (2006) suggested using the difference of these statistics

$$\sqrt{n}(RV - BV) \xrightarrow{\mathcal{L}^s} N \left( 0, T(\theta_1 - 2) \int_0^T \sigma_s^4 ds \right),$$

as a basis for a test for jumps as  $BV$  is known to be robust to jumps. The resulting test records

$$J = \frac{\sqrt{n}(RV - BV)}{\sqrt{T(\theta_1 - 2)\widehat{IQ}}} \xrightarrow{\mathcal{L}^s} N(0, 1),$$

where  $\widehat{IQ}$  is a jump robust consistent estimator of integrated quarticity  $IQ = \int_0^T \sigma_s^4 ds$ . This method has been used and extended extensively empirically e.g. Huang and Tauchen (2005), Andersen, Bollerslev, and Diebold (2007) and Lee and Mykland (2008).

The above results show that  $BV$  is not as efficient as  $RV$  when there are no jumps. It opens up the question of whether there is a simple generalisation of  $BV$  which is significantly more efficient and can this test be made more efficient as a result?

## 2.3 A different view of bipower variation

Before we introduce our generalisation of  $BV$  it is helpful to interpret  $BV$  in the following way. Think of  $|\Delta X_{t_i}|$  as roughly in law  $\chi_1 n^{-1/2} \sigma_{t_{j-1}}$ , while  $\int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt$  is roughly estimated by

$$RV_i = \sum_{\tau_{i-1} < t_j \leq \tau_i} (\Delta X_{t_j})^2, \quad i = 1, \dots, n.$$

That is the absolute value of returns is the square root of a very local realised variance estimator — computing  $RV_i$  over only a single observation. Then we can write

$$BV = \frac{1}{k_1} \sum_{i=2}^n |RV_i|^{\frac{1}{2}} |RV_{i-1}|^{\frac{1}{2}}.$$

There are then two natural questions.

- Are there gains to be made by taking longer blocks to compute a better local RV estimator, before computing  $BV$ ?
- Would the block version still be robust to jumps?

We study these questions in the next section.

### 3 Improving the efficiency of multipower variation

#### 3.1 Blocked bipower variation

Set temporal block boundaries as

$$\tau_i = M \frac{i}{n} T, \quad i = 0, 1, \dots, \lfloor n/M \rfloor = n_M,$$

where  $M$ , the number of high frequency returns within a block, is a positive integer so  $M \leq n$ . Then we define the (non-overlapping)  $i$ -th blocked realised variance as

$$RV_i = \sum_{\tau_{i-1} < t_j \leq \tau_i} (\Delta X_{t_j})^2, \quad i = 1, 2, \dots, n_M. \quad (9)$$

This computes the realised variance using all  $M$  high frequency observations inside the block  $(\tau_{i-1}, \tau_i]$  and estimates

$$\int_{\tau_{i-1}}^{\tau_i} \sigma_s^2 ds.$$

Thus if  $M = 1$  we get a single squared return, while if  $M = n$  we get back to the full sample RV.

If we do this for all feasible  $i$ , then we have a time series, within the day, of miniature non-overlapping realised variances  $RV_1, RV_2, \dots, RV_{n_M}$ . Each of these is quite noisy, but they can be averaged. If they were summed we reproduce the realised variance  $RV = \sum_{i=1}^{n_M} RV_i$ . Our focus is on the *block bipower variation*, which we define as

$$BV_M = \frac{M}{k_M} \sum_{i=2}^{n_M} (RV_{i-1} RV_i)^{1/2}, \quad (10)$$

where  $k_M$  is as defined in (7).

In practice it is better to use an alternative definition

$$BV_M = \frac{n}{(n_M - 1) k_M} \sum_{i=2}^{n_M} (RV_{i-1} RV_i)^{1/2}. \quad (11)$$

$BV_M$  is exactly unbiased for constant  $\sigma_t$ , see Remark 1 below. For most of the paper, we shall not distinguish between (10) and (11) since  $n/(n_M - 1) = M + O(M/n)$ . In particular, the two statistics are asymptotically equivalent for fixed  $M$ . If we do distinguish, we say so explicitly.

**Remark 1** *If  $X$  is scaled Brownian motion  $\sigma W$ , and with the definition (11), then*

$$E(BV_M) = \sigma^2 T, \quad \text{Var}(BV_M) = T^2 \sigma^4 d \sim T^2 \frac{\sigma^4}{n} M \theta_M,$$

where

$$d = \frac{1}{M(n_M - 1)} \left( M \theta_M - \frac{1}{(n_M - 1)} \psi_M \right), \quad \psi_M = 2M(M/k_M - 1) = 1 + \frac{1}{4M} + o(M^{-2}). \quad (12)$$

Proof. This is in the Appendix.

### 3.2 Properties of $BV_M$ for fixed $M$

Under the more general conditions set out in Appendix B.1, we produce the following result.

**Theorem 1** *Under Assumption 1-2, for fixed  $M$  as  $n \rightarrow \infty$ :*

$$n^{1/2} \left( BV_M - \int_0^T \sigma_t^2 dt \right) \xrightarrow{\mathcal{L}^s} N \left( 0, TM\theta_M \int_0^T \sigma_s^4 ds \right), \quad (13)$$

likewise

$$n^{1/2} (RV - BV_M) \xrightarrow{\mathcal{L}^s} N \left( 0, T(M\theta_M - 2) \int_0^T \sigma_s^4 ds \right). \quad (14)$$

**Proof.** Given in the Appendix.

**Remark 2** *Note that (see, for example, Abramowitz and Stegun (1970, 6.1.47) and Qi (2010, equation (1.11))) for large  $z$  (PER CAN YOU CHECK THIS)*

$$z^{1/2} \frac{\Gamma(z)}{\Gamma(z + \frac{1}{2})} \simeq \left( 1 + \frac{1}{8z} + \frac{3}{16} \frac{1}{24z^2} \right).$$

So for large  $M$

$$\frac{M}{k_M} = \left( \sqrt{\frac{M}{2}} \frac{\Gamma(\frac{M}{2})}{\Gamma(\frac{M}{2} + \frac{1}{2})} \right)^2 \simeq \left( 1 + \frac{1}{4M} + \frac{1}{16} \frac{1}{2M^2} \right)^2 \simeq 1 + \frac{1}{2M} + \frac{1}{8M^2}.$$

Thus

$$M\theta_M = M \left\{ \left( \frac{M^2}{k_M^2} - 1 \right) + 2 \left( \frac{M}{k_M} - 1 \right) \right\} \simeq 2 + \frac{3}{4M}.$$

This means that  $BV_M$  is roughly efficient if  $M$  is large. Table 1 shows the relative efficiency of  $BV_M$  as  $M$  varies. It suggests by the time  $M = 3$  we have obtained a modest reduction in the asymptotic variance, but that this gain is two-thirds of all the potential gains from using the blocking. It makes realised  $BV$  nearly as efficient as  $RV$ . Under the null of no jumps  $RV - BV_M$  has a much smaller variance which should importantly improve  $J_M$ 's power as a test.

$M$	1	2	3	5	10	25	50	$\infty$
$M\theta_M$	2.609	2.335	2.232	2.143	2.073	2.029	2.014	2
$M\theta_M - 2$	.609	.335	.232	.143	.073	.029	.014	0
$\frac{3}{4M}$	.750	.375	.250	.150	.075	.030	.015	0

Table 1: Efficiency of the blocked bipower variation (for  $T = 1$ ). The asymptotic efficiency of this procedure is determined by  $M\theta_M$ , with the best possible number being 2. In testing for jumps the efficiency is determined by  $M\theta_M - 2$ .

**Remark 3** *If  $M \rightarrow \infty$  with  $n$ , specifically,  $M \sim cn^{1-\beta}$ , with  $\beta > 1/2$ , then the asymptotic variance in (14) converges to that given by Theorem 7 below. In other words, the finite sample variance is consistent with results on the double asymptotics  $M \rightarrow \infty$  as  $n \rightarrow \infty$ . To see why, using the development in Remark 2 that  $n^{1/2}(BV - RV)$  has an asymptotic variance of the form*

$$T(M\theta_M - 2) \simeq \frac{3}{4cn^{1-\beta}}$$

What this says is that Theorem 1 predicts that

$$n^{1-\beta/2}(RV - BV) \xrightarrow{Ls} N\left(0, \frac{3}{4c}T \int_0^T \sigma_s^4 ds\right).$$

This is exactly the variance in the central limit Theorem 7 when  $\beta > 1/2$  where  $M$  increases with  $n$ . Thus,  $BV$  is essentially efficient for large  $M$ . See Section 3.5 for the formal development of a theory which allows  $M \rightarrow \infty$  as  $n \rightarrow \infty$ .

**Remark 4** If there are finitely many jumps in addition to the continuous part of  $X$ , this estimator is unaffected by these, asymptotically. Hence it is similar to bipower. This is driven by the fact that if there is no jump then  $\sqrt{RV_{i-1}} = O_p\left(\sqrt{\frac{M}{n}}\right)$ . Hence so long as  $M/n \rightarrow 0$  then under finite activity jumps  $BV_M \xrightarrow{p} \int_0^T \sigma_s^2 ds$  so this statistic is robust to finite activity jumps while the block bipower jump test statistic (15) will be consistent. A more detailed theoretical analysis of jumps will appear in Section 3.5.

The implication of this is that the block bipower test takes the form

$$J_M = \frac{n^{1/2}(RV - BV_M)}{\sqrt{(TM\theta_M - 2)\widehat{IQ}}} \xrightarrow{Ls} N(0, 1), \quad (15)$$

as  $n \rightarrow \infty$ , rejecting the null of no jump if  $J_M$  is significantly positive.

### 3.3 Blocked multipower variation

The above arguments extend to blocked multipower variation

$$MV_M^{(r,K)} = \left(\frac{n}{T}\right)^{\frac{r}{2}-1} \frac{M}{(k_{M,r/K})^K} \sum_{i=K}^{n_M} \prod_{k=0}^{K-1} (RV_{i-k})^{r/2K}, \quad \text{where}$$

$$k_{M,r} = E\{(\chi_M)^r\} = 2^{r/2} \frac{\Gamma\left(\frac{M+r}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}.$$

When  $K = 3$  this is called a blocked tripower estimator, while when  $K = 4$  it is a blocked quadpower estimator. Both are used frequently empirically when  $M = 1$ . As with (11), when  $M$  is moderately large, one is better off using a finite sample exact constant:

$$MV_M^{(r,K)} = \frac{n^{\frac{r}{2}}}{T^{\frac{r}{2}-1}} \frac{1}{(n_M - K + 1)(k_{M,r/K})^K} \sum_{i=K}^{n_M} \prod_{k=0}^{K-1} (RV_{i-k})^{r/2K}. \quad (16)$$

As with bipower variation, the two versions are equivalent when  $M/n = O_p(1)$ , but the version (16) is preferred for moderate  $M$ .

We now look at relative efficiency as a function of  $M$  and  $K$ .

**Remark 5** Assume that  $\sigma$  is constant. Then the  $RV_i$  are i.i.d.  $\chi_M^2 \sigma^2 T/n$ , and so

$$MV_M^{(r,K)} \stackrel{\mathcal{L}}{=} \sigma^r T \frac{1}{(n_M - K + 1)(k_{M,r/K})^K} \sum_{i=K}^{n_M} \prod_{k=0}^{K-1} v_{i-k}, \quad \text{where } v_i \stackrel{iid}{\sim} (\chi_M)^{r/K}.$$

Now the mean is  $\sigma^r T$  (this is exact in the case of (16)), and the variance is

$$\sigma^{2r} T^2 d_{n,M,K} \sim \sigma^{2r} T^2 \frac{M}{n} \theta_M^{(r,K)}, \quad (17)$$

as  $n \rightarrow \infty$ , where

$$d_{n,M,K} = \frac{1}{M(n_M - K + 1)^2} \frac{1}{(k_{M,r/K})^{2K}} \times \left\{ (n_M - K + 1) \text{Var} \left( \prod_{i=1}^K v_i \right) + 2 \sum_{k=1}^{K-1} (n_M - K + 1 - k) \text{Cov} \left( \prod_{i=1}^K v_i, \prod_{i=1}^K v_{i-k} \right) \right\}.$$

Here

$$\theta_M^{(r,K)} = \frac{1}{(k_{M,r/K})^{2K}} \left\{ \text{Var} \left( \prod_{i=1}^K v_i \right) + 2 \sum_{k=1}^{K-1} \text{Cov} \left( \prod_{i=1}^K v_i, \prod_{i=1}^K v_{i-k} \right) \right\}.$$

Now

$$\begin{aligned} \text{Cov} \left( \prod_{i=1}^K v_i, \prod_{i=1}^K v_{i-k} \right) &= \left[ \prod_{i=k+1}^K \text{E}(v_i^2) - \left\{ \prod_{i=k+1}^K \text{E}(v_i) \right\}^2 \right] \prod_{i=1}^k \text{E}(v_i)^2 \\ &= \left\{ (k_{M,2r/K})^{K-k} - (k_{M,r/K})^{2(K-k)} \right\} (k_{M,r/K})^{2k} \\ &= \left( \lambda_{M,r,K}^{K-k} - 1 \right) (k_{M,r/K})^{2K} \end{aligned}$$

where

$$\lambda_{M,r,K} = \frac{(k_{M,2r/K})}{(k_{M,r/K})^2} = \frac{\Gamma\left(\frac{M+2r/K}{2}\right)}{\Gamma\left(\frac{M+r/K}{2}\right)^2} \geq 1.$$

Thus

$$\begin{aligned} \theta_M^{(r,K)} &= (\lambda_{M,r,K}^K - 1) + 2 \sum_{k=1}^{K-1} (\lambda_{M,r,K}^{K-k} - 1) = 2\lambda_{M,r,K} \sum_{k=0}^{K-1} \lambda_{M,r,K}^k - \lambda_{M,r,K}^K - 2K + 1 \\ &= \frac{2\lambda_{M,r,K} (\lambda_{M,r,K}^{K-1} - 1)}{\lambda_{M,r,K} - 1} - (\lambda_{M,r,K}^K - 1) - 2K. \end{aligned}$$

### 3.4 Properties of $MV_M^{(r,K)}$ for fixed $M$

In the general case when  $\sigma$  is nonconstant, we can proceed as in the proof of Theorem 1, to obtain

**Theorem 2** Under Assumptions 1-2, for fixed  $M$  as  $n \rightarrow \infty$ :

$$n^{1/2} \left( MV_M^{(r,K)} - \int_0^T \sigma_t^r dt \right) \xrightarrow{\mathcal{L}_s^s} N \left( 0, TM \theta_M^{(r,K)} \int_0^T \sigma_s^{2r} ds \right). \quad (18)$$

**Proof.** Follows exactly the same lines as the proof of Theorem 1, so is omitted here.

Also, by the same development, we obtain,

$$n^{1/2} \left( MV_M^{(2,K)} - RV \right) \xrightarrow{\mathcal{L}_s^s} N \left( 0, T \left( M \theta_M^{(2,K)} - 2 \right) \int_0^T \sigma_s^4 ds \right), \quad (19)$$

which can be used for testing.

