Efficiency Gain of Integrated Variance Estimation in the Presence of Jumps and Market Microstructure Noise

Shuichi NAGATA*

Abstract

In this paper, we consider integrated variance estimation in the presence of jumps and market microstructure noise in the observable price process. Podolskij and Vetter (2009) proposed the use of modulated bi-power variation (MBV), which is a consistent estimator of integrated variance in the presence of noise and jumps. We propose a consistent estimator based on the MBV and show that this new estimator is asymptotically more efficient than the MBV. A simulation experiment assessing the finite sample behavior of our estimator supports the theoretical results. Finally, we conduct an empirical analysis applying our new estimator to high-frequency data from the Nikkei 225 Stock Price Index. The results of our empirical example suggests implications for improving the performance of volatility forecasting.
Keywords: modulated power variation, integrated variance, jumps, market microstructure noise, HAR model.

1. Introduction

Integrated variance (IV) estimation using high-frequency data is one of the most prominent topics in the recent literature on financial econometrics. Realized variance or realized volatility (RV), based on the theory of the quadratic variation of stochastic processes is the most common estimator of IV because it has a simple structure and is statistically efficient when a price process is a semimartingale. A number of studies are examined using RV (see McAleer and Medeiros (2008) for a review of this topic).

However, recent studies report that RV does not converge to IV in a practical setting. We identify at least two obstacles to consistent IV estimation and discuss their treatment in the literature. The first one is market microstructure noise (MMN) and as its name suggests, such noise is caused by the microstructure of financial markets. MMN includes frictions such as price discreteness and bid-ask bounce. It is widely accepted that the observation price process of a financial asset is contaminated by MMN and that RV is not preferable in this case. We review this topic in a later section. The treatment of the MMN problem has been discussed by numerous researchers such as Zhou (1996), Bandi and Russell (2008),

* Assistant Professor, School of Business Administration, Kwansei Gakuin University.

and Mancino and Sanfelici (2008). Consistent estimators have been proposed by Zhang et al. (2005) and Barndorff-Nielsen et al. (2008), among others. Secondly, it is also important to note that RV is inconsistent when jumps exist in the price process. Previous research suggests that it is important to separate and account for the jumps effect in IV estimation. Bi-power variation (BV) proposed by Barndorff-Nielsen and Shephard (2004) is a well-known consistent estimator in the presence of finite activity jumps. The common practice in the literature is to detect a jump effect using BV and the statistical test using BV. Since Andersen et al. (2007) showed that the jump component of volatility does not contribute to volatility forecasting, jump-robust estimators such as BV and realized absolute variation (RAV) have frequently appeared in the volatility forecasting literature, including the studies by Ghysels et al. (2006) and Forsberg and Ghysels (2007). Andersen et al. (2010) offered a detailed summary of the previous research on and solutions to these two problems.

In this paper, we consider the consistent estimation of IV in the simultaneous presence of the above two obstacles. A solution to this problem was provided by Podolskij and Vetter (2009). They proposed modulated bi-power variation (MBV) and showed that MBV is a consistent estimator for IV in the presence of jumps and noise in the observable price process. However, as mentioned by Podolskij and Vetter (2009), while we can estimate IV consistently by MBV, it has a relatively lower efficiency than the other IV estimators.[1] To respond to this concern, we propose a new IV estimator, applying an efficiency gain technique by Nagata (2012a) to MBV. We explain the technique below. We show that this new estimator is consistent and asymptotically approaches the mixture normal in the presence of MMN and jumps. Most importantly, it achieves higher efficiency than MBV.

In addition to derive the asymptotic results, we also conduct a Monte Carlo simulation, which assesses the finite sample behavior of our proposed estimator. We compare its performance against alternatives. In the simulation experiments, we generate artificial price data for general stochastic volatility models with and without MMN and jumps. We compare the accuracy of our estimator to those of the alternatives using computational bias and root mean squared error (RMSE) at sampling frequencies normally used in practice.

We also conduct empirical research applying our new estimator to the Nikkei 225 Stock Price Index. This is one of the most common stock price indexes of the Tokyo Stock Exchange. Recently researchers such as Ghysels et al. (2006), Andersen et al. (2007), and Nagata (2012b) have suggested that lags in measures of intraday variation other than RV may have predictive power for volatility. They show that volatility prediction performance

---

1 The same authors discuss the pre-averaging estimator, which is more efficient than MBV, in the no jump case. see Podolskij and Vetter (2009b).

can be improved by using alternative explanatory variables in some forecasting models. These authors employ BV or RAV as an explanatory variable because they mainly consider the jump effect. However, as discussed by Andersen et al. (2011), MMN can have a detrimental effect on volatility forecasting. That is, it is worthwhile to examine the performance of our proposed estimator for volatility prediction, because our estimator is robust for both jumps and noise. We employ our estimator as an explanatory variable in the heterogeneous autoregressive (HAR) model and compare its performance against alternatives.

The remainder of this paper is organized as follows. In Section 2, we introduce our theoretical framework and show the main results of this paper. We conduct a simulation to assess our theoretical results in Section 3. In Section 4, we report the empirical analysis using our estimator. Finally, we conclude the paper in Section 5.

2. Efficiency Gain of Integrated Variance Estimation

Here, we first review the IV estimation in the presence of finite activity jumps excluding MMN. In the next subsection, we consider the IV estimation in the presence of jumps and MMN simul-taneously and present the results for the same.

2.1 IV Estimation with Jumps

Assume that the logarithmic price process $P(t)$ is determined by the univariate stochastic differential equation (SDE)

$$dP(t) = \mu(t)dt + \sigma(t)dW(t), \qquad 0 < t \leq 1, \tag{1}$$

where $\sigma(t)$ is the spot volatility process, and $W(t)$ is a standard Brownian motion. Following the literature, we assume that $\sigma(t)$ is cadlag as in most of the major volatility models.

Our target, IV, is defined as follows:

$$IV := \int_0^1 \sigma^2(s)ds.$$

To take into account jump effects for IV estimation, we consider the following model;

$$X = P + J,$$

where J denotes a finite activity jump process, such as compound Poisson processes. Here, we consider the estimation of IV from an observable price process from process X. We introduce the effect of market microstructure noise in the next section. Suppose that we observe M intra day prices, X_1, X_2, ..., X_M at t_1, t_2, ..., t_M. Then, RV, is defined as

$$RV := \sum_{i=1}^{M} r_i^2, \tag{2}$$

where $r_i = X_{t_i} - X_{t_{i-1}}$. It is well known that RV is an inconsistent estimator for IV in the presence of jumps, if $M \rightarrow \infty$, then

$$RV \xrightarrow{p} QV = IV + \sum_{s\in[0,1]:dq(s)=1} \kappa^2(s),$$

where QV is a quadratic variation, $dq(t)$ is the counting process and κ(t) is the jump size at time t.

