KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE

OF

ECONOMIC RESEARCH

http://www.kier.kyoto-u.ac.jp/index.html

Discussion Paper No. 634

“Finite Sample Analysis of Weighted Realized Covariance with Noisy Asynchronous Observations”

Taro Kanatani

June 2007

KYOTO UNIVERSITY

KYOTO, JAPAN

Finite Sample Analysis of Weighted Realized Covariance with Noisy Asynchronous Observations*

Taro Kanatani[†‡]

JSPS Research Fellow, Institute of Economic Research, Kyoto University

Yoshida Honmachi, Sakyo-Ku, Kyoto 6068501, Japan

June 2007

Abstract

In this paper, we provide a framework to evaluate finite sample MSE of several realized covariance estimators when using nonsynchronous observations contaminated with microstructure noise. This framework enables us to examine different estimators. We propose some estimators as an application of the framework.

Keywords: High frequency data; Weighted realized covariance; Nonsynchronous (asynchronous) observation; Microstructure noise

JEL Classification: C14; C32; C63

---

*The author would like to thank seminar participants at the University of Florence, Hiroshima University, and Kyoto University for their comments. This research was supported by Grant-in-Aid for JSPS fellows.

†E-mail: `kanatani@kier.kyoto-u.ac.jp`

‡URL: `http://www.geocities.jp/kanatanit/` (The latest version is available here.)

1 Introduction

Recent availability of high-frequency data has been making realized-type estimators for volatility increasingly attractive. However, market microstructure noise has prevented researchers from using highest-frequency data such as transaction or quote data. Researchers had been compelled to choose moderate data frequency at which the effects of the noise might be negligible. For instance, Andersen et al. (2003), the most influential work among realized volatility studies, selected 15-min returns of foreign exchange rates. It was natural to desire a rigorous theory to select the data frequency. Therefore, Bandi and Russell (2005a) provided the optimal frequency for realized volatility based on finite sample MSE. On the other hand, for asymptotic theory, Zhang et al. (2005) first provided a consistent realized estimator in the presence of noise, which is called *two-scale estimator* (TSE). The *realized kernel* developed by Barndorff-Nielsen et al. (2006) unified several estimators including TSE and presented discussion of the asymptotic efficiency for different kernels.

Compared to volatility estimation, co-volatility has not been well studied. Only Bandi and Russell (2005b) and Griffin and Oomen (2006) derived the optimal frequency for the realized covariance in the presence of noise. One reason why we cannot apply the theories on realized estimators of volatility to those of co-volatility is nonsynchronicity of observations. Hayashi and Yoshida (2005) proposed an unbiased and consistent covariance estimator for asynchronous observation in the absence of noise. The estimator is still unbiased for independent noise; therefore, Griffin and Oomen (2006) examined how many observations should be used or discarded under a somewhat restricted situation in which volatilities are constant and prices are observed in a Poisson random manner. To handle nonsynchronicity and microstructure noise together, in this paper we examine *weighted realized covariance* (WRC) which was proposed as a general estimator by Kanatani (2004) under a more general situation. We provide a framework to evaluate a finite sample MSE of WRC to examine existing estimators and propose new estimators.

The remainder of the paper is organized as follows. In Section 2, we present assumptions on the true price process and microstructure noise. In Section 3, we calculate the finite sample MSE of WRC. Section 4 presents how to evaluate the MSE and some examples of weight functions are given in Section 5. In Section 6, we confirm the theory through a Monte Carlo study and Section 7 concludes the paper.

2 Assumptions

In this paper, we specifically examine a methodology for measuring covariance between financial assets in the presence of market microstructure noise.

We consider a multi-dimensional vector of logarithmic asset price $p(t)$ for $t \geq 0$. Without loss of generality, we set the dimension of p as 2. We assume that p is a continuous stochastic volatility semimartingale ($\mathcal{SVSM}^c$) with zero drift.[1]

$$p(t) = \int_0^t \Sigma(u) dz(u),$$

where Σ has elements that are all cadlag and z is a vector standard Brownian motion. We set the drift vector as 0 for the purpose of simplification.[2] The instantaneous or spot covariance matrix is defined as

$$\Omega(t) \equiv \Sigma(t)\Sigma(t)';$$

that is to say, cross volatility between the first and second asset is denoted as the $(1, 2)$ or $(2, 1)$ element of Ω:

$$\omega_{12}(t) = \sigma_{11}(t)\,\sigma_{21}(t) + \sigma_{12}(t)\,\sigma_{22}(t)\,.$$

[1]See Barndorff-Nielsen and Shephard (2004) for the $\mathcal{SVSM}^c$,

[2]This simplification is acceptable not only because it indicates an efficient market in financial economics, but also because, mathematically, the martingale component swamps the predictable portion over short time intervals.