The exact values of  $\theta_M^{(r,K)}$  are presented in Table 2 for  $r = 2$  and  $r = 4$ , the values most important in empirical work.

$M$	$r = 2$				$r = 4$			
	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 1$	$K = 2$	$K = 3$	$K = 4$
1	2.000	2.609	3.061	3.377	10.66	12.00	13.65	14.91
2	2.000	2.335	2.556	2.697	10.00	10.00	10.61	11.09
3	2.000	2.232	2.376	2.465	9.600	9.333	9.702	9.985
5	2.000	2.143	2.228	2.277	9.142	8.800	9.000	9.156
10	2.000	2.073	2.114	2.136	8.666	8.400	8.492	8.564
25	2.000	2.029	2.045	2.054	8.296	8.160	8.194	8.222
50	2.000	2.014	2.022	2.027	8.153	8.080	8.097	8.110
$\infty$	2.000	2.000	2.000	2.000	8.000	8.000	8.000	8.000
<i>approx</i>	2.000	$2 + \frac{3}{4M}$	$2 + \frac{1.14}{M}$	$2 + \frac{4}{3M}$	$8 + \frac{8}{M}$	$8 + \frac{4}{M}$	$8 + \frac{4.84}{M}$	$8 + \frac{11}{2M}$

Table 2: Relative efficiency is determined by  $M\theta_M^{(r,K)}$ . When  $r = 2$  the optimal value of  $M\theta_M^{(r,K)}$  is 2, when  $r = 4$  the best value of  $M\theta_M^{(r,K)}$  is 8.  $M$  is the degree of blocking,  $K$  is the degree of multipowering.  $K = 2$  is a bipower variation statistic,  $K = 4$  is quadpower.

The  $r = 4$  case is interesting. It is easy to see using the delta method that the efficiency bound for this is 8. When there are no jumps using sums of fourth power of returns is the most efficient estimator of these estimators, but all of them are pretty inefficient. By the time  $M$  reaches 5 all the estimators are substantially improved. This is particularly the case for the  $K = 3$  and  $K = 4$  statistics — which are attractive as they are robust to jumps.

The  $r = 2$  case is less interesting, the results are also given in Table 2. This shows the substantial efficiency loss of using higher values of  $K$  and that increasing  $M$  ameliorates the worse of these effects. Indeed by the time  $K = 5$  the difference in efficiency between the different estimators is very modest.

### 3.5 Properties of $MV_M^{(2,K)}$ for $M$ increasing with $n$

We now make the assumption that

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + J_t, \quad (20)$$

where  $J_t$  is a finite activity jump process. We allow, as  $n \rightarrow \infty$ ,

$$n^\beta \frac{M}{n} \rightarrow c,$$

that is  $M \approx cn^{1-\beta}$ , where  $\beta \in (0, 1)$ . The boundary case  $\beta = 1$  corresponds to  $M$  finite as  $n \rightarrow \infty$ .

**Theorem 3** *Let the infeasible realised variance*

$$RV^* = \sum_{0 < t_i \leq T} (\Delta X_{t_i} - \Delta J_{t_i})^2, \quad (21)$$

and let  $MV_M^{(K)} = MV_M^{(2,K)}$  be given as in (16). Assume that  $\sigma_t$  is a continuous semimartingale. Then, under the Assumptions 1-4 given in the Appendix,

$$\begin{aligned} n^{1/2} \left( MV_M^{(K)} - RV^* \right) &= K(Tc)^{\frac{K-1}{K}} n^{\frac{1}{2}-\beta\frac{K-1}{K}} \sum_{0 < t \leq T} \sigma_t^{2\frac{K-1}{K}} \Delta J_t^{2/K} - n^{-\beta+\frac{1}{2}\frac{K-1}{K}} c^{-1} [\sigma, \sigma]_T \\ &\quad + O_p(n^{\frac{1}{2}-\frac{3}{2}\beta} + n^{\frac{1}{2}(\beta-1)}) + o_p(n^{\frac{1}{2}-\beta\frac{K-1}{K}}). \end{aligned}$$

The term  $o_p(n^{\frac{1}{2}-\beta\frac{K-1}{K}})$  only appears when there are jumps.

**Proof.** Given in the Appendix.

In other words: for multipower variation of order  $K \geq 3$ , there is a range of  $\beta$  for which the integrated volatility ( $IV$ ) is estimated efficiently. The asymptotically equivalence of  $MV_M^{(K)}$  and  $RV^*$  is remarkable and theoretically extremely helpful. On the other hand, bipower variation is a boundary case which is consistent, but converges to  $IV$  only at rate  $n^{-\beta/2}$ . Hence bipower is not quite asymptotically equivalent to  $RV^*$ . This result is due to the presence of jumps.

This asymptotic equivalence result immediately implies a number of properties.

**Theorem 4** *Assume the condition of Theorem 3. Also assume that  $K \geq 3$  and*

$$\frac{1}{2} \frac{K}{K-1} < \beta < 1. \quad (22)$$

Then

$$n^{1/2} \left( MV_M^{(K)} - \int_0^T \sigma_t^2 dt \right) \xrightarrow{\mathcal{L}_s} N \left( 0, 2T \int_0^T \sigma_t^4 dt \right), \quad (23)$$

and

$$n^{1/2} \left( \left( RV - MV_M^{(K)} \right) - \sum_{0 < t \leq T} |\Delta J_t|^2 \right) \xrightarrow{\mathcal{L}_s} N \left( 0, 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2 \right). \quad (24)$$

**Proof.** The result (23) follows from Theorem 3. The second result (24) follows from (23) and from (1) in the Introduction.

A key practical feature of this result is that when there are no jumps

$$J_M = \frac{n^{1/2} \left( RV - MV_M^{(2,K)} \right)}{\sqrt{T \left( M\theta_M^{(2,K)} - 2 \right) \int_0^T \sigma_s^4 ds}} \xrightarrow{\mathcal{L}_s} N(0, 1), \quad (25)$$

while when there are jumps the average value of  $J_M$  should be

$$J_M \simeq \frac{n^{1/2} \sum_{0 < t \leq T} |\Delta J_t|^2}{\sqrt{T \left( M\theta_M^{(2,K)} - 2 \right) \int_0^T \sigma_s^4 ds}}.$$

Hence the power of the test should increase with the size of  $JV$  and  $n$  so long as  $K \geq 3$  and  $\frac{K}{K-1} < 2\beta$ .

**Remark 6** (*Optimal choice of  $\beta$  when estimating  $IV$* ). When both jumps and volatility of volatility is present, it follows from Theorem 3 that the jumps part will always dominate the  $[\sigma, \sigma]_T$  part, by a factor of order  $O_p(n^{\frac{\beta}{K}})$ . There is thus no trade-off between the two biases. The bias due to jump can, however, be traded off against the  $O_p(n^{\frac{1}{2}(\beta-1)})$  stochastic term. In view of Theorem 7, this term does not vanish. A bias-variance trade-off to minimize the error in  $MV_M^{(K)} - RV^*$  will thus seek to set these two orders to be equal. This yields the optimal choice of  $\beta$

$$\beta^* = \frac{2K}{3K-2}, \quad (26)$$

which means that as  $K$  increases so  $\beta^*$  converges from above to  $2/3$ . This choice of  $\beta$  satisfies condition (22) in Theorem 4. For tripower,  $\beta^* = 6/7$ . We will use this combination in our Monte Carlo experiments reported in Section 5.

**Remark 7** If  $\beta = 1/2$  and there are no jumps then Theorem 3 implies

$$n^{1/2} \left( MV_M^{(K)} - RV \right) = \frac{K-1}{3} c^{-1} [\sigma, \sigma]_T + O_p(n^{-1/4}), \quad (27)$$

an estimator of the quadratic variation of volatility,  $[\sigma, \sigma]_T$ . This is the choice of  $\beta$  which minimizes the estimation error for  $[\sigma, \sigma]_T$ . The error is conjectured to be mixed normal, cf. the results in Theorem 7 in Appendix A. At the same time, the case  $\beta = 1/2$  also shows the danger of carrying out jump tests with large  $M$  compared to  $n$ . When  $\beta > 1/2$  this bias disappears. Overall this theory suggests that if one used a jump robust blocks approach then we may be able to use it to estimate  $[\sigma, \sigma]_T$ . However, the treatment of this is sufficiently subtle that we have decided to delay its development to a follow up paper.

### 3.6 Optimality for jump estimation of $RV - MV_M^{(2,K)}$

In addition to estimating  $IV$  efficiently, we also obtain an efficient estimate of  $JV$ .

**Theorem 5** Assume that  $\widehat{JV}_n$  is an estimator of  $JV$ , and that  $n^{1/2}(\widehat{JV}_n - JV) \xrightarrow{\mathcal{L}^s} N(0, \gamma^2)$ , where  $\gamma^2$  is measurable with respect to the underlying filtration. Also assume that  $J_t$  is compound Poisson, independent of  $X_t^c$ , with jump sizes that are independent of the jump times and of each other, and that the distributions of jumps are absolutely continuous with respect to Lebesgue measure.<sup>3</sup> Then, under Assumptions 1-2 of Appendix B.1,

$$\gamma^2 \geq 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2 \quad (28)$$

with probability one.

**Proof.** Given in the Appendix.

### 3.7 Reengineering $MV_M^{(2,K)}$ to be robust to noise

**Remark 8** (Strong approximation under microstructure.) In the case where there is microstructure in the data, convergence rates for the estimation of  $IV$  are reduced from  $n^{-1/2}$  to  $n^{-\alpha}$ , where  $\alpha \leq 1/4$ .<sup>4</sup> The value of  $\alpha$  depends on the specific estimator. Now define  $MV_M^{(K)}$  as before, but with  $RV$  replaced by a microstructure-robust estimator  $\widehat{RV}$ . Let  $\widehat{RV}^*$  define the corresponding estimator based on a process without jumps. Following the development in Section A.1.2, along with the proof in Appendix B.4, we obtain

$$\begin{aligned} n^\alpha \left( MV_M^{(K)} - \widehat{RV}^* \right) &= K(Tc)^{\frac{K-1}{K}} n^{\alpha - \beta \frac{K-1}{K}} \sum_{0 < t \leq T} \sigma_t^{2 \frac{K-1}{K}} \Delta J_t^{2/K} - n^{-\beta + \alpha} \frac{K-1}{3} c^{-1} [\sigma, \sigma]_T \\ &\quad + n^{-\alpha + \beta} b_1 \\ &\quad + O_p(n^{-\alpha + \frac{1}{2}\beta} + n^{\alpha - \frac{3}{2}\beta}) + o_p(n^{\frac{1}{2} - \beta \frac{K-1}{K}}), \end{aligned}$$

where  $b_1$  is given by equation (A.12) below. The term  $o_p(n^{\alpha - \beta \frac{K-1}{K}})$  only appears when there are jumps.

<sup>3</sup>The assumptions are made to invoke results on superefficiency, see LeCam (1953) as well as Chapter 6 of Lehmann (1983). There is no minimal set of conditions, and for other relevant results on superefficiency, see Bahadur (1964) and Bahadur (1980), as well as the literature on Hájek-LeCam convolution, as in Hájek (1969). There are circumstances where the bound (28) can be improved on, in particular when the jump sizes are known in advance.

<sup>4</sup>Two-scales realised volatility (Zhang, Mykland, and Ait-Sahalia (2005)) has  $\alpha = 1/6$ . The multi-scale, realised kernel, pre-averaging and quasi-likelihood estimators have  $\alpha = 1/4$  (see, respectively, Zhang (2006), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), Jacod, Li, Mykland, Podolskij, and Vetter (2009), and Xiu (2010)).

A second bias term,  $b_1$ , has thus turned up. This term is eliminated in the no-microstructure case by being very accurate about standardization of  $MV$  (see Appendix B.4 for details of the subtlety). In the microstructure case, however, since  $\alpha \leq 1/4$ , there is no pair  $(\alpha, \beta)$  that makes all the bias terms vanish.

By linear combination with different choices of  $c$  (but fixed  $\beta$ ), all the bias terms can be eliminated to relevant order. The most important term to remove is  $n^{-\alpha+\beta}b_1$ , which is new to the microstructure case, and we here show how to remove this. For fixed  $K$  and  $\beta$ , let  $c$  take on two different values  $c_1$  and  $c_2$ . We can define corresponding multipower variations  $MV_M^{(K,c_1)}$  and  $MV_M^{(K,c_2)}$ . Define a “two scale” estimator

$$MV_M^{(K,c_1,c_2)} = (c_1 - c_2)^{-1}(c_1 MV_M^{(K,c_1)} - c_2 MV_M^{(K,c_2)}).$$

Again following Section A.1.2, we obtain

$$\begin{aligned} n^\alpha \left( MV_M^{(K,c_1,c_2)} - \widehat{RV}^* \right) &= KT^{\frac{K-1}{K}} \frac{c_1^{\frac{-1}{K}} - c_2^{\frac{-1}{K}}}{c_1 - c_2} n^{\alpha-\beta\frac{K-1}{K}} \sum_{0 < t \leq T} \sigma_t^{2\frac{K-1}{K}} \Delta J_t^{2/K} \\ &\quad + n^{-\beta+\alpha} \frac{K-1}{3} \frac{1}{c_1 c_2} [\sigma, \sigma]_T \\ &\quad + O_p(n^{-\alpha+\frac{1}{2}\beta} + n^{\alpha-\frac{3}{2}\beta}) + o_p(n^{\frac{1}{2}-\beta\frac{K-1}{K}}). \end{aligned}$$

With this expression, the orders become analogous to those in Theorem 3, with  $\alpha$  replacing  $1/2$ . In particular, the other bias terms disappear to order  $o_p(1)$  provided

$$\alpha \frac{K}{K-1} < \beta < 1, \tag{29}$$

generalising (22). Notice that as  $\alpha \leq 1/4$ , the permissible range for  $\beta$  is much bigger. The reason for this is that the bias terms scale with  $n$  as  $n^{\alpha-\beta}$  so low values of  $\alpha$  allow  $\beta$  to be lower.

A version of Theorem 4 for microstructure thus follows.