The most common consistent estimator of IV is BV in this setting. BV is defined as

$$BV = \mu_1^{-2} \sum_{i=1}^{M} |r_i||r_{i+1}|. \tag{3}$$

where $\mu_1 = E\{|u|\} = \sqrt{\frac{2}{\pi}}$, and $u \sim N(0, 1)$. The efficiency gain of BV is discussed by Nagata (2012a).

2.2 IV Estimation with Jumps and Noise

Next, we consider IV estimation in the presence of not only jumps, but also MMN. It is widely accepted that the observable log price process Y is contaminated by MMN as seen in the following model:

$$Y_{t_i} = P_{t_i} + U_i, \quad U_i \sim i.i.d.(0, \omega^2). \tag{4}$$

where u_i is MMN, and its mean and variance are 0 and ω^2, respectively. It is also well known that when using observable price set Y_i, RV and BV do not converge to IV. Following Podolskij and Vetter (2009), We assume Y is defined on the filtered probability space $(\Omega, F, (F_t)_{t\in[0,1]}, P)$ and eq(4) is a its decomposition.

Same as no noise case, we consider the jump effect as $Z = Y + J$, so here we discuss the estimation of IV from an observable price process from process Z. To estimate IV consistently in this setting, Podolskij and Vetter (2009) proposed using MBV. First, unscaled MBV (UMBV) is defined as

$$UMBV_M(Z) = \sum_{i=1}^{N} |\bar{Z}_i^{(K)}||\bar{Z}_{i+1}^{(K)}|,$$

where

$$\bar{Z}_i^{(K)} = \frac{1}{\frac{M}{N} - K + 1} \sum_{j=\frac{(i-1)M}{N}}^{\frac{iM}{N}-K} (Z_{j+K} - Z_j),$$

and $K = c_1 M^{1/2}$, $N = M/c_2 K$, $c_1 > 0$, and $c_2 > 1$. Note that $\bar{Z}_i^{(K)}$ is the mean of all increments of length K within the interval $[t_{\frac{(i-1)M}{N}}, t_{\frac{M}{N}}]$.

Assumption

A.1: $E[|u|^{4+\varepsilon}] < \infty$ for some $\varepsilon > 0$.

A.2: filtered probability space $(\Omega, F, (F_t)_{t\in[0,1]}, P)$ supports another Brownian motion B, that is independent of the process X, such that $u_i = \sqrt{M}(B_{i/M} - B_{(i-1)/M})$.

A.3: Volatility function σ_t satisfies $\sigma = \sigma_0 + \int_0^t a'_s ds + \int_0^t \sigma'_s dW_s + \int_0^t v'_s dV_s$, where a' , σ' v' are adapted cadlag processes, a' being predictable and locally bounded. and V is a Brownian motion independent of W.

Under the above conditions, Podolskij and Vetter (2009) show the following important results about MBV.

Consistency of MBV (Proposition 3, Podolskij and Vetter (2009)) A.1 is hold and $M \to \infty$, then

$$MBV_M := \frac{\mu_1^{-2} c_1 c_2 UMBV_M(Z) - \nu_2 \hat{\omega}^2}{\nu_1} \xrightarrow{p} IV.$$

Asymptotic mixtured normarity (Corollary 2, Podolskij and Vetter (2009)) A.1 A.2 and A.3 are hold, then

$$\frac{M^{1/4}(MBV_M(Z) - IV)}{\beta_{mpv}} \sim N(0,1),$$

where

$$\beta_{mpv}^2 = c_1 c_2 V_{bv} \int_0^1 (\nu_1 \sigma^2(s) + \nu_2 \omega^2)^2 ds,$$

$$\nu_1 = \frac{c_1(3c_2 - 4 + (2 - c_2)^3 \vee 0)}{3(c_2 - 1)^2} \qquad \nu_2 = \frac{2((c_2 - 1) \wedge 1)}{c_1(c_2 - 1)^2},$$

and $V_{bv} = \mu_1^{-4} + 2\mu_1^{-2} - 3$.

However, as mentioned in the Introduction, MBV has relatively lower efficiency than the other IV estimators. Here, we propose a modulated three-step variation (MTV). Unscaled MTV (UMTV) can be calculated as

$$UMTV_{M,\alpha}(Z) = \frac{1}{N} \sum_{s=1}^{m} q_{\alpha,s}(Z)^2, \tag{6}$$

$$q_{\alpha,s}(Z) = \frac{1}{\eta_s} \sum_{j=1}^{n_s} |\bar{Z}_{\nu_{s-1}+j}^{(K)}|, \tag{7}$$

where m and n_i are positive integers, determined to satisfy $N = \sum_{i=1}^{m} n_i$, $m = O(N^{1-\alpha})$, and $n_i = O(N^{\alpha})$, where $\alpha \in (0, 1)$, $\nu_i = \sum_{k=1}^{i} n_k$, $\nu_0 = 0$, and $\eta_i = t_{\nu i} - t_{\nu i-1}$

Then, MTV is defined as follows:

$$MTV_{M,\alpha}(Z) := \frac{\mu_1^{-2} c_1 c_2 U MTV_{M,\alpha}(Z) - \nu_2 \hat{\omega}^2}{\nu_1}.$$

Considering the asymptotic properties of MTV, we can show the following theorems.

Theorem 1 Assumption A.1 is satisfied. If = $M \to \infty$, then

$$MTV_{M,\alpha}(Z) \xrightarrow{p} IV$$

Proof. The proof is given in Appendix A.

Theorem 2 Assumption A.1, A.2 and A.3 are satisfied. If $M \to \infty$, then

$$\frac{M^{1/4}(MTV_{M,\alpha}(Z) - IV)}{\sqrt{V_{mtv}}} \sim N(0,1),$$

where

$$V_{mtv} = c_1 c_2 V_{trv} \int_0^1 (\nu_1 \sigma^2(s) + \nu_2 \omega^2)^2 ds. \tag{8}$$

with $V_{trv} = 4(\mu_1^{-2} - 1)$.

Proof. The proof is given in Appendix A.