Our target is not spot covariance, but integrated covariance over $[0, T]$:

$$IC \equiv \int_0^T \omega_{12}(t)\, dt.$$

For the case of $i = j$, we call $IV_i \equiv \int_0^T \omega_{ii}(t)dt = \int_0^T (\sigma_{i1}(t)^2 + \sigma_{i2}(t)^2)dt$ the integrated variance. For estimation of the integrated covariance matrix, the following quadratic variation formula is a theoretical basis for using the sum of the outer product of the return vector. If all assets are synchronously observed at simultaneous points

$$0 = t_0 < t_1 < \cdots < t_N <= T,$$

and

$$\lim_{N\to\infty} \max_n (t_n - t_{n-1}) = 0$$

then

$$p - \lim_{N\to\infty} \sum_{n=1}^{N} (p(t_n) - p(t_{n-1}))(p(t_n) - p(t_{n-1}))' = \int_0^T \Omega(t)dt.$$

See e.g. Barndorff-Nielsen and Shephard (2004).

However, each ith asset price is observed nonsynchronously at different time points

$$0 = t_{0_i} < t_{1_i} < \cdots < t_{N_i} <= T.$$

Usually in practice, nonsynchronous data are transformed into synchronous data using some data manipulation scheme such as previous-tick interpolation. However, such manipulation should cause a bias on the realized covariance estimator, which is known as the *Epps effect*. See e.g. Kanatani and Renò (2007) or Zhang (2006). Using raw data, Hayashi and Yoshida (2005) proposed a new estimator, which can solve the nonsynchronous bias problem in the absence of the observation error.

More crucially, the efficient prices are considered to be contaminated by market microstructure noise. The noise is interpreted as an observation error in the recent literature on realized volatility. Define the observed logarithmic asset price as

$$p_i^o(t_{n_i}) \equiv p_i(t_{n_i}) + e_i(t_{n_i}),$$

where $e_i(t)$ is independent with any other variables and $E(e_i(t)) = 0, V(e_i(t)) = \sigma_i^2$.[3] Define observed return as

$$r_i^o(t_{n_i}) \equiv r_i(t_{n_i}) + u_i(t_{n_i}),$$

where $r_i^o(t_{n_i}) \equiv p_i^o(t_{n_i}) - p_i^o(t_{n_i-1})$, $r_i(t_{n_i}) \equiv p_i(t_{n_i}) - p_i(t_{n_i-1})$, and $u_i(t_{n_i}) \equiv e(t_{n_i}) - e(t_{n_i-1})$. Notice that $r_i(t_{n_i})$ and $u_i(t_{n_i})$ have zero mean, but have different variances of $\int_{t_{n_i-1}}^{t_{n_i}} \omega_{ii}(t)dt$ and $2\sigma_i^2$, which are respectively at orders of $O(t_{n_i} - t_{n_i-1})$ and $O(1)$. Therefore, under a high-frequency situation where $t_{n_i} - t_{n_i-1}$ is sufficiently small, the true return $r_i(t_{n_i})$ is overwhelmed by the noise term $u_i(t_{n_i})$.

We concentrate on measuring the integrated covariance from a given observation and do not make any hypothesis on the structure of the underlying probability space. Therefore, our analysis is conditioned on $\{\Sigma(t)\}$ and $\{t_{n_i}\}$, in other words, we can consider $\Sigma(t)$ and t_{n_i} as deterministic functions.

3 MSE of weighted realized covariance

In this section, we investigate the *weighted realized covariance* (WRC), which was proposed in Kanatani (2004). In fact, WRC is the general form of realized estimators nesting low frequency RV, subsampling methods, TSE, Fourier estimator, and Realized kernels; it enables us to unify the discussion

[3] Voev and Lunde (2007) consider a more general type of noise and examine its effect on the Hayashi-Yoshida estimator. Ubukata and Oya (2007) propose how to test correlation of noises between different assets.

related to all of them. We specifically examine the finite sample MSE-based analysis in this paper.