**Theorem 6** *Assume the setting of this section, that  $K \geq 2$  (PER SURELY THIS SHOULD BE  $K \geq 3$ ) and that (29) holds. Then*

$$n^\alpha \left( MV_M^{(K)} - \int_0^T \sigma_t^2 dt \right) \xrightarrow{\mathcal{L}^s} N \left( 0, AVAR(\widehat{RV}^*) \right), \tag{30}$$

where  $AVAR(\widehat{RV}^*)$  is the asymptotic variance of  $n^\alpha \left( \widehat{RV}^* - \int_0^T \sigma_t^2 dt \right)$ , and

$$n^{1/2} \left\{ \left( \widehat{RV} - MV_M^{(K)} \right) - \sum_{0 < t \leq T} |\Delta J_t|^2 \right\} \xrightarrow{\mathcal{L}^s} N \left( 0, 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2 \right). \tag{31}$$

The optimal  $\beta$  is for estimating  $IV$  and  $JV$  is, similarly to (26), given by

$$\beta^* = 2\alpha \frac{2K}{3K-2}. \tag{32}$$

If  $K = 3$  and  $\alpha = 1/4$ , then  $\beta^* = 3/7$ . Hence, when there is microstructure it makes sense to have much bigger blocks due to the slow rate of convergence.

Also note that when there are no jumps, one can similarly remove the bias term due to volatility of volatility. The resulting jump test converges at rate  $O_p(n^{-3\alpha/2})$ , cf. the development in Appendix B.4. This, obviously, also applies to the no-microstructure case, with  $\alpha = 1/2$ .

PER: SPELL THIS OUT

### 3.8 Estimating the asymptotic variance

In order to carry out inference robustly to jumps it is helpful to be able to consistently estimate the “integrated quarticity”  $IQ = \int_0^T \sigma_t^4 dt$  and the “squared jump times vol”  $SJV = \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2$ . Barndorff-Nielsen and Shephard (2006) solved the former problem using multipower variation. We can improve the efficiency of their approach by using a blocked version to more efficiently estimate  $IQ$ . In this subsection our focus will be on estimating  $\sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2$ .

Using a blocked approach this problem is straightforward. Rearranging terms

$$RV - BV_M = \sum_{i=1}^{n_M} \left[ RV_i - \frac{M}{k_M} (RV_i RV_{i-1})^{1/2} \right] \xrightarrow{p} \sum_{0 < t \leq T} |\Delta J_t|^2,$$

as  $n \rightarrow \infty$  if  $M$  is fixed or very slowly increasing. We now extend this using multipower ideas, so now

$$\sum_{i=1}^{n_M} \left\{ RV_i - \frac{M}{k_M} (RV_i RV_{i-1})^{1/2} \right\} \left\{ \frac{1}{k_M} (RV_{i-2} RV_{i-3})^{1/2} \right\} \xrightarrow{p} \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2.$$

Hence all our new distribution theory is feasible. Veraart (2010) provided an alternative estimator of  $SJV$ , which is somewhat more complicated.

## 4 Finite sample improvement

### 4.1 Transformation

The jump test (15) can be transformed using the delta method so we look at

$$J_M = \frac{n^{1/2} (\log RV - \log BV_M)}{\sqrt{T(M\theta_M - 2) \frac{\widehat{IQ}}{(BV_M)^2}}} \xrightarrow{\mathcal{L}_s} N(0, 1). \quad (33)$$

This log-transform was used by Barndorff-Nielsen and Shephard (2002) and Barndorff-Nielsen and Shephard (2006) to produce better size properties of their tests and confidence intervals. This was improved upon by Barndorff-Nielsen and Shephard (2005) who noted that by Jensen’s inequality the volatility of volatility ratio

$$R = \frac{\int_0^T \sigma_s^4 ds}{\left( \int_0^T \sigma_s^2 ds \right)^2} \geq \frac{1}{T},$$

so it makes sense to replace (33) by

$$J_M = \frac{n^{1/2} (\log RV - \log BV_M)}{\sqrt{T(M\theta_M - 2) \max\left(\frac{\widehat{IQ}}{(BV_M)^2}, \frac{1}{T}\right)}} \xrightarrow{\mathcal{L}_s} N(0, 1). \quad (34)$$

In practice we advocate using the exact constant (12) in empirical and simulation work, which produces

$$J_M = \frac{\log RV - \log BV_M}{\sqrt{Td \max\left(\frac{\widehat{IQ}}{(BV_M)^2}, \frac{1}{T}\right)}} \xrightarrow{\mathcal{L}_s} N(0, 1).$$

We will see this makes quite a large difference if  $M$  is large. Of course further refinements may be possible, such as the use of bootstrapping suggested by Goncalves and Meddahi (2004).

## 4.2 Time-change and diurnality

Quite a substantial amount of the variation of the volatility process within a day is caused by the average diurnal pattern playing out. This is one of the causes of  $R > 1$  in practice. Can we perform a time-change to bring it down closer to one, making the test more efficient?

Consider the absolutely continuous time-change, and the  $T = 1$  case for ease of exposition.

$$h_t^2 > 0, \quad \int_0^1 h_t^2 dt = 1, \quad H_t = \int_0^t h_s^2 ds$$

then  $X_{H_t}$  is a Brownian semimartingale with spot variance at time  $t$  of  $\tilde{\sigma}_t^2 = \sigma_{H_t}^2 h_t^2$ , while  $[X \circ H]_1 = [X]_1$ . Hence if we know the path  $\sigma_t$  then we can design the time-change  $h_t$  so that  $\tilde{\sigma}_t^2$  is time-invariant and so  $R = 1$ . Of course in general this is impossible, but we can time-change to take out the expected diurnal pattern. We can do this safely in our model, for if the expected diurnality is modelled incorrectly then our inference procedure will still be valid, for the time-change process is always a Brownian semimartingale whatever absolutely continuous time-change we use.

Consider the diurnal model  $\sigma_t^2 = \psi_t d_t$  then  $\tilde{\sigma}_t^2 = \psi_{H_t}^2 d_{H_t} h_t^2 = \psi_{H_t}^2$  if we design  $h_t$  so that  $h_t^{-2} = d \circ \int_0^t h_u^2 du$ .

How do we implement this idea? A natural method to control the diurnal effect is to estimate the diurnal shape non-parametrically using block Bipower-type statistics. Define the  $n - 2M + 1$  overlapping unnormalised ‘‘Bipower blocks’’ as

$$\widetilde{BV}_{M,i} = \sqrt{RV_{M,i}RV_{M,i+M}}, \quad i = 1, 2, \dots, n - 2M + 1.$$

The diurnal pattern can be estimated using  $R$  days of historical data,

$$\hat{\delta}_i^2 = R^{-1} \sum_{t=i-R}^{i-1} \left[ \frac{1}{u_i - l_i} \sum_{i=l_i}^{u_i} \widetilde{BV}_{M,i,t} \right]$$

where  $u_i = \min(n - 2M + 1, i)$  and  $l_i = \max(1, i - 2M + 1)$ . This estimator uses all  $M$ -sample blocks which contain return  $i$  to estimate the diurnal pattern in the data. Except for period near the beginning or end of the day, the diurnal effect is estimated by averaging the  $2M$  Bipower blocks that contain return  $i$ . Blocks at the beginning or end use between 1 and  $2M - 1$  blocks.

## 5 Monte Carlo study of efficient bipower variation

For purposes of simulation, we take  $T = 1$  and always use 10,000 replications. Throughout when constructing estimators and  $t$  tests we use versions which are exactly unbiased in the Brownian motion case, rather than the simpler asymptotic equivalent versions. This does make a great deal of difference in terms of Monte Carlo performance.

### 5.1 Single factor SV model

To assess the usefulness of the asymptotic results we have conducted two Monte Carlo experiments. The first, and simplest, uses a single factor SV model. It has the following log-normal structure

$$\begin{aligned} dX_t &= \mu dt + \exp(\beta_0 + \beta_1 v_t) dW_{p,t} + dJ_t, \\ dv_t &= \alpha_v v_t dt + dW_{v,t}. \end{aligned}$$

Here  $W_p$  and  $W_v$  are correlated standard Brownian motions. This process was implemented using the values given in Table 3, which follow from the experiments reported in Huang and Tauchen (2005). This was designed to have relatively fast mean reversion with a half-life of 10 days, and so there is some variation in the volatility during each day.

Parameter	$\mu$	$\beta_0$	$\beta_1$	$\alpha_v$	$\rho$
Value	0.030	0	0.125	-0.100	-0.620

Table 3: Parameters which index the single factor SV simulation model

### 5.1.1 Finite activity jumps

Throughout our experiments we simulate 50,000 days. Within each day the high frequency data was simulated using 23,400 steps per day, and then “1-minute” returns were sampled by skipping 60 steps. The result is  $n = 390$ . Jumps, when present, have a random location uniformly distributed on  $1, \dots, 23,400$  and were a deterministic percentage of the day’s integrated variance (e.g. a jump, when squared, represents 10% of IV, or  $JV/IV = 0.1$ ). Throughout our simulations only a single jump occurred on each day.

### 5.1.2 One factor SV model with diurnal effects

The simulated model was augmented to include a diurnal effect  $\delta_t^2$  which produces variances which differ deterministically by a factor of 3 between the middle of the day and the beginning or end.

$$\begin{aligned} dX_t &= \mu dt + \exp(\beta_0 + \beta_1 v_t) \sqrt{\delta_t} dW_{p,t} + dJ_t, \\ dv_t &= \alpha_v v_t dt + dW_{v,t}, \\ \delta_t &= 0.6 + 4.8(t - 0.5)^2, \end{aligned}$$

where the average effect  $\int_0^1 \delta_t^2 dt = 1$  and the start and the end of the day  $\delta_0^2 = \delta_1^2 = 0.6 + 4.8 \times 0.5^2 = 1.8 = 3\delta_{0.5}^2$ .

KEVIN: CAN WE PROVIDE A REFERENCE TO TRY TO BACK UP THIS DGP CHOICE?

## 5.2 Bias, variance, size and power of jump tests

Size was computed by setting the jump size to 0. The test statistic was

$$J_M = \frac{RV - BV_M}{\sqrt{d_{n,M,K} \widehat{IQ}}}$$

where  $d_{n,M,K}$  is given by (17). We estimate integrated quarticity using the unbiased blocked tripower quarticity estimator (16) with  $r = 4$ ,  $K = 3$  and the same value of  $M$  as used for  $BV_M$ . The logged version is

$$J_{M,\log} = \frac{\log RV - \log BV_M}{\sqrt{d_{n,M,K} \max\left(\frac{\widehat{IQ}}{(BV_M)^2}, 1\right)}}.$$

All tests were conducted using 1-sided upper tailed critical values. We also carried out the Monte Carlo using a unblocked tripower variation statistic to estimate  $IQ$  and there was little difference in the corresponding size.

Table 4 reports the empirical size both the standard 1-factor model and the model that includes deterministic diurnal effects. We only report the size of a 5% one-sided test since both 10% and 1% sized tests performed analogously.

In the Table we use the following important notation.

- **FS.** This reports results using the finite sample variance for a standard Brownian motion, using equation (B.17). The results using the large sample values for  $M\theta_M$  were poor and

$M$	No Diurnality				Diurnality			
	FS		FS-DC		FS		FS-DC	
	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$
1	6.4	5.5	6.5	5.6	7.3	6.5	6.7	5.7
2	6.1	5.5	6.2	5.6	8.8	8.3	6.7	6.0
3	6.4	5.9	6.4	5.8	10.9	10.5	6.1	5.6
5	7.0	6.6	6.4	6.0	14.7	14.2	6.4	5.8
10	6.7	6.4	6.9	6.5	33.6	33.0	7.0	6.8
15	7.3	7.0	6.9	6.5	56.1	55.4	7.5	7.1
30	8.1	7.6	8.3	7.9	90.2	89.5	8.6	8.1

Table 4: Empirical size of SV1F with and without (left panel) deterministic diurnality using one minute simulated returns. Values under  $J_M$  correspond to the standard version of  $BV_M$  and  $J_{M,\log}$  correspond to the log version of the jump test statistic. Values under FS use the finite sample version in (B.17). Values below FS-DC use finite sample variance and an estimated diurnality correction.

so we do not recommend their use and so we are not reporting them here. The distortions to the size of the test statistics computed using the large sample  $M\theta_M$  is clearly caused by the overconfidence in the precision, and block sizes larger than 5 lead to unacceptable large deviations from 1.

- **FS-DC.** This results using the finite sample variance and applying the diurnal correction developed in Section 4.2. Throughout this uses the previous  $R = 100$  days of simulated data. The test statistics were computed by first constructing a modified set of returns,  $\Delta X_i/\hat{\delta}_i$ , where  $\hat{\delta}_i$  is the diurnal correction for the relevant block. This is exactly the same as having standardised the local realised variance estimator by the estimated diurnal feature of the data. We then compute  $RV$ ,  $BV_M$  and  $TPQ$  using the transformed data.

In the non-diurnal simulation results, using the large sample variance of the jump test leads to large distortions when the block size is moderately large. The finite sample results show little size distortion for any block size. When the deterministic diurnality is introduced a large distortion, driven by the difference in volatility across blocks, appears which increases with the block length.

Table 5 reports the mean and variance of the test statistics under the same treatments. The means are very close to 0 when no diurnal component is present. When the diurnality is introduced the mean of the test statistic grows monotonically with the block size and when the block size is larger than 10 the average test statistic would reject using a 5% upper-tailed test.

Table 6 reports the empirical size-adjusted power of the of the test statistics when using a 5% test. The power peaks with a block size of 5 which is consistent with the minimum of finite sample variance for  $n = 390$ . The size-adjusted power falls off for large blocks due to both the increase in variance, and in the simulations with the diurnal component, the substantial size distortions.

Table 4 contains size results for the jump test constructed using the finite sample variance and the diurnality correction (FS-DC). When there is no diurnality present the modification has no effect on size. When the diurnality is present the transformation produces an empirical size which is virtually identical to the size of the FS estimator in the simulation without the diurnality. Table 5 shows that the improvement comes from the improvement in the mean of the test statistic which is only slightly worse than the FS estimator in the non-diurnal case. The diurnal correction has no effect on the variance of the test statistic. Finally, Table 6 shows that the use of the diurnal correction has no effect on size-adjusted power when the data are not diurnal, and improves the size-adjusted power when the data contains a diurnal component.

Mean								
$M$	No Diurnality				Diurnality			
	FS		FS-DC		FS		FS-DC	
	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$
1	0.05	0.02	0.05	0.03	0.09	0.07	0.05	0.03
2	0.03	0.02	0.03	0.01	0.17	0.15	0.03	0.01
3	0.03	0.01	0.03	0.01	0.29	0.27	0.02	0.01
5	0.04	0.03	0.01	0.00	0.50	0.49	0.05	0.03
10	0.03	0.02	0.04	0.03	1.25	1.22	0.04	0.03
15	0.04	0.03	0.03	0.02	2.01	1.96	0.05	0.04
30	0.06	0.04	0.08	0.07	3.82	3.61	0.10	0.09

Variance								
$M$	No Diurnality				Diurnality			
	FS		FS-DC		FS		FS-DC	
	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$
1	1.00	0.95	1.03	0.98	1.04	1.01	1.02	0.96
2	0.99	0.96	1.00	0.96	1.09	1.06	1.04	1.01
3	1.01	0.98	1.01	0.98	1.14	1.11	0.99	0.96
5	1.05	1.02	1.01	0.99	1.24	1.20	0.99	0.97
10	1.02	1.00	1.03	1.01	1.68	1.58	1.03	1.01
15	1.08	1.06	1.06	1.04	2.17	1.99	1.07	1.05
30	1.14	1.11	1.17	1.14	3.42	2.85	1.16	1.13

Table 5: Mean and variance of the test statistic  $J_M$  and  $J_{M,\log}$  in the Monte Carlo. These values should be around 0 and 1 if the asymptotic distribution is a good approximation. The left panel contains results from a simulation without a diurnal component. Values under FS use the finite sample version in (12). Values below FS-DC use finite sample variance and an estimated diurnality correction.