Remark 1 MTV is asymptotically more efficient than MBV. As shown by Podolskij and Vetter(2009), the asymptotic conditional variance of MBV is $\beta_{mbv}^2 = c_1 c_2 V_{bv} \int_0^t (\nu_1 \sigma^2(s) + \nu_2 \omega^2)^2 ds$. If we assume constant volatility, then $\beta_{mbv}^2 \simeq 22.87\sigma^3\omega$, which is smaller than $\beta_{mbv}^2 \simeq 26.14\sigma^3\omega$.

3. Monte Carlo Simulation

3.1 Simulation Design

In this section, we conduct Monte Carlo simulation experiments with several settings. The simulations are carried out to evaluate and compare the finite sample performance of our proposed estimator and the major integrated variance estimators. We calculate RV, BV, MBV, MTV, and realized kernel (RK). As the discussion in Barndorff-Nielsen et al. (2008), RK is a consistent estimator in the presence of MMN.[2]

2 A brief review of RK is provided in Appendix B.

Here, we experiment with the following four variations on the existence of noise and jumps: we conduct the simulations with (1) no jumps and noise, (2) MMN only, (3) jumps only, and (4) MMN and jumps.

First, we generate the artificial data from the following SV model.

$$dp(t) = \mu + \sigma(t)dW_1(t) \quad \sigma(t) = \exp(\beta_0 + \beta_1\tau(t)) \tag{9}$$

$$d\tau(t) = \theta\tau(t)dt + dW_2(t) \quad \mathrm{Corr}(dW_1, dW_2) = \rho, \tag{10}$$

where ρ is a leverage parameter. Simulating with this model is common in the literature; for example, see Barndorff-Nielsen et al. (2008) and Podolskij and Vetter (2009). Following the above studies, we set the parameter values of (5) and (6) as μ = 0.03, β_0 = 0.3125, β_1 = 0.12, θ = -0.025, ρ and = -0.3. We set the time interval to [0, 1] for simplicity. Sample paths of equation (5) are generated using Euler–Maruyama discretization with time step 1/ 23400; thus, [0, 1] spans 6.5 hours (from 9:30 to 16:00). Then, we construct sparse sampled returns as $r_i = p_i - p_{i-1}$, and compute the bias and mean squared error (MSE) of all six estimates for M = 39, 78, 130, 390, 780, 2,340, 4,680, and 23,400. For example, the case of 1-minute returns is presented by M = 390 in this setting. We compute the bias and MSE of MTV with α = 0.5.[3] Thus, $m = n_i = \sqrt{N}$ for all i.[4] Following the simulation setting of Podolskij and Vetter (2009), we set c_1 = 0.125 and c_2 = 2 for MBV and MTV in all cases. The results are summarized in Table 1.

Second, we examine the performance of the estimation in the presence of MMN.

$$p_i^* = p_i + u_i, \quad u_i \sim N(0, \omega^2), \tag{11}$$

where ω^2 is the variance of the noise. Following Podolskij and Vetter (2009), we set ω^2 = 0.001. The volatility model and its parameters are the same as in the no MMN case. The results are reported in Table 2.

Third, we examine the performance of the IV estimation in the presence of jumps.

$$dp(t) = \mu + \sigma(t)dW_1(t) + \zeta(t)dq(t), \tag{12}$$

where $\zeta(t)$ is the jump size, which we set as $N(0, h)$-distributed with h = 0.20, where the value of h is chosen to give a jump contribution of 14% of IV on average, based on the empirical results of the S&P 500 stock index analyzed by Andersen et al. (2007). $dq(s)$ is a

3 We also compute TRV with some different settings of α, and α = 0.5 gives the most efficient and unbiased estimates in our simulation comparison. Hence, we only report the results with this setting in this paper.

4 If $\sqrt{N}$ is not an integer, we set $m = [\sqrt{N}]$, $n_1 = M - (m-1)[N/m]$, and $n_i = [N/m]$ for i = 2, ..., m, where $[\cdot]$ denotes the floor function.

counting process with intensity λ, which we set as $\lambda = 1$. The volatility model and its parameters are the same as in the no jump case. The results are reported in Table 3.

Finally, we examine the performance of the simulation in the presence of MMN and jumps. This is the most practical setting: therefore, we focus on this situation in this paper. The results are reported in Table 4.

3.2 Simulation Results

Table 1 reports the biases and RMSEs of all five estimates for the no jump and no noise case. The biases are generally very low and seem to be lower for higher sampling frequencies. The RMSEs are lower if higher sampling frequencies are used, which is as expected. From Table 1, we can conclude that, as anticipated, RV performs best when there are no obstacles. In addition, we observe that our estimator is more efficient than MBV in terms of the RMSE, which corroborates the asymptotic theoretical result.

Table 1. Monte Carlo comparison of the biases and RMSEs of RV, BV, RK, MBV, and TRV for the SV model (no MMN and no jumps). The number of replications is 8,000.

	Bias in $\widehat{IV}$					RMSE of $\widehat{IV}$				
M	RV	BV	RK	MBV	MTV	RV	BV	RK	MBV	MTV
78	0.002	-0.014	0.035	2.033	-0.158	0.099	0.113	0.431	1.042	0.372
130	0.002	-0.005	0.015	1.919	-0.362	0.076	0.086	0.335	0.970	0.329
390	0.000	-0.004	0.008	1.488	-0.153	0.044	0.050	0.188	0.741	0.255
520	0.003	-0.001	0.008	1.440	-0.317	0.038	0.044	0.163	0.699	0.239
780	0.001	0.001	0.007	1.466	-0.154	0.031	0.035	0.133	0.672	0.212
2340	0.000	0.000	0.002	1.104	0.031	0.018	0.020	0.077	0.506	0.172
4680	0.000	0.000	0.001	1.023	0.053	0.012	0.014	0.055	0.453	0.146
23400	0.000	0.000	0.001	0.700	-0.038	0.006	0.006	0.024	0.304	0.094

Next, we show the simulation result with MMN. To focus on the effect of MMN for IV estimation, we generate data with MMN and without jumps. We observe that RV and BV have serious bias problems if noise exists. The biases of RK, MBV, and MTV become smaller because they are designed as being robust to noise, whereas RV and BV have positive biases when M is large. This corroborates the theoretical result.

Table 3 reports the biases and RMSEs of all estimates with jumps. We observe that RV, RK, and MBV have serious bias problems if jumps exist. The biases of BV, MBV, and

Table 2. Monte Carlo comparison of the biases and RMSEs of RV, BV, RK, MBV, and TRV for the SV model with MMN. The number of replications is 8,000.