10% Jump Size								
$M$	No Diurnality				Diurnality			
	FS		FS-DC		FS		FS-DC	
	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$
1	41.4	41.9	41.4	41.9	37.7	37.7	48.2	48.7
2	50.6	50.7	50.6	50.7	46.9	46.9	56.9	57.0
3	54.9	55.4	54.9	55.4	50.8	50.7	60.3	60.5
5	57.9	58.1	57.9	58.1	52.3	52.5	63.4	63.6
10	52.9	53.0	52.9	53.0	41.1	41.3	59.1	59.3
15	44.1	44.2	44.1	44.2	31.4	31.3	51.7	51.9
30	26.4	26.2	26.4	26.2	15.7	15.6	29.0	29.0

20% Jump Size								
$M$	No Diurnality				Diurnality			
	FS		FS-DC		FS		FS-DC	
	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$
1	85.8	86.0	85.8	86.0	82.3	82.2	83.5	83.9
2	93.4	93.6	93.4	93.6	91.0	91.0	90.1	90.2
3	95.9	96.1	95.9	96.1	93.4	93.4	92.3	92.4
5	96.8	96.9	96.8	96.9	93.9	93.9	93.4	93.5
10	95.8	95.8	95.8	95.8	89.6	89.7	92.7	92.8
15	91.8	92.0	91.8	92.0	80.5	80.5	90.0	90.1
30	68.5	68.4	68.5	68.4	43.5	43.2	75.4	75.3

Table 6: Empirical size-adjusted power for jump sizes representing 10% (top panel) and 20% (bottom panel) of the integrated variance. The left 4 columns correspond to the SV1F model and the right columns add a deterministic diurnal effect. FS indicates that the finite sample variance was used and FS-DC indicates that both the finite sample variance and the diurnal correction were used to compute the statistic.

$n$	$M$	No Diurnality				Diurnality			
		FS		FS-DC		FS		FS-DC	
		$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$	$J_M$	$J_{M,\log}$
Empirical size									
780	3	6.1	5.8	6.1	5.8	8.7	8.3	6.1	5.8
1,560	3	5.5	5.3	5.5	5.3	7.4	7.1	5.5	5.3
1,560	4	5.6	5.5	5.6	5.5	7.9	7.7	5.6	5.5
Power: 10% Jump Size									
780	3	87.3	87.4	87.3	87.4	85.2	85.2	85.2	85.3
1,560	3	99.6	99.7	99.6	99.7	99.0	99.0	97.2	97.3
1,560	4	99.8	99.8	99.8	99.8	99.3	99.3	98.0	98.0

Table 7: Impact of increasing  $n$  and  $M$  together. Empirical size and size-adjusted power for jump sizes representing 10% of the integrated variance. The left 4 columns correspond to the SV1F model and the right columns add a deterministic diurnal effect. FS indicates that the finite sample variance was used and FS-DC indicates that both the finite sample variance and the diurnal correction were used to compute the statistic.

### 5.3 $MV_M^{(2,K)}$ for $M$ increasing with $n$

We will now repeat the above experiments but allow  $n$  to vary, driving some movements in  $M$ . Throughout this subsection we will take  $K = 3$  and the corresponding optimal  $\beta = 6/7$  following Remark 6, where

$$M = cn^{1-\beta}.$$

As  $cn^{1-\beta}$  will not be an integer we take it to be the nearest integer but requiring it also to be greater than or equal to one. We do not have a theory driven value for  $c$ , so we take it from the above Monte Carlos which suggest for  $n = 390$  then  $M = 3$  provides satisfactory results. This suggests taking  $c = M/n^{1/7} \simeq 1.28$  as we vary  $n$ .

In our Monte Carlo we will take double  $n$  to 780 and double it again to 1,560. The first change leaves  $M$  unaltered at 3, the second change drives  $M$  up to 4. Throughout we use the finite  $M$  form of the t-statistic in implementing the test.

The results are in Table 7. They suggest an improving size performance as  $n$  and  $M$  increase together, while the power naturally increases very significantly with an increase in the sample size. Hence these results are encouraging for the blocking and its asymptotic analysis. The Table also shows the results from having  $M$  is fixed as  $n$  increases in the  $n = 1,560$  and  $M = 3$  case (the theory suggests  $M$  should have nudged up to 4 due to the increase in  $n$ ) The results suggest a loss in power in that case, although the differences are mild.

### 5.4 Confidence intervals for $JV$

A new result in this paper is that we can construct confidence intervals on  $JV$  using multipower variation in the presence of jumps. Recall that

$$n^{1/2} \left\{ \left( RV - MV_M^{(K)} \right) - \sum_{0 < t \leq T} |\Delta J_t|^2 \right\} \xrightarrow{\mathcal{L}^s} N \left( 0, 4T \sum_{0 < t \leq T} |\Delta J_t|^2 \sigma_t^2 \right), \quad (35)$$

which requires that  $K \geq 3$ . Here we will take  $M = cn^{1-\beta}$  and  $K = 3$ ,  $c = 1.28$ .

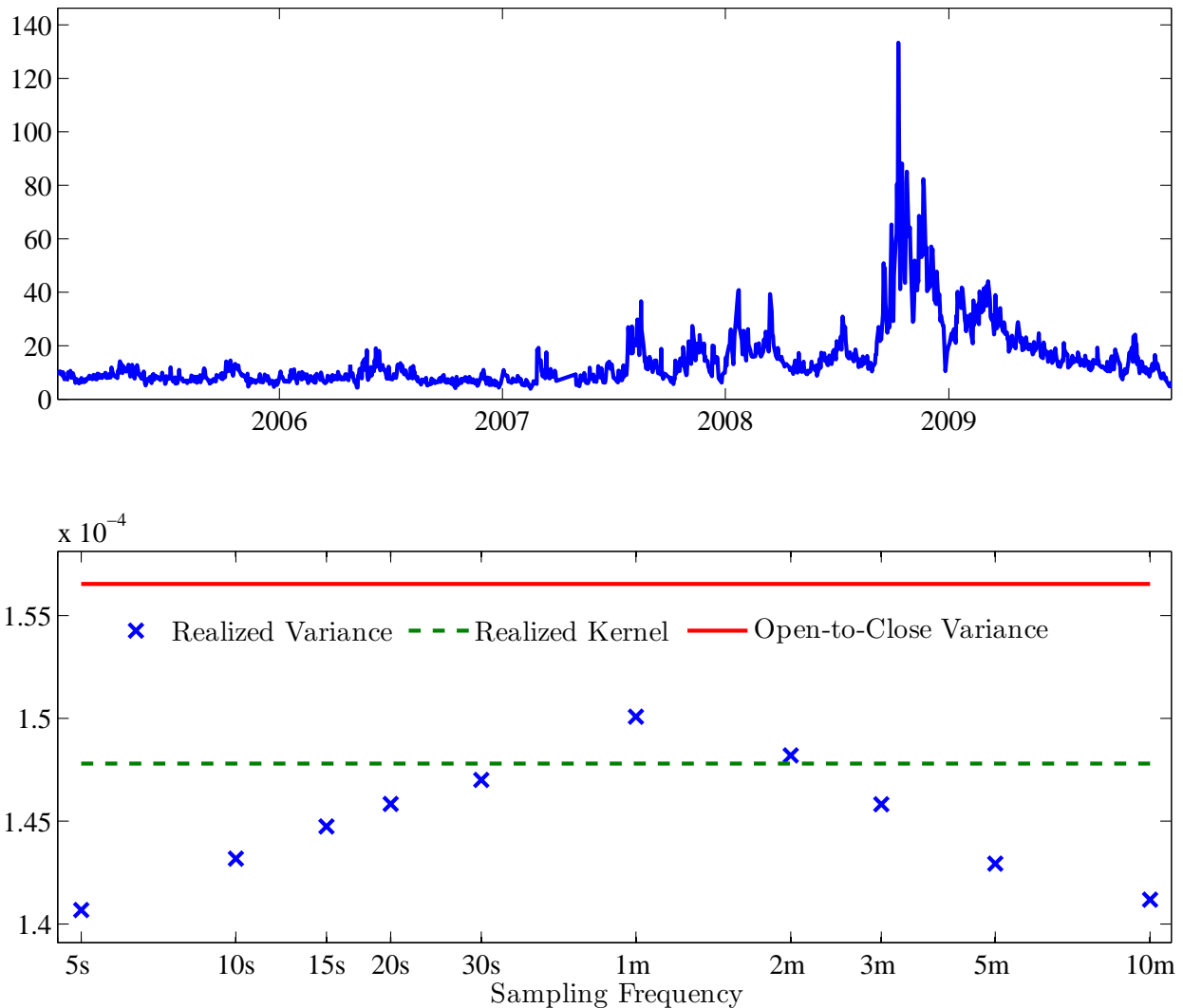


Figure 1: The top panel contains the annualized realised volatility for the S&P500 SPDR computed using one-minute returns from January 2005 to December 2009. The bottom panel contains the volatility signature plot for realized variance using sample frequencies from 5-seconds to 10-minutes, the average value of realized kernels computed daily, and the variance of daily open-to-close returns.

## 5.5 Confidence intervals for $QV$

# 6 Empirical results

## 6.1 Database

Our application will be based around trade data for the SPDR S&P 500 ETF (SPY), which is an exchange traded fund (ETF) that tracks the S&P 500 Index. This is the most liquid equity in U.S. markets and typically has a spread of .01 which represents approximately 0.01% of the price of the instrument.

The small spread allows for frequent sampling without substantial market microstructure noise, and so we employ returns computed every 60 seconds. As a result  $n = 390$ .

Our database contains transactions from January 3, 2005 until December 31, 2009. We will not compute statistics on days which are short (i.e. open or close early) or days where there are sequences of no trades for more than 60 consecutive seconds. These rules leave us with 1,226 full days of high frequency data. Prices were cleaned to remove outliers in a similar way to Barndorff-

$M$	$J_M$			$J_{M,\log}$		
	5%	1%	.1%	5%	1%	.1%
1	48.2	31.2	19.9	49.3	33.6	21.4
2	58.0	41.3	24.5	68.3	56.2	43.4
3	59.7	42.7	26.0	74.5	65.0	53.4
5	60.2	44.0	30.0	79.6	72.6	63.1
10	63.1	46.6	30.8	85.4	80.8	74.6
15	68.2	52.2	36.9	88.7	84.4	79.7

Table 8: Empirical size using the blocked bipower jump tests for different levels of blocking  $M$ .

Nielsen, Hansen, Lunde, and Shephard (2009). The top panel of Figure 1 contains a plot of the annualized volatility (%) computed from the daily realised variance estimates, and the bottom panel contains the volatility signature plot for sampling times between 5 seconds and 10 minutes, as well as the average value of daily realized kernels and the open-to-close return variance.<sup>5</sup> Sampling using 1-minute prices does not appear to be affected by market microstructure noise.

## 6.2 Empirical rejection rates

Both the standard jump test statistic and the log-version FS-DC were computed daily with  $n = 390$ . All tests were implemented as one-sided upper-tail tests, rejecting for large values of the test statistic. Throughout we estimate integrated quarticity using TPQ. The empirical rejection rates are given in Table 8. The table contains results from the test statistics applied directly to returns, and show that the rejection rate is monotonically increasing with the block size for both test statistics.

## 6.3 Diurnality Correction

The top panel of figure 2 shows an estimated averaged diurnality in jumps through the day. This figure was computed using the estimator described in the Monte Carlo with two changes:

1. High frequency returns are standardized by the daily  $\sqrt{BV_5}$  prior to estimating the diurnality. This is done to mitigate the effect of changing volatility throughout the sample, and is similar to weighting in a generalised least squares regression.
2. The entire sample was used to compute the diurnal volatility pattern.

The Figure indicates that volatility is substantially higher at the start of the day, even when controlling for the transaction rate, and that volatility rises at the end of the day.

Table 9 contains test statistics computed using returns transformed by an estimate of the diurnal effect. This diurnal correction was estimated using the previous returns from the previous 100 days. This correction removes the increase in the rejection rate as the block size grows, and the rejection rate peaks with  $M = 3$  or 5, consistent with the Monte Carlo study. In both panels the log version of the test rejects substantially more often than the standard version.

## 6.4 Validation

- Compute the time-series average of  $BV_M$  divided by the time-series average of RV. If the bias is important one would expect this to decline sharply as  $M$  increases. Figure 3 shows this ratio for all three versions of the test statistics. The diurnality correction appear to reverse the decline in the average BV as the block size decreases.

---

<sup>5</sup>Realized Kernels computed using all returns following the method detailed in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009).

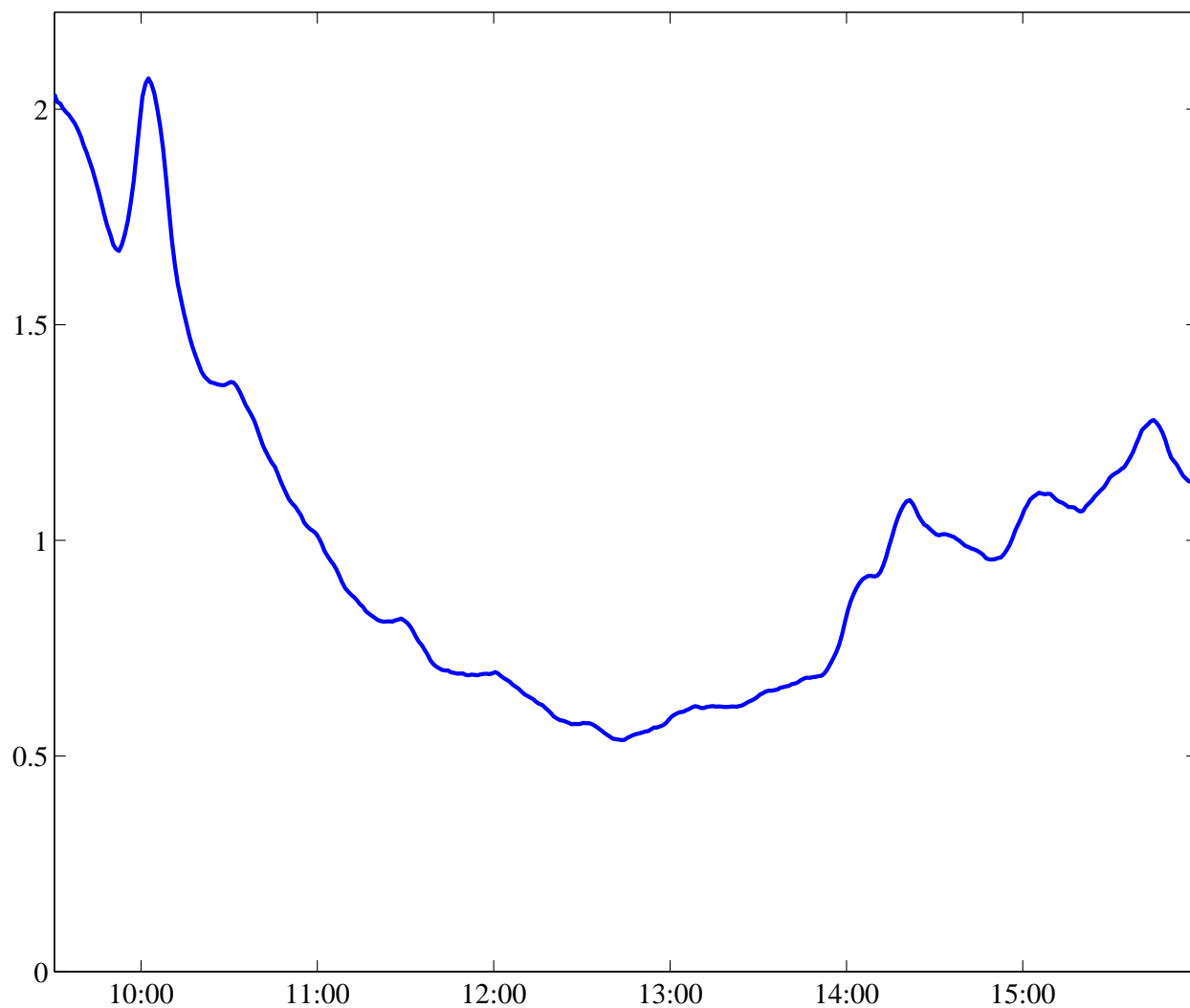


Figure 2: *Full sample estimated dirunal pattern of volatility ( $\hat{\delta}^2$ ) for the S&P500 SPDR.*

- Table 10 contains the percentage of the “bipower blocks” which are 0. In a BSM 0 returns should not occur, although due to market microstructure noise they do arise approximately 10% of the time when sampling using 1 minute returns. Larger blocks have substantially fewer 0s which may be an additional advantage.

$M$	No Correction						Diurnal Correction					
	$J_M$			$J_{M,\log}$			$J_M$			$J_{M,\log}$		
	5%	1%	.1%	5%	1%	.1%	5%	1%	.1%	5%	1%	.1%
1	48.2	31.2	19.9	49.3	33.6	21.4	55.4	38.7	25.4	58.9	42.3	28.4
2	58.0	41.3	24.5	68.3	56.2	43.4	62.9	46.1	31.1	73.6	63.5	51.1
3	59.7	42.7	26.0	74.5	65.0	53.4	63.1	47.1	30.4	78.4	69.1	59.2
5	60.2	44.0	30.0	79.6	72.6	63.1	59.6	44.9	30.7	78.0	72.8	64.3
10	63.1	46.6	30.8	85.4	80.8	74.6	51.3	35.6	23.1	75.4	69.7	63.9
15	68.2	52.2	36.9	88.7	84.4	79.7	44.9	32.7	20.6	70.2	65.1	58.8

Table 9: Empirical size using the blocked bipower jump tests for different levels of blocking  $M$ . The left two panels contains rejection rates for the usual version of the test statistic. The right panels contains rejection rates for the test statistic based on returns with a diurnal correction.

$M$	1	2	3	5	6	10	13	15
Calendar Time	11.7	2.0	0.5	0.0	0.0	0.0	0.0	0.0

Table 10: Percentage of “bipower blocks” which are 0 for each block size.

## 7 Some additional points

### 7.1 Blocked power variation

Barndorff-Nielsen and Shephard (2004) formalised power variation into economics<sup>6</sup>, finding it empirically attractive for low powers of  $r$ . Recall it is of the form

$$PV^{(r)} = n^{r/2-1} \frac{1}{k_r} \sum_{0 < t_i \leq 1} |\Delta X_{t_i}|^r, \quad k_r = \mathbb{E}|U|^r, \quad U \sim N(0, 1).$$

When there are no jumps  $PV^{(r)} \xrightarrow{P} \int_0^1 \sigma_s^r ds$ , integrated power volatility. Barndorff-Nielsen and Shephard (2004) show this is robust to jumps for  $r < 2$  for any jump is divided by  $n^{r/2-1}$ . This compares to bipower, which scales any jump by a  $O_p(n^{-1/2})$  term). The use of power variation type statistics is becoming popular in econometrics, although mostly they are combined with some form of truncation. See, for example, the work of Jacod (2008a).

One of the problems with power variation is that it is inflexible over which power volatility it estimates, so one cannot strip out the jumps and estimate quadratic variation of the continuous component. This lead Barndorff-Nielsen and Shephard (2004) to the development of multipower variation.

We can now generalise the blocking approach to provide another approach to utilising power variation, computing blocks of scaled power variation. Within each block the target is  $\int_{t_{i-1}}^{t_i} \sigma_s^r ds$  so the local power variation statistics we need is

$$PV_i^{(r)} = n^{r/2-1} \left( \frac{1}{k_r} \sum_{\tau_{i-1} < t_j \leq \tau_i} |\Delta X_{t_j}|^r \right).$$

Then to estimate  $IV$  we calculate the blocked power variation

$$PV^{(2,r)} = n_M^{2/r-1} \sum_{i=1}^{n_M} \left\{ PV_i^{(r)} \right\}^{2/r}.$$

<sup>6</sup>Andersen and Bollerslev (1998b) and Andersen and Bollerslev (1997) empirically studied the properties of sums of absolute values of intra-day returns on speculative assets. This was empirically attractive, for using absolute values is less sensitive to possible large movements in high frequency data. However, the approach was abandoned in their subsequent work reported in Andersen and Bollerslev (1998a) and Andersen, Bollerslev, Diebold, and Labys (2001) due to the lack of appropriate theory for the sum of absolute returns as  $n \rightarrow \infty$ , although Shiryaev (1999, pp. 349–350) and Maheswaran and Sims (1993) mention interest in the limit of sums of absolute returns.

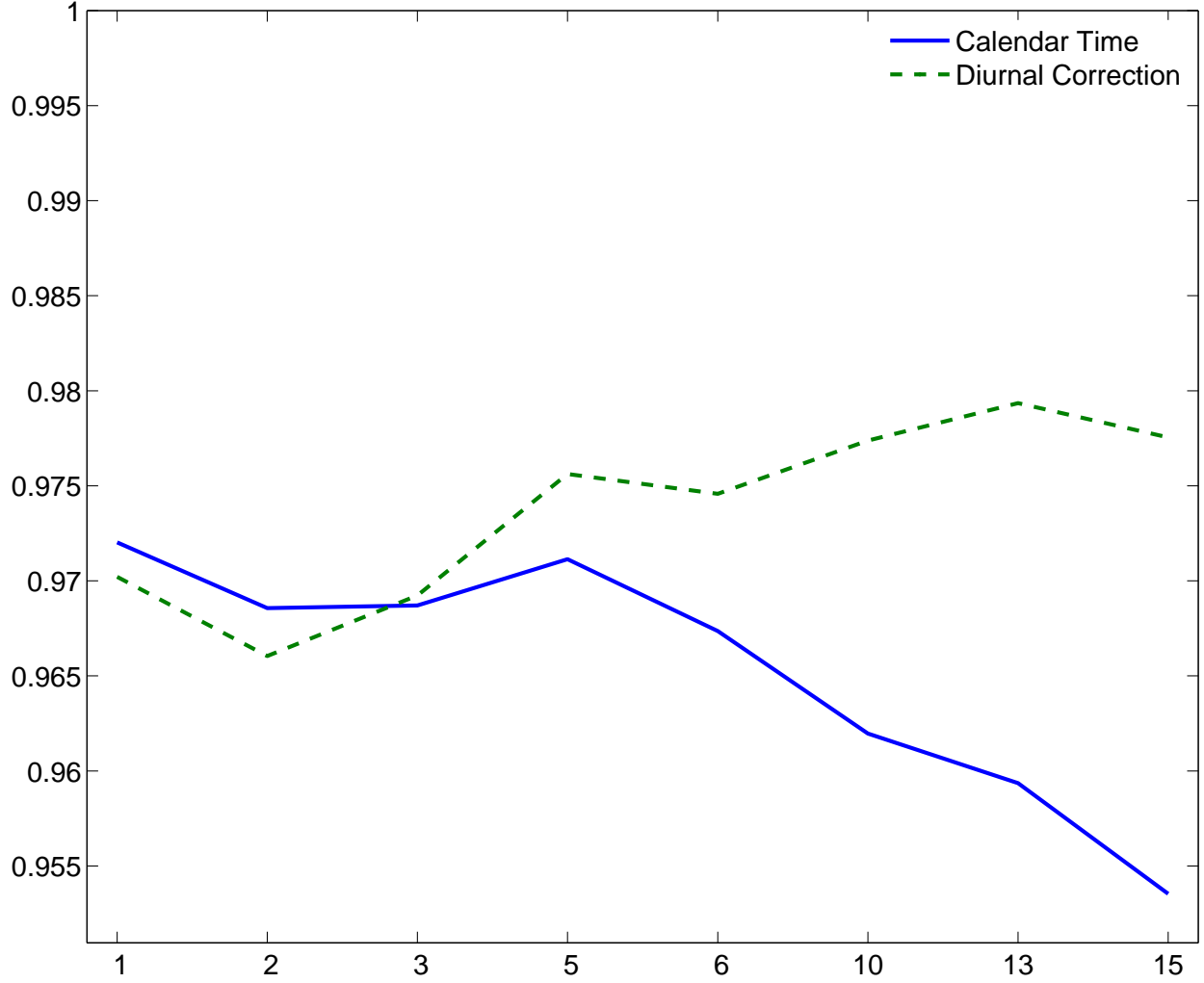


Figure 3: Full sample estimated ratio of average  $BM_M$  to average  $RV$  for prices sampled in calendar time (solid), diurnally corrected returns (dashed) and prices sampled in business time (dashed).

Is this as efficient as blocked bipower variation? The following remark shows the answer to this is no.

**Remark 9** If  $X$  is a scaled Brownian motion  $\sigma W$ , then  $PV_i^{(r)}$  are i.i.d. with

$$PV_i^{(r)} \stackrel{L}{=} \frac{M}{n} \sigma^r v_j, \quad v_j = \left( \frac{1}{k_r} \frac{1}{M} \sum_{j=1}^M |U_j|^r \right),$$

where  $U_j \stackrel{iid}{\sim} N(0,1)$ , so its mean is  $Mn^{-1}\sigma^r = n_M^{-1}\sigma^r$ . Likewise

$$PV^{(2,r)} \stackrel{L}{=} \left( \frac{M}{n} \right)^{1-2/r} \sum_{i=1}^{n_M} \left\{ \frac{M}{n} \sigma^r v_j \right\}^{2/r} = \sigma^2 \left( \frac{1}{n_M} \sum_{i=1}^{n_M} v_j^{2/r} \right).$$

Now as  $M$  increases

$$\sqrt{M} (v_i - 1) \xrightarrow{d} N \left\{ 0, \frac{k_{2r} - k_r^2}{k_r^2} \right\},$$

so by the delta rule

$$\sqrt{M} \left( v_i^{2/r} - 1 \right) \xrightarrow{d} N \left\{ 0, \left( \frac{2}{r} \right)^2 \frac{k_{2r} - k_r^2}{k_r^2} \right\}.$$

Hence the mean of  $PV^{(2,r)}$  is roughly  $\sigma^2$  and the variance is roughly

$$\frac{\sigma^4}{n} \left( \frac{2}{r} \right)^2 \frac{k_{2r} - k_r^2}{k_r^2}.$$

Of course efficiency of realised variance means that

$$\left( \frac{2}{r} \right)^2 \left( \frac{k_{2r}}{k_r^2} - 1 \right) \geq 2$$

with the minimum achieved at  $r = 2$ . Now when  $r = 1$

$$\left( \frac{2}{r} \right)^2 \left( \frac{k_{2r}}{k_r^2} - 1 \right) = 4 (k_1^{-2} - 1) \simeq 2.28$$

hence this scheme is more efficient than bipower variation but less effective than blocked bipower variation or realised variance.

## 7.2 More general forms of multipower variation

Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006) introduced a more abstract class of multipower statistics which take on the form of

$$\sum_{i=2}^{n_M} \prod_{k=0}^{K-1} g_k(\Delta X_{i-k}),$$

where  $g_k$  is an even non-random function (Kinnebrock and Podolskij (2008) study the more general case of where  $g_k$  is not even). A useful blocked version of this is clearly

$$\sum_{i=2}^{n_M} \prod_{k=0}^{K-1} g_k(\sqrt{RV_{i-k}}).$$

We will not study this more general statistic here.

## 8 Conclusion

Jumps are important in financial economics. Recent development of econometric methods to analyse jumps using high frequency data has quantified the importance of jumps. The recent empirical work has suggested they are quite common empirically.

One of the leading methods for detecting jumps is bi- and multi-power power variation. In this paper we have suggested ways of improving the efficiency of these estimators, making them efficient, and increasing the power of jump tests. We also show a path towards the estimation of the volatility of volatility.