	Bias in $\widehat{IV}$					RMSE of $\widehat{IV}$				
M	RV	BV	RK	MBV	MTV	RV	BV	RK	MBV	MTV
78	154.0	173.5	2.160	2.591	-0.310	49.6	56.2	3.567	1.989	1.229
130	257.7	290.2	2.253	2.033	-1.433	82.5	93.2	4.017	1.696	1.030
390	777.8	876.5	1.839	1.642	-0.737	246.9	278.5	5.731	1.287	0.793
520	1038	1171	2.058	1.574	-1.277	329.3	371.5	6.592	1.203	0.767
780	1559	1759	2.111	2.342	-0.423	494.0	557.6	7.613	1.311	0.674
2340	4679	5277	3.075	1.535	-0.080	1481	1670	13.08	0.941	0.535
4680	9361	10557	1.679	1.721	0.189	2961	3340	18.26	0.880	0.459
23400	46807	52789	1.211	1.212	-0.098	14803	16695	39.55	0.597	0.299

Table 3. Monte Carlo comparison of the biases and RMSEs of RV, BV, RK, MBV, and TRV for the SV model with jumps. The number of replications is 8,000.

	Bias in $\widehat{IV}$					RMSE of $\widehat{IV}$				
M	RV	BV	RK	MBV	MTV	RV	BV	RK	MBV	MTV
78	0.405	0.187	0.442	2.218	0.139	0.212	0.146	0.484	1.103	0.467
130	0.404	0.154	0.418	2.193	-0.108	0.200	0.113	0.390	1.062	0.381
390	0.408	0.103	0.410	2.104	0.229	0.187	0.067	0.265	0.998	0.336
520	0.405	0.085	0.411	1.880	-0.034	0.185	0.058	0.245	0.867	0.267
780	0.407	0.075	0.410	1.618	-0.010	0.183	0.048	0.226	0.720	0.227
2340	0.407	0.045	0.407	1.217	0.249	0.182	0.028	0.196	0.539	0.224
4680	0.407	0.033	0.406	1.173	0.371	0.181	0.020	0.188	0.494	0.229
23400	0.406	0.015	0.407	0.835	0.140	0.180	0.009	0.182	0.342	0.129

MTV become smaller, whereas RV and RK have similar positive biases for all M. This corroborates the asymptotic result. BV is the best in terms of RMSE because it was designed for this setting. Importantly, we observe that our estimator outperforms MBV; it is the best in terms of RMSE in all cases.

Table 4 reports the case with MMN and jumps. This is the most important group of settings in our simulation. We report the biases and RMSEs once again. The results in Table

Table 4. Monte Carlo comparison of the biases and RMSEs of RV, BV, RK, MBV, and TRV for the SV model with MMN and jumps. The number of replications is 8,000.

	Bias in $\widehat{IV}$					RMSE of $\widehat{IV}$				
M	RV	BV	RK	MBV	MTV	RV	BV	RK	MBV	MTV
78	154.6	173.9	2.430	2.710	-0.134	49.9	56.4	3.581	1.998	1.275
130	259.1	291.9	2.505	2.299	-1.147	82.9	93.7	4.027	1.762	1.042
390	779.0	878.0	2.186	2.233	-0.303	247.3	279.0	5.760	1.471	0.827
520	1038.3	1171	2.747	2.004	-0.989	329.3	371.5	6.651	1.337	0.775
780	1558	1756	2.494	2.536	-0.186	493.5	556.7	7.655	1.355	0.686
2340	4674	5269	3.170	1.740	0.227	1478	1667	13.05	0.977	0.580
4680	9359	10555	3.113	1.885	0.577	2960	3339	18.07	0.923	0.533
23400	46807	52789	1.599	1.385	0.118	14803	16695	39.53	0.643	0.324

4 show that RV, BV, and RK have serious bias problems because of either MMN or jumps, or both. The biases of MBV and MTV become smaller. It is important to note that MTV outperforms MBV in terms of RMSE in all cases, and this finding agrees with the asymptotic theoretical result.

4. An Empirical Application

4.1 Data and IV Estimation

Here, we show an empirical application of our estimator. As mentioned in the Introduction, we compare forecasting accuracy performance by employing our estimator and alternatives as explanatory variables in the forecasting model. It should be noted that the purpose of this application is to find the best estimator as the explanatory variable for predicting volatility; it is not to find the best true volatility proxy. Following the literature, we first set RV with a five-minute interval ($RV_{t,5m}$) as a volatility proxy.

We use high-frequency data from the Nikkei 225 Stock Index during the period from February 6th, 2004 through December 28th, 2007 (983 days). In Table 5, we report some basic statistics of daily returns R_t, $RV_{t,5m}$, and daily standardized returns.

In Table 5, $Q(20)$ and JB are the Ljung–Box test statistic with 20 degrees of freedom and the Jarque–Bera test statistic. Table 5 shows stylized features of financial time series documented in the literature. For example, it is difficult to say that the daily return distribution is normal, and the daily volatility series may have a long-run dependence. The last column shows the statistics for daily returns standardized by $RV_{t,5m}^{1/2}$. Compared to daily

Table 5. Basic statistics of daily returns, RV, and standardized daily returns of the Nikkei 225 Index.

	Mean	S.D.	Skewness	Kurtosis	$Q(20)$	JB
R_t	0.00	0.01	-0.36	4.40	24.65	101.92
$\log(RV_{t,5m}^{1/2})$	-4.73	0.38	-0.07	3.02	19790.4	0.83
$R_t/RV_{t,5m}^{1/2}$	0.08	1.03	-0.02	2.45	15.74	12.60

returns, standardized daily returns are more Gaussian. This is also consistent with the literature.

Figure 1 shows the estimation results using RV with a five-minute interval and MTV from the Nikkei 225 Index. The red lines indicate RV estimates. The blue lines indicate MTV outturns. From this figure, it is RV apparent that MTV are smaller than RV and less effective with the jumps. This may be attributed to the possibility that RV might be upper-biased by noise and jumps.

Figure 2 shows the estimation results using BV and MTV. The green and blue lines indicate BV and MTV outturns, respectively. We can observe that MTV is smaller than BV on average. We know both estimators are robust to jumps. That is, it is reasonable to think that the differences between these two estimates are caused by the market microstructure noise effect.