## References

- Abramowitz, M. and I. A. Stegun (1970). *Handbook of Mathematical Functions*. New York: Dover Publications Inc.
- Aït-Sahalia, Y. and J. Jacod (2007). Volatility estimators for discretely sampled Lévy processes. *Annals of Statistics* 35, 335–392.
- Aït-Sahalia, Y. and J. Jacod (2008). Fisher’s information for discretely sampled Lévy processes. *Econometrica* 35, 727–761.
- Andersen, T. G. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance* 4, 115–158.
- Andersen, T. G. and T. Bollerslev (1998a). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, 885–905.
- Andersen, T. G. and T. Bollerslev (1998b). Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance* 53, 219–265.
- Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics* 89, 707–720.
- Andersen, T. G., T. Bollerslev, and F. X. Diebold (2009). Parametric and nonparametric measurement of volatility. In Y. Aït-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*, pp. 67–138. Amsterdam: North Holland.
- 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.
- Andersen, T. G., D. Dobrev, and E. Schaumburg (2009). Jump-robust volatility estimation using nearest neighbor truncation. Unpublished paper.
- Andersen, T. G., D. Dobrev, and E. Schaumburg (2011). A functional filtering and truncation approach to integrated return quarticity estimation. Unpublished paper.
- Bahadur, R. (1964). On Fisher’s bound for asymptotic variances. *Ann. Math. Statist.* 35, 1545–1552.
- Bahadur, R. (1980). Hodges superefficiency. In *Encyclopedia of Statistical Sciences*, Volume 3 (F-H). New York: Wiley.
- Bandi, F. M. and J. R. Russell (2008). Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies* 75, 339–369.
- Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, and N. Shephard (2006). Limit theorems for realised bipower variation in econometrics. *Econometric Theory* 22, 677–719.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* 76, 1481–1536.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realised kernels in practice: trades and quotes. *Econometrics Journal* 12, C1–C32.
- Barndorff-Nielsen, O. E., S. Kinnebrock, and N. Shephard (2009). Measuring downside risk: realised semivariance. In T. Bollerslev, J. Russell, and M. Watson (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, pp. 117–136. Oxford University Press.
- 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). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* 2, 1–48.
- Barndorff-Nielsen, O. E. and N. Shephard (2005). How accurate is the asymptotic approximation to the distribution of realised volatility? In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models. A Festschrift in Honour of T.J. Rothenberg*, pp. 306–331. Cambridge: Cambridge University Press.
- 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.

- Barndorff-Nielsen, O. E. and N. Shephard (2007). Variation, jumps and high frequency data in financial econometrics. In R. Blundell, T. Persson, and W. K. Newey (Eds.), *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress*, Econometric Society Monographs, pp. 328–372. Cambridge University Press.
- Barndorff-Nielsen, O. E., N. Shephard, and M. Winkel (2006). Limit theorems for multipower variation in the presence of jumps. *Stochastic Processes and Their Applications* 116, 796–806.
- Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies* 22, 4463–4492.
- Bollerslev, T. and V. Todorov (2011). Estimation of jump tails. Unpublished paper: Northwestern University.
- Christensen, K. and M. Podolskij (2006). Asymptotic theory for range-based estimation of quadratic variation of a discontinuous semi-martingale. Unpublished paper: Aarhus School of Business.
- Delbaen, F. and W. Schachermeyer (2006). *The Mathematics of Arbitrage*. Springer.
- Dellacherie, C. and P. Meyer (1982). *Probabilities and Potential B*. Amsterdam: North-Holland.
- 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.
- Goncalves, S. and N. Meddahi (2004). Bootstrapping realized volatility. *Econometrica* 77, 283–306.
- Hájek, J. (1969). A characterization of limiting distributions of regular estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 14, 323–330.
- Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variation. *Journal of Financial Econometrics* 3, 456–499.
- Jacod, J. (2008a). Asymptotic properties of realized power variations and related functionals of semi-martingales. *Stochastic Processes and Their Applications* 118, 517–559.
- Jacod, J. (2008b). Statistics and high frequency data. Unpublished paper.
- Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and Their Applications* 119, 2249–2276.
- 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.
- Kinnebrock, S. and M. Podolskij (2008). A note on the central limit theorem for bipower variation of general functions. *Stochastic Processes and Their Applications* 118, 1056–1070.
- LeCam, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes’ estimates. *Univ. of California Publ. in Statistics* 1, 277–330.
- Lee, S. and P. A. Mykland (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Studies* 21, 2535–2563.
- Lehmann, E. (1983). *Theory of Point Estimation*. New York: Wiley.
- Maheswaran, S. and C. A. Sims (1993). Empirical implications of arbitrage-free asset markets. In P. C. B. Phillips (Ed.), *Models, Methods and Applications of Econometrics*, pp. 301–316. Basil Blackwell.
- Mancini, C. (2004). Estimation of the characteristics of jump of a general Poisson-diffusion process. *Scandinavian Actuarial Journal* 1, 42–52.
- Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics* 36, 270–296.
- Mykland, P. A. (1994). Bartlett type identities for martingales. *Annals of Statistics* 22, 21–38.
- Mykland, P. A. and L. Zhang (2006). ANOVA for diffusions and Ito processes. *Annals of Statistics* 34, 1931–1963.
- Mykland, P. A. and L. Zhang (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica* 77, 1403–1455.
- Mykland, P. A. and L. Zhang (2010). The econometrics of high frequency data. (to appear in *Statistical Methods for Stochastic Differential Equations*, M. Kessler, A. Lindner, and M. Sørensen, eds., Chapman and Hall/CRC Press).

- Patton, A. J. and K. K. Sheppard (2009). Good volatility, bad volatility: signed jumps and persistence of volatility. Unpublished paper: Oxford-Man Institute, University of Oxford.
- Protter, P. (2004). *Stochastic Integration and Differential Equations*. New York: Springer-Verlag.
- Qi, F. (2010). Bounds for the ratio of two gamma functions. *Journal of Inequalities and Applications* 2010, 1–84.
- Rootzén, H. (1980). Limit distributions for the error in approximations of stochastic integrals. *Annals of Probability* 8, 241–251.
- Shiryaev, A. N. (1999). *Essentials of Stochastic Finance: Facts, Models and Theory*. Singapore: World Scientific.
- Veraart, A. E. D. (2010). Inference for the jump part of quadratic variation of ito semimartingales. *Econometric Theory* 26, 331–368.
- Xiu, D. (2010). Quasi-maximum likelihood estimation of volatility with high frequency data. *Journal of Econometrics* 159, 235–250.
- Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli* 12, 1019–1043.
- 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.

# APPENDIX

## A Further results in the continuous case

In the following section, the observed process is assumed continuous, *i.e.*,  $J_t \equiv 0$ .

### A.1 Multipower variation as $M \rightarrow \infty$

#### A.1.1 An approximation theorem for bipower variation

To transparently study more deeply the case where  $M \rightarrow \infty$  as  $n \rightarrow \infty$ , we use the simplified estimator ("semi-bipower variation"):

$$SBV = 2 \sum_{1 \leq i; (2i+1)M \leq n} (RV_{2i-1}RV_{2i})^{1/2}. \quad (\text{A.1})$$

We use the limiting value 1 of  $k_M/M$ . Also, for simplicity, we avoid overlapping intervals, though see above, and also Remark 13 below for the more general situation.

The results in this and the next subsection are stronger than those of Theorem 3 in the sense that they provide stable convergence to a normal limit in place of an  $O_p$  term. This gives reasonable certainty that a similar stable convergence holds in the earlier theorem, and thus that the  $O_p$  term cited there does not vanish. Since such a result would mainly be relevant for the estimation of volatility of volatility, such a development is beyond the scope of this paper. The non-vanishing of the  $O_p$  terms, however, is important in Remarks 6-7.

The relationship between RV and SBV can be written

$$RV_n - SBV_n = \sum_{1 \leq i; (2i+1)M \leq n} f(RV_{2i-1}, RV_{2i}), \quad (\text{A.2})$$

where  $f(x, y) = (y^{1/2} - x^{1/2})^2$ . With  $\Delta\tau = M\Delta t = MT/n$ , we can also write

$$RV_n - SBV_n = \sum_{1 \leq i; (2i+1)M \leq n} f(RV_{2i-1}/\Delta\tau, RV_{2i}/\Delta\tau)\Delta\tau. \quad (\text{A.3})$$

Based on this representation, we obtain the following result under model (6). The symbol  $[\cdot, \cdot]_t$  represents the quadratic variation (integrated variance) of a continuous process.

**Theorem 7** *Under the assumptions of Theorem 3, and if  $X$  has no jumps, set*

$$a^2 = \int_0^T c^{-1} \left( \frac{1}{2}T\sigma_t^2 + \frac{2}{3}c^2[\sigma, \sigma]'_t \right)^2 dt \text{ and } b_1 = c^{-1}\frac{1}{2}T \int_0^T \sigma_t^2 dt \text{ and } b_2 = \frac{1}{3}c [\sigma, \sigma]_T. \quad (\text{A.4})$$

*Then, under Assumption 1, if  $\beta = 1/2$ ,*

$$n^{\frac{3}{4}} \left( RV_n - SBV_n - (n^{-1+\beta}b_1 + n^{-\beta}b_2) \right) \xrightarrow{\mathcal{L}} a \times Z \quad (\text{A.5})$$

*where  $Z$  is standard normal and independent of the underlying filtration.*

*In the case where  $\beta > 1/2$ , only keep the first term inside the brackets in  $a^2$ , and replace  $n^{\frac{3}{4}}$  by  $n^{1-\frac{1}{2}\beta}$  in (A.5). In the case where  $\beta < 1/2$ , only keep the second term inside the brackets in  $a^2$ , and replace  $n^{\frac{3}{4}\alpha}$  by  $n^{\frac{3}{2}\beta}$  in (A.5).*

The result is shown in Remark 11 below. The proof of how to get from here to Theorem 3 is given in Appendix B.4.

A path to understanding the bias is that following the proof in Section B.5,

$$RV - SBV = \frac{1}{2} \sum_{1 \leq i; (2i+1)M \leq n} \frac{(RV_{2i} - RV_{2i-1})^2}{RV_{2i-1}} + O_p(n^{-3/4}), \quad (\text{A.6})$$

cf., in particular, equation (B.22).

### A.1.2 Towards a more general theory: microstructure, and multipower

To show that it is possible to pursue a similar theory in the more general case of estimation under microstructure, we provide a theorem where “ $RV_i$ ” is a general estimator of the integrated volatility in interval  $\# i$ . (Denoted by  $\tilde{R}V$  in Section 3.5.)

**Theorem 8** *Assume that under the statistical risk neutral distribution<sup>7</sup>, there is a sequence of continuous martingales  $(M_t^{(n)})_{0 \leq t \leq T}$  so that*

$$RV_i = \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt + M_{\tau_i}^{(n)} - M_{\tau_{i-1}}^{(n)} = \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt + \Delta M_{\tau_i}^{(n)}. \quad (\text{A.7})$$

Suppose that as  $n \rightarrow \infty$ ,

$$n^{2\alpha} [M^{(n)}, M^{(n)}]_t \xrightarrow{p} \int_0^t f_s^2 ds \text{ and } n^\alpha [M^{(n)}, W^{(i)}]_t \xrightarrow{p} 0 \text{ for } i = 1, \dots, p. \quad (\text{A.8})$$

Also assume (B.15)-(B.16). In the case when  $\beta = \alpha$ , set

$$a^2 = \int_0^T c^{-1} \left( \frac{1}{4} \sigma_t^{-2} f_t^2 + \frac{2}{3} c^2 [\sigma, \sigma]'_t \right)^2 dt \text{ and } b_1 = \frac{1}{4} c^{-1} \int_0^T \sigma_t^{-2} f_t^2 dt \text{ and } b_2 = \frac{1}{3} c [\sigma, \sigma]_T. \quad (\text{A.9})$$

Then, under Assumption 1,

$$n^{\frac{3}{2}\alpha} \left( RV_n - SBV_n - (n^{-2\alpha+\beta} b_1 + n^{-\beta} b_2) \right) \xrightarrow{\mathcal{L}^s} a \times Z \quad (\text{A.10})$$

where  $Z$  is standard normal and independent of the underlying filtration.

In the case where  $\beta > \alpha$ , only keep the first term inside the brackets in  $a^2$ , and replace  $n^{\frac{3}{2}\alpha}$  by  $n^{2\alpha-\frac{1}{2}\beta}$  in (A.10). In the case where  $\beta < \alpha$ , only keep the second term inside the brackets in  $a^2$ , and replace  $n^{\frac{3}{2}\alpha}$  by  $n^{\frac{3}{2}\beta}$  in (A.10).

**Proof.** Given in the Appendix.

**Remark 10** *Assumption 1, along with (A.8) assures that  $n^\alpha M_t^{(n)}$  converges stably in law to  $M_t$  (as a process), with  $M_t = \int_0^t f_s dB_s$ , where  $B$  is a Brownian motion independent of  $\mathcal{F}_T$ . (See Theorem 6 in Mykland and Zhang (2010), or similar results in Jacod and Shiryaev (2003).)*

In the classical case of no microstructure, we obtain

**Remark 11** *(No microstructure). To verify Theorem 7 from Theorem 8, we note that in this case,  $\alpha = 1/2$ .  $M_t^{(n)}$  is the usual martingale associated with realised volatility, and stable convergence follows from Jacod and Protter (1998), Barndorff-Nielsen and Shephard (2002), and Mykland and Zhang (2006). The conditions (A.8) is one path to verifying the stable convergence, and otherwise (A.8) follows from the stable convergence via Corollary VI.6.30 (p. 385) in Jacod and Shiryaev (2003). From the same literature,  $f_t^2 = 2T\sigma_t^4$ .*

<sup>7</sup>Section 2.2 of Mykland and Zhang (2009).

**Remark 12** *It is seen from the proof that the choice (B.16) minimizes the order of both the bias and stochastic terms, cf. the statement just after equation (B.26), and similar considerations for the stochastic term. The optimal block size is  $M = O(n^{1-\alpha})$ .*

**Remark 13** *Using overlapping intervals, or a rolling window, will produce the same bias terms, but will gain efficiency for the stochastic term. The order  $O_p(n^{-\frac{3}{2}\alpha})$  of this term is, however, unlikely to change.*

Finally, to study multipower variation, we introduce the "semi-multipower":

$$SMV_n^{(K)} = K \sum_{1 \leq i; (Ki+1)M \leq n} (RV_{2i-K+1} \times \dots \times RV_{2i})^{1/K}. \quad (\text{A.11})$$

We here only study the first order behavior, and the order of the difference with RV

**Theorem 9** *Assume the conditions of Theorem 8. Set*

$$b_1 = \frac{1}{2}c^{-1} \frac{K-1}{K} \int_0^T \sigma_t^{-2} f_t^2 dt \text{ and } b_2 = \frac{K-1}{3}c [\sigma, \sigma]_T. \quad (\text{A.12})$$

Then

$$RV_n - SMV_n^{(K)} = n^{-2\alpha+\beta}b_1 + n^{-\beta}b_2 + O_p(n^{-2\alpha+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}). \quad (\text{A.13})$$

**Proof.** Given in the Appendix.

## B Assumptions and Proofs

### B.1 Assumptions

**Assumption 1** . *The continuous part of the process  $X_t$  is adapted to a filtration  $(\mathcal{F}_t)$  which is generated by Brownian motions  $W_t^{(1)}, \dots, W_t^{(p)}$  (for some  $p$ ). That*

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + J_t, \quad (\text{B.14})$$

where  $\mu$  is a predictable locally bounded drift,  $W$  is a standard Brownian motion and  $J$  is a finite activity jump process.  $\sigma_t$  is also locally bounded away from zero, and is itself a continuous semimartingale (Itô process).