We show the estimation results for RK and MTV in Figure 3. The pink and blue lines indicate RK estimates and MTV outturns, respectively. This figure indicates that MTV is

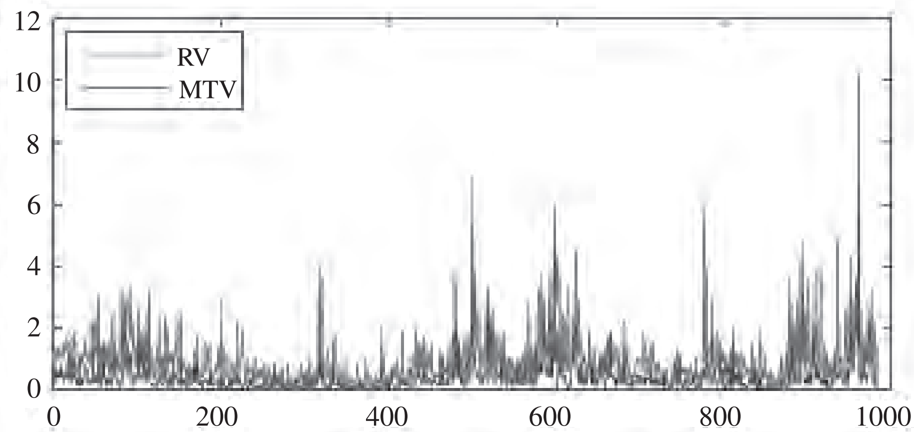

Figure 1. Estimation results of daily RV with five-min interval and MTV from the Nikkei 225 during the period from February 6, 2004 through December 28, 2007 (983 days).

less sensitive than RK in the presence of jumps. This is in line with the Monte Carlo result obtained in the former section.

4.2 Volatility Forecasting and Its Evaluation

Using the above estimation result, we compare performance in terms of the forecasting volatility. We employ the HAR model proposed by Corsi (2009) as our forecasting model. Although it is a simple model, Corsi (2009) showed that the HAR model can capture long memory, which is a well-known feature of the volatility process of the financial asset price. Define the s-day multi-period RV as $\overline{RV}^{1/2}_{i,(s)} = s^{-1}(RV^{1/2}_{i+1} + \cdots + RV^{1/2}_{i+s})$. Using $\overline{RV}^{1/2}_{i,i+s}$, the HAR-RV model for the one-day forecast is defined by

$$\log(RV^{1/2}_{t+1}) = \beta_0 + \beta_D \log(RV^{1/2}_t) + \beta_W \log(\overline{RV}^{1/2}_{t-5,(5)}) + \beta_M \log(\overline{RV}^{1/2}_{t-22,(22)}) + \varepsilon_{t+1}. \quad (13)$$

The HAR model uses one-day, one-week, and four-week (one-month) RVs as regressors. We also examine the forecast of the HAR model using BV, RK, MBV, and MTV as explanatory variables. The s-day multi-period of these estimates is defined in the same manner as for RV.

To evaluate volatility prediction performance, we calculate the out-of-sample error and compare its values. First, we compare the performance in terms of the RMSE, because this loss function is robust.[5] See Patton (2011) for a detailed discussion of robust loss functions. The RMSE of the one-day forecast is calculated as follows:

$$\mathrm{RMSE}_n = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\widehat{h}^{1/2}_t - QV^{1/2}_t\right)}, \quad (14)$$

where $h^{1/2}_{t,} = \exp(\log(h^{1/2}_t))$, and $h^{1/2}_t$ is the square root of forecasting volatility on the tth day. Here, we employ $RV_{t,5m}$ and RK_t as proxies of QV_t to calculate RMSE.

4.3 Forecasting Results

Table 6 presents the in-sample R^2 and out-of-sample RMSE. Bold entries indicate the best performance from among the five volatility estimators. Table 6 shows that regardless of the setting of the volatility proxy, MTV shows the highest R^2 and RMSE of the five volatility estimators. This result indicates that there is a possibility to increase the

5 Hansen and Lunde (2005) called the loss function robust if it gives a consistent ranking of forecasting when the volatility proxy has an error in the context of volatility forecasting.

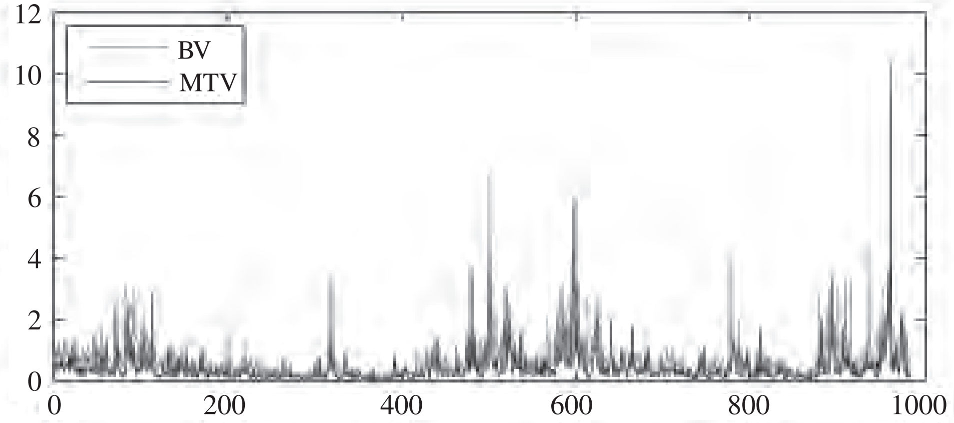

Figure 2. Estimation results of daily BV and MTV from the Nikkei 225 during the period from February 6, 2004 to December 28, 2007 (983 days).

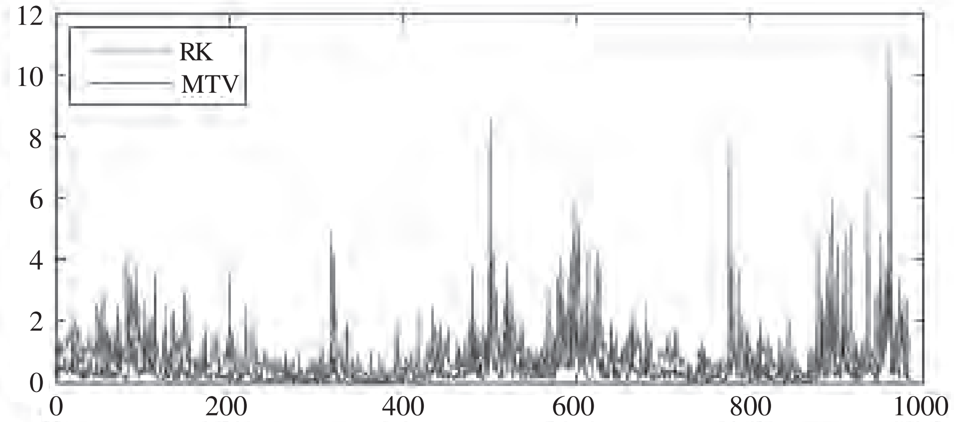

Figure 3. Estimation results of daily RK and MTV from the Nikkei 225 during the period from February 6, 2004 to December 28, 2007 (983 days).