**Assumption 2** . *Observations*

$$t_i = \frac{i}{n}, \quad i = 0, 1, 2, \dots, n.$$

With blocks of size  $M$  we have block boundaries

$$\tau_i = \frac{Mi}{n}, \quad i = 0, 1, \dots, \lfloor n/M \rfloor = n_M.$$

**Assumption 3** . *As  $n \rightarrow \infty$ ,*

$$n^\beta \frac{M}{n} \rightarrow \frac{c}{T}, \quad (\text{B.15})$$

where  $\beta \in (0, 1)$ .

**Assumption 4**

$$0 < \beta < 1. \quad (\text{B.16})$$

## B.2 Proof of Remark 1.

For simplicity, take  $T = 1$ . Now

$$BV_M \stackrel{\mathcal{L}}{=} \frac{\sigma^2}{n} \frac{n}{(n_M - 1) k_M} \sum_{i=2}^{n_M} \chi_{M,i-1} \chi_{M,i} \stackrel{\mathcal{L}}{=} \frac{\sigma^2}{(n_M - 1) k_M} \sum_{i=2}^{n_M} \chi_{M,i-1} \chi_{M,i},$$

which has a expectation and variance of

$$\frac{\sigma^2}{(n_M - 1) \{E(\chi_M)\}^2} (n_M - 1) \{E(\chi_M)\}^2 = \sigma^2$$

and noting that

$$\begin{aligned} \text{Var}(\chi_{M,1} \chi_{M,2}) &= E(\chi_{M,1}^2 \chi_{M,2}^2) - \{E(\chi_{M,1})\}^4 = M^2 - k_M^2 \\ \text{Cov}(\chi_{M,1} \chi_{M,2}, \chi_{M,3} \chi_{M,2}) &= M k_M - k_M^2 \end{aligned}$$

we have

$$\begin{aligned} & \frac{\sigma^4}{(n_M - 1)^2 k_M^2} [(n_M - 1) \text{Var}(\chi_{M,1} \chi_{M,2}) + 2(n_M - 2) \text{Cov}(\chi_{M,1} \chi_{M,2}, \chi_{M,3} \chi_{M,2})] \\ &= \frac{\sigma^4}{(n_M - 1)^2 k_M^2} \{(n_M - 1)(M^2 - k_M^2) + 2(n_M - 2)(M k_M - k_M^2)\} \\ &= \frac{\sigma^4}{(n_M - 1)^2} \left\{ (n_M - 1) \frac{(M^2 + 2M k_M - 3k_M^2)}{k_M^2} - \frac{2(M k_M - k_M^2)}{k_M^2} \right\} \\ &= \frac{\sigma^4}{M (n_M - 1)^2} \{(n_M - 1) M \theta_M - \psi_M\}, \quad \psi_M = 2M (M/k_M - 1) = 1 + o(1), \\ &= \frac{\sigma^4}{M (n_M - 1)} \left( M \theta_M - \frac{1}{(n_M - 1)} \psi_M \right) \sim \frac{\sigma^4}{n} (M \theta_M), \end{aligned} \tag{B.17}$$

the latter transition for large  $n$  so long as  $M/n = o(1)$ . In practice it makes sense to use the finite sample constant

$$d = \frac{1}{M (n_M - 1)} \left( M \theta_M - \frac{1}{(n_M - 1)} \psi_M \right).$$

## B.3 Proof of Theorem 1

For compactness of notation, set  $\theta = \int_0^T \sigma_t^2 dt$ , and let  $\hat{\theta}_n$  be  $BV_M$  based on  $n$  observations. (Not to be confused with  $\theta_M$  in (8).)  $RV_i$  is based on blocks of size  $M$ . Let  $\mathcal{M}$  be a large integer. Decompose

$$\begin{aligned} \hat{\theta}_n &= \hat{\theta}_n^{(1)} + \hat{\theta}_n^{(2)}, \text{ where} \\ \hat{\theta}_n^{(1)} &= \frac{M}{k_M} \sum_k \sum_{i=k\mathcal{M}+1}^{(k+1)\mathcal{M}-1} (RV_i RV_{i+1})^{1/2} \text{ and} \\ \hat{\theta}_n^{(2)} &= \frac{M}{k_M} \sum_k (RV_{k\mathcal{M}} RV_{k\mathcal{M}+1})^{1/2}. \end{aligned}$$

For the first term, note that,

$$n^{1/2} \sum_k \left( \frac{M}{k_M} \sum_{i=k\mathcal{M}+1}^{(k+1)\mathcal{M}-1} (RV_i RV_{i+1})^{1/2} - \frac{\mathcal{M}-1}{\mathcal{M}} \sigma_{\tau_{n,k}}^2 \Delta\tau \right)$$

is a sum of  $(\mathcal{Y}_{n,k}, Q_n)$  martingale increments in the sense of the definitions in Section 3 and 4 (on p. 1417 and p. 1421) of Mykland and Zhang (2009). The corresponding block size is  $M\mathcal{M}$  in lieu of  $M$ .  $\tau_{n,k}$  is also as defined in this paper.

The martingale has quadratic variation

$$\begin{aligned}
& n \sum_k \text{Var}_{Q_n} \left( \frac{M}{k_M} \sum_{i=k\mathcal{M}+1}^{(k+1)\mathcal{M}-1} (RV_i RV_{i+1})^{1/2} - \frac{\mathcal{M}-1}{\mathcal{M}} \sigma_{\tau_{n,k}}^2 \Delta\tau \mid \mathcal{Y}_{n,k} \right) \\
&= n \sum_k \sigma_{\tau_{n,k}}^4 \left( \frac{M}{k_M} \right)^2 \Delta t_n^2 \text{Var} \left( \sum_{i=1}^{\mathcal{M}-1} \chi_{M,i} \chi_{M,i+1} \right) \\
&\rightarrow \frac{T}{M\mathcal{M}} \left( \frac{M}{k_M} \right)^2 \text{Var} \left( \sum_{i=1}^{\mathcal{M}-1} \chi_{M,i} \chi_{M,i+1} \right) \int_0^T \sigma_t^4 dt, \tag{B.18}
\end{aligned}$$

in probability as  $n \rightarrow \infty$ , where the  $\chi_{M,i}$  are i.i.d.  $\chi_M$ .

By the same arguments as in Mykland and Zhang (2009), we thus get that  $n^{1/2}(\hat{\theta}_n^{(1)} - \theta(\mathcal{M} - 1)/\mathcal{M})$  converges stably under  $Q_n$  to a mixed normal distribution with mean zero and variance (B.18). It is easy to see that there is no adjustment to  $P_n^*$ ,  $P^*$ , or  $P$ , as these measures are defined in the referenced paper.

Similar arguments, using a first block of size  $\mathcal{M}$ , and then blocks of size  $M\mathcal{M}$ , yield that  $n^{1/2}(\hat{\theta}_n^{(2)} - \theta/\mathcal{M})$  converges stably, to a variance which is of  $o_p(1)$  as  $\mathcal{M} \rightarrow \infty$  ( $\mathcal{M}$  is sent to infinity *after*  $n$ .) Standard weak convergence arguments involving tightness and subsequences of subsequences thus yield that  $n^{1/2}(\hat{\theta}_n - \theta)$  converges stably under  $P$  to a mixed normal limit with mean zero and a random variance which is the limit of (B.18) as  $M \rightarrow \infty$ , in other words,

$$TM\theta_M \int_0^T \sigma_t^4 dt, \tag{B.19}$$

where this  $\theta_M$  is as in (8). This shows the first part of the theorem. The second part follows similarly.

## B.4 Proof of Theorem 3

To get from Theorems 7 and 9 to Theorem 3, consider first bipower variation and the case of no jumps. Set

$$SBV'_n = \sum_{i=2}^{n_M} (RV_{i-1} RV_i)^{1/2}.$$

By symmetry, it follows from Theorem 7 that

$$\begin{aligned}
RV_n - SBV'_n &= n^{-1+\beta} b_1 + n^{-\beta} b_2 + O_p(n^{-1+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}) \\
&= n^{-1+\beta} (c^{-1} \frac{T}{2} RV_n + O_p(n^{-\frac{1}{2}})) + n^{-\beta} b_2 + O_p(n^{-1+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}) \\
&= n^{-1+\beta} c^{-1} \frac{T}{2} RV_n + n^{-\beta} b_2 + O_p(n^{-1+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}) \tag{B.20}
\end{aligned}$$

where the  $O_p$  terms are asymptotically unbiased to this same order. Now set

$$SBV''_n = \frac{n}{n_M k_M} SBV'_n. \tag{B.21}$$

Because of the exact unbiasedness of  $SBV_n''$  when  $\sigma$  is constant, we get by uniform integrability that  $\sigma^2 T = E(SBV_n'') = \frac{n}{n_M k_M} E(SBV_n') = \frac{n}{n_M k_M} \left( E(RV) - n^{-1+\beta} c^{-1} \frac{1}{2} T^2 \sigma^2 + o(n^{-1+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}) \right)$ , whence

$$1 = \frac{n}{n_M k_M} \left\{ 1 - n^{-1+\beta} c^{-1} \frac{1}{2} T + o(n^{-1+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta}) \right\}.$$

Substituting the latter formula into equation (B.20) yields the result of the theorem.

For multipower variation, invoke Theorem 9 directly, with the values given in Remark 11, and then use the same unbiasedness argument as for bipower.

The extension to jumps is straightforward.

## B.5 Proof of Theorem 8

We Taylor expand  $f(RV_{2i-1}/\Delta\tau, RV_{2i}/\Delta\tau)$  around  $f(\sigma_{\tau_{2i-2}}^2, \sigma_{\tau_{2i-2}}^2)$ . Note that  $f(z, z) = f_x(z, z) = f_y(z, z) = 0$ , while  $f_{x,x}(z, z) = f_{y,y}(z, z) = -f_{x,y}(z, z) = 1/2z$ . Hence, up to second order,

$$\begin{aligned} RV - SBV &\approx \frac{1}{2} \Delta\tau \sum_{1 \leq i; (2i+1)M \leq n} \left\{ \left( \frac{RV_{2i-1}}{\Delta\tau} - \sigma_{\tau_{2i-2}}^2 \right)^2 f_{x,x}(\sigma_{\tau_{2i-2}}^2, \sigma_{\tau_{2i-2}}^2) \right. \\ &\quad + \left( \frac{RV_{2i}}{\Delta\tau} - \sigma_{\tau_{2i-2}}^2 \right)^2 f_{y,y}(\sigma_{\tau_{2i-2}}^2, \sigma_{\tau_{2i-2}}^2) \\ &\quad \left. + 2 \left( \frac{RV_{2i-1}}{\Delta\tau} - \sigma_{\tau_{2i-2}}^2 \right) \left( \frac{RV_{2i}}{\Delta\tau} - \sigma_{\tau_{2i-2}}^2 \right) f_{x,y}(\sigma_{\tau_{2i-2}}^2, \sigma_{\tau_{2i-2}}^2) \right\} \\ &= \frac{1}{4} \Delta\tau \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-2} \left( \frac{RV_{2i}}{\Delta\tau} - \frac{RV_{2i-1}}{\Delta\tau} \right)^2 \\ &= \frac{1}{4} \frac{1}{\Delta\tau} \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-2} (RV_{2i} - RV_{2i-1})^2 \end{aligned} \quad (\text{B.22})$$

The remainder term in the above is of order  $o_p(n^{-\frac{3}{2}\alpha})$ , as explained at the end of the proof. Since both the bias and stochastic terms are of order  $O_p(n^{-\frac{3}{2}\alpha})$ , the remainder can be ignored.

To get a further handle on this quantity recall (A.7). By Itô's formula,

$$\int_{\tau_{2i-1}}^{\tau_{2i}} \sigma_t^2 dt = \sigma_{\tau_{2i-1}}^2 \Delta\tau + \int_{\tau_{2i-1}}^{\tau_{2i}} (\tau_{2i} - t) d\sigma_t^2 \quad \text{and} \quad \int_{\tau_{2i-2}}^{\tau_{2i-1}} \sigma_t^2 dt = \sigma_{\tau_{2i-1}}^2 \Delta\tau - \int_{\tau_{2i-2}}^{\tau_{2i-1}} (t - \tau_{2i-2}) d\sigma_t^2. \quad (\text{B.23})$$

Thus,  $RV_{2i} - RV_{2i-1} = L_{\tau_{2i}}^{(n)} - L_{\tau_{2i-1}}^{(n)}$ , where  $dL_t^{(n)} = -dM_t^{(n)} - (t - \tau_{2i-2}) d\sigma_t^2$  for  $\tau_{2i-2} < t < \tau_{2i-1}$ , and  $dL_t^{(n)} = dM_t^{(n)} + (\tau_{2i} - t) d\sigma_t^2$  for  $\tau_{2i-1} < t < \tau_{2i}$ . In particular,

$$\begin{aligned} d[L^{(n)}, L^{(n)}]_t &= d[M^{(n)}, M^{(n)}]_t + (t - \tau_{2i-2})^2 d[\sigma^2, \sigma^2]_t + 2(t - \tau_{2i-2}) d[M^{(n)}, \sigma^2]_t \quad \text{for } \tau_{2i-2} < t < \tau_{2i-1} \\ d[L^{(n)}, L^{(n)}]_t &= d[M^{(n)}, M^{(n)}]_t + (\tau_{2i} - t)^2 d[\sigma^2, \sigma^2]_t + 2(\tau_{2i} - t) d[M^{(n)}, \sigma^2]_t \quad \text{for } \tau_{2i-1} < t < \tau_{2i} \end{aligned}$$

The cross term involving  $[M^{(n)}, \sigma^2]_t$  will disappear in the following to relevant order, from Condition (A.8).

Recall that  $n^\alpha M_t^{(n)}$  converges in law to  $M_t$  (as a process, from Remark 10), so that in particular,  $n^{2\alpha} [M^{(n)}, M^{(n)}]_t$  converges to  $[M, M]_t$ . Note that

$$[L^{(n)}, L^{(n)}]_{\tau_{2i}} - [L^{(n)}, L^{(n)}]_{\tau_{2i-2}} \approx n^{-2\alpha} ([M, M]_{\tau_{2i}} - [M, M]_{\tau_{2i-2}})$$

$$\begin{aligned}
& + \int_{\tau_{2i-2}}^{\tau_{2i-1}} (t - \tau_{2i-2})^2 d[\sigma^2, \sigma^2]_t + \int_{\tau_{2i-1}}^{\tau_{2i}} (\tau_{2i} - t)^2 d[\sigma^2, \sigma^2]_t \\
& \approx n^{-2\alpha} ([M, M]_{\tau_{2i}} - [M, M]_{\tau_{2i-2}}) + \frac{2}{3} (\Delta\tau)^3 [\sigma^2, \sigma^2]'_{\tau_{2i-2}},
\end{aligned} \tag{B.24}$$

where ‘‘prime’’ denotes differentiation with respect to time.