Table 6. Comparing the forecasting performance of six estimators with the HAR model

Volatility Proxy	R^2 (T = 250)					RMSE (n = 711)				
	RV	BV	RK	MBV	MTV	RV	BV	RK	MBV	MTV
$RV_{t,\,5m}$	0.647	0.675	0.644	0.637	**0.672**	0.367	0.365	0.371	0.379	**0.362**
RK_t	0.64	0.667	0.639	0.635	**0.667**	0.415	0.412	0.418	0.423	**0.406**

forecasting performance by addressing jumps and noise. BV shows the second-best performance of the remaining estimators. BV is a jump-robust estimator; hence, this result meets the empirical findings reported in the literature. To evaluate improvement of the accuracy of volatility forecasts using MTV, we examine the Diebold–Mariano and West tests with the null hypothesis that predictive accuracy is equal. However, we find that the improvement of MTV is not statistically significant. Our empirical result using the Nikkei 225 as an indicator only shows the potential of our estimator for improving volatility prediction.

5. Concluding Remarks

The existence of market microstructure noise and jumps in the price process complicates integrated variance estimation. One answer to this problem is using MBV, which is derived by Podolskij and Vetter (2009). MBV is a consistent integrated variance estimator in the presence of these obstacles. In this paper, we extended the idea of MBV and showed that our new estimator achieves higher efficiency than the original one.

We conducted simulation experiments to investigate the finite sample properties of the new estimator in several settings. We found that our estimator provides the best forecasts in terms of RMSE when jumps and noise exist simultaneously.

In a brief empirical application, we examined the forecasting volatility of the Nikkei 225 Stock Index. We applied our estimator as an explanatory variable of the HAR model. Similar to our simulation, we employed various variance estimators that are robust for jumps or noise, or both effects. We reported that our proposed estimator shows the best performance in terms of RMSE. This result indicates that the performance of forecasting volatility can be improved by taking care 14 of jumps and noise simultaneously. However, the improvement was not statistically significant; therefore, more data should be collected and analyzed in order to reach a firmer conclusion. Finally, our results suggest that our estimator offers potential to improve forecasting performance. For example, it would be interesting to employ our estimator with more advanced models, such as realized-GARCH proposed by Hansen et al. (2010), or realized-SV by Takahashi et al. (2009). This examination would comprise a part of our future research.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: Including jump components in the measurement modeling and forecasting of return volatility. *Review of Economics and Statistics* 89, pp.701-720.

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2010. Parametric and nonparametric measure-ment of volatility. In: Ait-Sahalia, Y., Hansen, L.P. (Eds.), *Handbook of Financial Econometrics*. North Holland, Amsterdam, pp.67-138.

Andersen, T.G., Bollerslev, T., Medahi, N., 2011. Realized volatility forecasting and market microstructure noise. *Journal of Econometrics* 160, pp.220-234.

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

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

Barndorff-Nielsen, O.E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* 2, pp.1-48.

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

Barndorff-Nielsen, O.E., Shephard, N., Winkel, M., 2006. Limit theorems for multipower variation in the presence of jumps. *Stochastic Process and their Applications* 116, pp.796-806.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7, pp.174-196.

Forsberg, L., Ghysels, E., 2007. Why do absolute returns predict volatility so well? *Journal of Financial Econometrics* 5, pp.31-67.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2006. Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131, pp.59-96.

Hansen, P.R., Huang, Z., Shek, H.H., 2012. Realized GARCH: A complete model of returns and realized measures of volatility. *Journal of Applied Econometrics*, 27-6, pp.877-906.

Mancino, M.E., Sanfelici, S., 2008. Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise. *Computational Statistics & Data Analysis*, 52, pp.2966-2989.

McAleer, M., Medeiros M.C., 2008. Realized Volatility: A Review. *Econometric Reviews*, 27, pp.10-45.

Nagata, S., 2012a. Consistent estimation of integrated volatility using intraday absolute returns for SV jump diffusion processes. *Economics Bulletin*, 32, pp.306-314.

Nagata, S., 2012b. Predicting volatility with realized absolute values: Evidence from the Tokyo stock exchange. *Empirical Economics Letters*, 11-2, pp.551-558.

Patton, A. J., 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160, pp.246-256.

Podolskij, M., Vetter M., 2009. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli* 15, pp.634-658.

Takahashi, M., Omori, Y., Watanabe, T., 2009. Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics & Data Analysis* 53, pp.2404-2426.

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

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

Appendix A

We provide the proofs of Theorems 1 and 2 in this appendix A. The results derived by Podolskij and Vetter (2009) and Nagata (2012a) comprise the main body of this proof.

First, we recall the definition of MTV assuming n evenly spaced observed price-averaged returns are obtained in each period. Thus, we set $n_i = n = N^{\alpha}$ for all i, and $m = N^{1-\alpha}$, where N is the sample size of averaged returns defined as $K = c_1 M^{1/2}$, $N = M/\ c_2 K$, $c_1 > 0$, and $c_2 > 1$. It is obvious that the asymptotic results under this assumption also hold for the more general case $n_i = O(N)$.

Then, MTV for evenly spaced data from process Y with sample size M is given as

$$UMTV_{M,\alpha}(Y) = \frac{1}{n} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} |\bar{Y}_{n(i-1)+j}^{(K)}| \right)^2 ,$$

and

$$\bar{Y}_i^{(K)} = \frac{1}{\frac{M}{N} - K + 1} \sum_{j=\frac{(i-1)M}{N}}^{\frac{iM}{N}-K} (Y_{(j+K)M} - Y_{j/M}).$$

Proof of Theorem 1

Here, we consider to derive the proof of consistency of MTV. Following Podolskij and Vetter (2009), we first show consistency of MTV with the data from the process Y (no jump case) and discuss jump effect later.

We first introduce following quantities;

$$UMTV^M = \sum_{i=1}^{m} \eta_i^M, \qquad UMTV'^M = \sum_{i=1}^{m} \eta_i'^M,$$

$$UMTV'^M = \sum_{i=1}^{N} \eta_i''^M,$$

where η_i^M, $\eta_i'^M$, $\eta_i''^M$ are defined by

$$\eta_i^M = \frac{M^{1/2}}{c_1c_2n}\left\{\sum_{j=1}^{n} E[(\bar{y}_{a_i+j})^2|\mathcal{F}_{(a_i+j-1)/n}] + 2\sum_{j=1}^{n-1}\sum_{l=1}^{n-j} E[|\bar{y}_{a+j}||\bar{y}_{a_i+j+l}||\mathcal{F}_{(a_i+j-1)/n}]\right\},$$
$$\eta_i^{\prime M} = \frac{\mu_1^2}{c_1c_2n}\left\{\sum_{j=1}^{n}(\sigma^2_{a_i+j} + \nu_2\omega^2) + 2\sum_{j=1}^{n-1}\sum_{l=1}^{n-j}(\sigma^2_{a+j}\sigma_{a_i+j+l} + \nu_2\omega^2)\right\},$$
$$\eta_i^{\prime\prime M} = \frac{\mu_1^2}{c_1c_2}(\nu_1\sigma^2_{(i-1)/N} + \nu_2\omega^2)$$

with, $a_i = (i-1)n$.

Assuming Riemann integrability, we obtain

$$\frac{1}{N}\mathrm{UMTV}^{\prime\prime M} \xrightarrow{p} \mathrm{UIV} = \frac{\mu_1^2}{c_1c_2}\int_0^1 (\nu_1\sigma_s^2 + \nu_2\omega^2)ds$$

Therefore, now we consider to check the convergence following three steps sepalately to complete the ploof of Theorem 1,

$$\text{(step1)} \qquad UMTV_{M,\alpha}(y) - \frac{1}{N}UMTV^M \xrightarrow{p} 0,$$
$$\text{(step2)} \qquad \frac{1}{m}(UMTV^M - UMTV^{\prime M}) \xrightarrow{p} 0,$$
$$\text{(step3)} \qquad \frac{1}{N}(\frac{N}{m}UMTV^{\prime M} - UMTV^{\prime\prime M}) \xrightarrow{p} 0.$$

Following the discussions in Podolskij and Vetter (2009), we can prove step1 and step2 using Lemma 1 and lemma 2 of Podolskij and Vetter (2009), respectively. For checking step 3 property, we extend the proof of BN-S(2004) for their bi-power variation's case. No-jump case is completed. As stated in PV(2009), following the discution in Barndorff-Nielsen, Shephard and Winkel (2006), the same result also can be derived in the presence of jumps. Because $\omega^2 - \hat{\omega}^2 = O_p(\mathrm{M}^{1/2})$, the variance of the noise can be estimated consistently.

Therefore, if $0 < \alpha < 1$, then

$$MTV_{M,\alpha}(Z) = \frac{\mu_1^{-2}c_1c_2UMTV_{M,\alpha}(Z) - \nu_2\hat{\omega}^2}{\nu_1} \xrightarrow{p} \int_0^1 \sigma_s^2 ds.$$

Hence, the proof of Theorem 1 is completed. □

Proof of Theorem 2

We first derive the conditional asymptotic variance of the estimator. As shown by Podolskij and Vetter (2009), the joint distributions of $\bar{r}_1^*$, ..., $\bar{r}_N^*$ and Z_1, ..., Z_N are asymptotically equivalent, where $Z_i = \sigma_{\frac{(i-1)}{N}}\bar{W}_i + \bar{U}_i$, $\lim_{M\to\infty}\mathrm{M}^{1/4}\bar{W}_i \sim N(0, \nu_1)$, and $\lim_{M\to\infty}$

$M^{1/4}\bar{U}_i \sim N(0, \nu_2\omega^2)$.

Then,

$$UMTV_{N,\alpha} \overset{\mathcal{L}}{=} \frac{1}{n}\sum_{i=1}^{m}\left\{\sum_{j=1}^{n}(Z_{a+j})^2 + 2\sum_{j=1}^{n-1}\sum_{l=1}^{n-j}|Z_{a+j}||Z_{a+j+l}|\right\}.$$

That is, the conditional mean of *MTV* is then

$$\begin{aligned}
\mathrm{E}[UMTV_{M,\alpha}|\sigma^2] &= \frac{1}{nM^{1/2}}\sum_{i=1}^{m}\left\{\sum_{j=1}^{n}(E[|Z_j|^2|\sigma^2]) + 2\sum_{j=1}^{n-1}\sum_{l=1}^{n-j}(E[|Z_{a+j}||Z_{a+j+l}||\sigma^2])\right\} \\
&= \frac{1}{nM^{1/2}}\sum_{i=1}^{m}\left\{\sum_{j=1}^{n}(\nu_1^{(M)}\sigma_{a+j}^2 + \nu_2\omega^2)\right. \\
&\qquad \left. +2\mu_1^2\sum_{j=1}^{n-1}\sum_{l=1}^{n-j}(\nu_1^{(M)}\sigma_{a+j}\sigma_{a+j+l} + \nu_2\omega^2)\right\} + o(1/N), \qquad (15)
\end{aligned}$$

where

$$\nu_1^{(M)} = \nu_1 + \frac{(3-c_2)\wedge\frac{1}{c_2-1}}{(c_2-1)\sqrt{M}} + O(1/M).$$

The second equivalence in (17) is from the proof of Theorem 2 of Podolskij and Vetter (2009).

We now prove Theorem 2, letting

$$D := M^{1/4}(UMTV_{M,\alpha} - \mathrm{E}[UMTV_{M,\alpha}|\sigma^2]) = \sum_{i=1}^{m}\theta_i$$

where

$$\theta_i := \frac{1}{nM^{1/4}}\left\{\sum_{j=1}^{n}(\nu_1^{(M)}\sigma_{a+j}^2 + \nu_2\omega^2)v_{a+j} + 2\sum_{j=1}^{n_1}\sum_{l=1}^{n-j}(\nu_1^{(M)}\sigma_{a+j}\sigma_{a+j+l} + \nu_2\omega^2)w_{a+j,a+j+l}\right\}, \quad (16)$$

$v_i = |u_i|^2 - 1$, $w_{i,j} = |u_i||u_j| - \mu_1^2$, and $n_1 = n - 1$.