We therefore get

$$\begin{aligned}
& \text{r.h.s. of (B.22)} \\
& \approx \frac{1}{4} \frac{1}{\Delta\tau} \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-2} \left( [M^{(n)}, M^{(n)}]_{\tau_{2i}} - [M^{(n)}, M^{(n)}]_{\tau_{2i-2}} + \frac{2}{3} (\Delta\tau)^3 [\sigma^2, \sigma^2]'_{\tau_{2i-2}} \right)
\end{aligned} \tag{B.25}$$

$$\begin{aligned}
& \approx \frac{1}{4} \frac{1}{\Delta\tau} \int_0^T \sigma_t^{-2} d[M^{(n)}, M^{(n)}]_t + \frac{1}{12} (\Delta\tau) \int_0^T \sigma_t^{-2} d[\sigma^2, \sigma^2]_t \\
& \approx \frac{1}{4} \frac{1}{n^{2\alpha} \Delta\tau} \int_0^T \sigma_t^{-2} d[M, M]_t + \frac{1}{12} \Delta\tau \int_0^T \sigma_t^{-2} d[\sigma^2, \sigma^2]_t.
\end{aligned} \tag{B.26}$$

From (B.16) we obtain

$$\begin{aligned}
& \text{r.h.s. of (B.22)} \approx n^{-2\alpha+\beta} \frac{1}{4} c^{-1} \int_0^T \sigma_t^{-2} d[M, M]_t + n^{-\beta} \frac{1}{12} c \int_0^T \sigma_t^{-2} d[\sigma^2, \sigma^2]_t \\
& = n^{-2\alpha+\beta} \frac{1}{4} c^{-1} \int_0^T \sigma_t^{-2} d[M, M]_t + n^{-\beta} \frac{1}{3} c [\sigma, \sigma]_T
\end{aligned} \tag{B.27}$$

since  $d\sigma_t^2 = 2\sigma_t d\sigma_t + d[\sigma, \sigma]_t$ .

Meanwhile, the error in approximation in (B.26) has quadratic variation

stoch term qv

$$\begin{aligned}
& = \left( \frac{1}{4} \frac{1}{\Delta\tau} \right)^2 \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-4} 4 \int_{\tau_{2i-2}}^{\tau_{2i}} ([L^{(n)}, L^{(n)}]_t - [L^{(n)}, L^{(n)}]_{\tau_{2i-2}}) d[L^{(n)}, L^{(n)}]_t \\
& = \left( \frac{1}{4} \frac{1}{\Delta\tau} \right)^2 \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-4} 2 ([L^{(n)}, L^{(n)}]_{\tau_{2i}} - [L^{(n)}, L^{(n)}]_{\tau_{2i-2}})^2 \\
& \approx \left( \frac{1}{4} \frac{1}{\Delta\tau} \right)^2 \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-4} 2 (n^{-2\alpha} ([M, M]_{\tau_{2i}} - [M, M]_{\tau_{2i-2}}) + \frac{2}{3} (\Delta\tau)^3 [\sigma^2, \sigma^2]'_{\tau_{2i-2}})^2 \\
& \approx \left( \frac{1}{4} \frac{1}{\Delta\tau} \right)^2 \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-4} 2 (n^{-2\alpha} \Delta\tau [M, M]'_{\tau_{2i-2}} + \frac{2}{3} (\Delta\tau)^3 [\sigma^2, \sigma^2]'_{\tau_{2i-2}})^2
\end{aligned} \tag{B.28}$$

where the relative error is of smaller order in probability, from (B.24). Under (B.16), consider first the case where  $\beta = \alpha$ . Then

$$\begin{aligned}
& \text{stoch term qv} = \frac{1}{16} \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-4} 2n^{-3\alpha} c^{-1} \Delta\tau ([M, M]'_{\tau_{2i-2}} + \frac{2}{3} c^2 [\sigma^2, \sigma^2]'_{\tau_{2i-2}})^2 + o_p(n^{-3\alpha}) \\
& = n^{-3\alpha} \frac{1}{16} \int_0^T c^{-1} \sigma_t^{-4} ([M, M]'_t + \frac{2}{3} c^2 [\sigma^2, \sigma^2]'_t)^2 dt + o_p(n^{-3\alpha}) \\
& = n^{-3\alpha} \int_0^T c^{-1} (\frac{1}{4} \sigma_t^{-2} [M, M]'_t + \frac{2}{3} c^2 [\sigma, \sigma]'_t)^2 dt + o_p(n^{-3\alpha}).
\end{aligned} \tag{B.30}$$

Meanwhile, by similar arguments, it is easy to see that the covariation of the stochastic term with underlying processes is of order  $o_p(n^{-\frac{3\alpha}{2}})$ . Since these derivations remain valid over intervals  $[0, t]$  (replacing  $T$  with  $t$ ), we have shown that the right hand side of (B.22) converges as specified in Theorem 8. This follows from Theorem 6 in Mykland and Zhang (2010).

In the case where  $\beta \neq \alpha$ , consider separately the case where  $\beta > \alpha$  and  $\beta < \alpha$ . In the former case, the first term inside the brackets in (B.29)-(B.30) dominate, and you replace  $c^{-1}n^{-3\alpha}$  by  $n^{-4\alpha}\Delta\tau^{-1} = c^{-1}n^{-4\alpha+\beta}$ . In the latter case, the second term dominates, and you replace  $c^3n^{-3\alpha}$  by  $\Delta\tau^3 = c^3n^{-3\beta}$ .

To complete the proof, we need to show that the remainder terms in (B.22) are of lower order. For simplicity of exposition, we just show that when  $\beta = \alpha$ , the remainder is of order  $o_p(n^{-\frac{3}{2}\alpha})$ . For this, note that for  $p + q \geq 2$ ,  $\partial^{p+q}f(x, y)/\partial x^p\partial y^q = -2a_p a_q x^{\frac{1}{2}-p}y^{\frac{1}{2}-q}$ , where  $a_0 = 1$ , and  $a_k = \frac{1}{2}\dots(\frac{1}{2} - k + 1)$  ( $k$  factors) for  $k \geq 1$ . Thus, if we truncate at the fourth derivative, by the usual truncation and stopping arguments, the expectation of this term is of order proportional to

$$\begin{aligned} \Delta\tau E & \sum_{1 \leq i; (2i+1)M \leq n} \left( \frac{RV_{2i}}{\Delta\tau} - \frac{RV_{2i-1}}{\Delta\tau} \right)^4 \\ & \leq c\Delta\tau^{-3} \sum_{1 \leq i; (2i+1)M \leq n} E([M^{(n)}, M^{(n)}]_{\tau_{2i}} - [M^{(n)}, M^{(n)}]_{\tau_{2i-2}})^2 \\ & = O(\Delta\tau^{-4}(n^{-3\alpha})^2) = O(n^{-2\alpha}) \end{aligned}$$

from the Burkholder-Davis-Gundy inequality (see Section 3 of Ch. VII of Dellacherie and Meyer (1982), or p. 193 and 222 in Protter (2004)) from (B.16) and (B.24). – This leaves the third order term, which is

$$-\frac{1}{3!}\Delta\tau \sum_{1 \leq i; (2i+1)M \leq n} \sigma_{\tau_{2i-2}}^{-4} \left( \frac{RV_{2i}}{\Delta\tau} - \frac{RV_{2i-1}}{\Delta\tau} \right)^3.$$

The Burkholder-Davis-Gundy bound is in this case of exact order  $O(n^{-\frac{3\alpha}{2}})$ . A long but tedious calculation shows, however, that the term goes away at this order. (Heuristically, this is because the third cumulant of an asymptotically normal sequence will vanish to first order. An exact derivation involves the Bartlett-type identities for martingales (Mykland (1994)).

This completes the proof of Theorem 8.

## B.6 Proof of Theorem 9

Set  $f(x_1, \dots, x_K) = \sum_{k=1}^K x_k - K \prod_{k=1}^K x_k^{1/K}$ . With  $\Delta\tau = M\Delta t = MT/n$ , we can write

$$RV_n - SMV_n^{(K)} = \sum_{1 \leq i; (Ki+1)M \leq n} f(RV_{Ki-K+1}/\Delta\tau, \dots, RV_{Ki}/\Delta\tau)\Delta\tau. \quad (\text{B.31})$$

As in the bipower case, note that  $f(z, \dots, z) = f_{x_i}(z, \dots, z) = 0$ , while  $f_{x_i x_i}(z, \dots, z) = z^{-1}(K-1)/K$  and  $f_{x_i x_j}(z, \dots, z) = z^{-1}(-1)/K$  for  $i \neq j$ . It follows that

$$\begin{aligned} f(x_1, \dots, x_K) &= \frac{1}{2}z^{-1} \left( \sum_{k=1}^K (x_k - z)^2 - \frac{1}{K} \left( \sum_{k=1}^K (x_k - z) \right)^2 \right) \\ &+ \text{terms involving derivatives of order greater than two.} \end{aligned}$$

Now set  $x_k = RV_{Ki-k+1}/\Delta\tau$  and  $z = \sigma_{\tau_{K(i-1)}}^2$ . For the same reasons as those discussed in detail in the proof of Theorem 8, we thus obtain that up to error of order  $O_p(n^{-2\alpha+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta})$ ,

$$RV_n - SMV_n^{(K)} \approx$$

$$\frac{1}{2\Delta\tau} \sum_{1 \leq i; (Ki+1)M \leq n} \sigma_{\tau_{K(i-1)}}^{-2} \left( \sum_{k=1}^K (RV_{Ki-k+1} - \sigma_{\tau_{K(i-1)}}^2 \Delta\tau)^2 - \frac{1}{K} \left( \sum_{k=1}^K (RV_{Ki-k+1} - \sigma_{\tau_{K(i-1)}}^2 \Delta\tau) \right)^2 \right) \quad (\text{B.32})$$

As in equation (B.23), we note that

$$\int_{\tau_a}^{\tau_b} \sigma_t^2 dt = \sigma_{\tau_a}^2 \Delta\tau + \int_{\tau_a}^{\tau_b} (\tau_b - t) d\sigma_t^2.$$

Thus, to sufficient approximation (when aggregated in (B.32))

$$\begin{aligned} & \sum_{k=1}^K (RV_{Ki-k+1} - \sigma_{\tau_{K(i-1)}}^2 \Delta\tau)^2 \\ &= \sum_{k=1}^K \left( \Delta M_{\tau_{Ki-k+1}}^{(n)} + \int_{\tau_{Ki-k}}^{\tau_{Ki-k+1}} (\tau_{Ki-k} - t) d\sigma_t^2 + \sigma_{\tau_{Ki-k}}^2 \Delta\tau - \sigma_{\tau_{K(i-1)}}^2 \Delta\tau \right)^2 \\ &\approx \sum_{k=1}^K \left( \Delta[M^{(n)}, M^{(n)}]_{\tau_{Ki-k+1}} + \int_{\tau_{Ki-k}}^{\tau_{Ki-k+1}} (\tau_{Ki-k+1} - t)^2 d[\sigma^2, \sigma^2]_t \right. \\ &\quad \left. + ([\sigma^2, \sigma^2]_{\tau_{Ki-k}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}}) \Delta\tau^2 \right) \\ &\approx ([M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}}) + \frac{1}{3} \Delta\tau^2 ([\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}}) \\ &\quad + \frac{K-1}{2} \Delta\tau^2 ([\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}}) \\ &\approx ([M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}}) + \left( \frac{K}{2} - \frac{1}{6} \right) \Delta\tau^2 ([\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}}), \end{aligned} \quad (\text{B.33})$$

while

$$\begin{aligned} & -\frac{1}{K} \left( \sum_{k=1}^K (RV_{Ki-k+1} - \sigma_{\tau_{K(i-1)}}^2 \Delta\tau) \right)^2 \\ &= -\frac{1}{K} \left( M_{\tau_{Ki}}^{(n)} - M_{\tau_{K(i-1)}}^{(n)} + \int_{\tau_{K(i-1)}}^{\tau_{Ki}} (\tau_{Ki} - t) d\sigma_t^2 \right)^2 \\ &\approx -\frac{1}{K} \left( [M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}} + \int_{\tau_{K(i-1)}}^{\tau_{Ki}} (\tau_{Ki} - t)^2 d[\sigma^2, \sigma^2]_t \right) \\ &\quad - \frac{1}{K} \left( [M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}} \right) - \frac{1}{3} K \Delta\tau^2 \left( [\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}} \right). \end{aligned} \quad (\text{B.34})$$

Combining (B.33)-(B.34) in (B.32), we obtain, again up to error of order  $O_p(n^{-2\alpha+\frac{1}{2}\beta} + n^{-\frac{3}{2}\beta})$ ,

$$\begin{aligned} & RV_n - SMV_n^{(K)} \\ &\approx \frac{1}{2\Delta\tau} \sum_{1 \leq i; (Ki+1)M \leq n} \sigma_{\tau_{K(i-1)}}^{-2} \frac{K-1}{K} \left( [M^{(n)}, M^{(n)}]_{\tau_{Ki}} - [M^{(n)}, M^{(n)}]_{\tau_{K(i-1)}} \right) \\ &\quad + \frac{1}{2\Delta\tau} \sum_{1 \leq i; (Ki+1)M \leq n} \sigma_{\tau_{K(i-1)}}^{-2} \frac{1}{6} (K-1) \Delta\tau^2 \left( [\sigma^2, \sigma^2]_{\tau_{Ki}} - [\sigma^2, \sigma^2]_{\tau_{K(i-1)}} \right) \end{aligned}$$

$$\approx \frac{1}{2} \frac{K-1}{K} \Delta\tau^{-1} n^{-2\alpha} \int_0^T \sigma_t^{-2} f_t^2 dt + \frac{K-1}{3} \Delta\tau[\sigma, \sigma]_T. \quad (\text{B.35})$$

Substituting by (B.16) gives the result of the theorem.

## B.7 Proof of Theorem 5

Since  $J$  is independent of  $X^c$ , we can condition on the  $J$  process. Also, it is enough to show the result when it is known that the jumps are in intervals  $(t_{i_1}, t_{i_1+1}]$ ,  $\dots$ ,  $(t_{i_\nu}, t_{i_\nu+1}]$ , where  $\nu$  is nonrandom (this adds information, and so can only reduce any lower bound). By the absolute continuity condition in the statement of the theorem, we thus only need to show the result when the jumps are of nonrandom size almost everywhere (with respect to Lebesgue measure). The problem then reduces to estimating the jump sizes (as parameters) when the observations are  $X_{t_{i_1+1}} - X_{t_{i_1}}, \dots, X_{t_{i_\nu+1}} - X_{t_{i_\nu}}$ . If these increments had normal distributions around jumps, the result would follow from the superefficiency results in LeCam (1953). The generalization to the non-normal case follows from Mykland and Zhang (2009, Thm 1, p. 1411), since the contiguity correction does not alter the asymptotic variance.