It is easy to obtain

$$\begin{aligned}
&\mathrm{E}[v_i] = 0, && \mathrm{var}[v_i] = 2, \\
&\mathrm{cov}[v_i, v_j] = 0, && \mathrm{E}[w_{i,j}] = 0, \\
&\mathrm{var}[w_{i,j}] = 1 - \mu_1^4, && \mathrm{cov}[w_{i,j}, w_{i,k}] = \mathrm{cov}[w_{i,j}, w_{k,i}] = \mu_1^2(1-\mu_1^2), \\
&\mathrm{cov}[w_{i,j}, w_{k,l}] = 0, && \mathrm{cov}[v_i, w_{j,i}] = \mathrm{cov}[v_i, w_{i,j}] = \mu_1^2, \\
&\mathrm{cov}[v_i, w_{j,k}] = 0.
\end{aligned}$$

Hence, $\mathrm{E}[D|\sigma^2] = 0$ and

$$\mathrm{var}[D|\sigma^2] = \frac{1}{n^2 M^{1/2}} \sum_{i=1}^{m} \mathrm{E}\left[\left(\sum_{j=1}^{n}(\sigma_{a+j}^2 + \nu_2\omega^2)v_{a+j} + 2\sum_{j=1}^{n_1}\sum_{l=1}^{n-j}(\sigma_{a+j}\sigma_{a+j+l} + \nu_2\omega^2)w_{a+j+l}\right)^2\right]$$

Because $N = M^{1/2}(c_1c_2)^{-1}$, we can show

$$\mathrm{var}[D|\sigma^2] = \frac{(c_1c_2)^{-1}}{mn^3}\Bigg(\mathrm{E}\left[v_i^2\right]\sum_{i=1}^{N}(\nu_1^{(M)}\sigma_i^4 + \nu_2\omega^2) + 4\mathrm{E}\left[w_{i,j}^2\right]\sum_{i=1}^{m}\sum_{j=1}^{n_1}\sum_{l=1}^{n-j}(\nu_1^{(M)}\sigma_{a+j}^2\sigma_{a+j+l}^2 + \nu_2\omega^2)$$

$$+ 4\mathrm{cov}\left[v_i, w_{i,j}\right]\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{\{1\le l\le n_1: l\neq j\}}(\nu_1^{(M)}\sigma_{a+j}^3\sigma_{a+l} + \nu_2\omega^2)$$

$$+8\mathrm{cov}\left[w_{i,j}, w_{i,l}\right]\sum_{i=1}^{m}\sum_{j=1}^{n_1}\sum_{\{1\le l\le s\le n_1: l\neq s\neq j\}}(\nu_1^{(M)}\sigma_{a+j}^2\sigma_{a+l}\sigma_{a+s} + \nu_2\omega^2)\Bigg).$$

It is easy to show that

$$\frac{1}{mn^3}\sum_{i=1}^{N}(\nu_1^{(M)}\sigma_i^4 + \nu_2\omega^2) = o(1/M),$$

$$\frac{1}{mn^3}\sum_{i=1}^{m}\sum_{j=1}^{n_1}\sum_{l=1}^{n-j}(\nu_1^{(M)}\sigma_{a+j}^2\sigma_{a+j+l}^2 + \nu_2\omega^2) = o(1/M), and$$

$$\frac{1}{mn^3}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{\{1\le l\le n_1: l\neq j\}}(\nu_1^{(M)}\sigma_{a+j}^3\sigma_{a+l} + \nu_2\omega^2) = o(1/M).$$

Thus,

$$\mathrm{var}[D|\sigma^2] = \frac{(c_1c_2)^{-1}}{N}8\mathrm{cov}\left[w_{i,j}, w_{i,l}\right]\frac{1}{n^2}\sum_{i=1}^{m}\sum_{j=1}^{n_1}\sum_{\{1\le l\le s\le n_1: l\neq s\neq j\}}(\sigma_{a+j}^2\sigma_{a+l}\sigma_{a+s} + \nu_2\omega^2) + o(M^{-1}).$$

Assuming Riemann integrability, we obtain the conditional asymptotic variance of D.

$$\lim_{M\to\infty}\mathrm{var}[D|\sigma^2] = \frac{1}{c_1c_2}V\int_0^t(\nu_1\sigma^2(s) + \nu_2\omega^2)^2 ds, \tag{17}$$

where

$$V = 4\mathrm{cov}[w_{i,j}, w_{i,l}] = 4\mu_1^2(1-\mu_1^2).$$

Since $\omega^2 - \hat{\omega}^2 = O_p(\mathrm{M}^{1/2})$, the error of this estimation does not affect the form of (17).

Finally, to consider the convergence in distribution of UMTV, we recall the sequence θ_i defined in eq (16). Under the assumption A.1 A.2, and A.3 are hold, we can directly apply lemma 3 by Podolskij and Vetter (2009) for $\sum_{i=1}^{m} \theta_i$ to prove its stable convergence property. □

Appendix B

Here, we briey review realized kernel (RK), which is a consistent estimator in the presence of MMN. RK is defined as

$$RK = \gamma_0 + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)(\gamma_h + \gamma_{-h}). \tag{18}$$

where $\gamma_h = \sum_{i=1}^{n} r_{t,i} r_{t,i-h}$ with $h = -H, \ldots,$ H, and $k(x)$ is a kernel function. In our simulation, following Podolskij and Vetter (2009), we choose a modified Tukey-Hamming kernel and $k(x) = 1 - \cos\pi(1-x)^2$, $H = \hat{c}\zeta M^{1/2}$, and $\zeta^2 = \omega^2/\sqrt{IQ}$. c is

$$c = \sqrt{\xi \frac{k_1}{k_0}\left\{1 + \sqrt{1 + \frac{3k_0 k_2}{\xi k_1^2}}\right\}}, \tag{19}$$

where $\hat{\xi} = IV / \sqrt{IQ}$, $k_0 = 0.219$, $k_1 = 1.71$, and $k_2 = 41.7$. ω^2 can be estimated by $\hat{\omega}^2 = \frac{1}{2M} \sum_{i=1}^{M} r_i^2$ at the highest frequencies. IQ is estimated by $RQ = \frac{M}{3} \sum_{i=1}^{M} r_i^4$ (realized quadracity) with low-frequency returns such as 15-minute returns. See Barndorff-Nielsen et al. (2008) for a detailed discussion of RK.