The WRC is defined as follows.

$$WRC = (r_1^o)' W r_2^o,$$

where $r_i^o = (r_i^o(0), ..., r_i^o(t_{n_i}), ..., r_i^o(T))'$ and W is $N_1 \times N_2$ matrix. WRC is decomposed into

$$WRC = r_1' W r_2 + r_1' W u_2 + u_1' W r_2 + u_1' W u_2,$$

where $r_i = (r_i(0), ..., r_i(t_{n_i}), ..., r_i(T))'$ and $u_i = (u_i(0), ..., u_i(t_{n_i}), ..., u_i(T))'$. The first term represents the WRC in the absence of the noise. For convenience, we introduce some notation. We denote the elements in W as

$$\begin{cases} w_{n_1 n_2}^{diag} & \text{if } (t_{n_1-1}, t_{n_1}] \cap (t_{n_2-1}, t_{n_2}] \neq \emptyset, \\ w_{n_1 n_2}^{off} & \text{otherwise.} \end{cases}$$

We denote the piecewise integrated covariance as

$$IC_{n_1 n_2} = \begin{cases} \displaystyle\int_{\max\{t_{n_1-1}, t_{n_2-1}\}}^{\min\{t_{n_1}, t_{n_2}\}} \omega_{12}(t) dt & \text{if } (t_{n_1-1}, t_{n_1}] \cap (t_{n_2-1}, t_{n_2}] \neq \emptyset, \\ 0 & \text{otherwise,} \end{cases}$$

and also denote the piecewise integrated variance as $IV_{n_i} = \int_{t_{n_i-1}}^{t_{n_i}} \omega_{ii}(t) dt$. Using the properties of the independent increment of Brownian motion and uncorrelated noise, the bias of WRC is calculated as

$$E[WRC - IC] = \sum_{n_1, n_2} (w_{n_1 n_2}^{diag} - 1) IC_{n_1 n_2}. \tag{3.1}$$

Therefore, WRC is unbiased if $w_{n_1 n_2}^{diag} = 1$. In the special case of $w_{n_1 n_2}^{diag} = 1$ and $w_{n_1 n_2}^{off} = 0$, WRC is equivalent with Hayashi and Yoshida (2005)'s estimator. The independent noise does not affect the expectation of WRC, the bias arises from nonsynchronicity only. However, the noise does affect the MSE, which is calculated as

$$\begin{aligned} MSE &= E[WRC - IC]^2 \\ &= \underbrace{E[r_1' W r_2 - IC]^2}_{A} + \underbrace{E[r_1' W u_2]^2}_{B} + \underbrace{E[u_1' W r_2]^2}_{C} + \underbrace{E[u_1' W u_2]^2}_{D}, \end{aligned} \tag{3.2}$$

where

$$A = \sum_{n_1,n_2} \left(w_{n_1 n_2}^{diag} IC_{n_1 n_2} \right)^2 + \sum_{n_1,n_2} w_{n_1 n_2}^2 IV_{n_1} IV_{n_2} + \left\{ \sum_{n_1,n_2} (w_{n_1 n_2}^{diag} - 1) IC_{n_1 n_2} \right\}^2,$$

$$B = 2\sigma_2^2 \sum_{n_1,n_2} w_{n_1 n_2} (w_{n_1 n_2} - w_{n_1 n_2 - 1}) IV_{n_1},$$

$$C = 2\sigma_1^2 \sum_{n_1,n_2} w_{n_1 n_2} (w_{n_1 n_2} - w_{n_1 - 1 n_2}) IV_{n_2},$$

$$D = \sigma_1^2 \sigma_2^2 \sum_{n_1,n_2} w_{n_1 n_2} \left\{ 4 w_{n_1 n_2} + 2 w_{n_1 - 1 n_2 - 1} + 2 w_{n_1 - 1 n_2 + 1} - 4 (w_{n_1 - 1 n_2} + w_{n_1 n_2 - 1}) \right\},$$

$w_{n_1 n_2} = 0$ if $n_i \leq 0$ or $n_i \geq N_i$.

See Appendix for details of calculation. In those equations, A is the MSE of WRC in the absence of noise. In the absence of noise, Kanatani (2004) derived the optimal weight that minimizes A.

4 Feasible evaluation of MSE

Since WRC is a bit too general to minimize the MSE, we need to select a specific form of weight function. In the next section, we present several examples of one parameter function. Furthermore, for simplicity, we limit our discussion to unbiased estimators; in other words, we set $w_{n_1 n_2}^{diag} = 1$. This enables avoidance of evaluating the bias and $\sum \left(w_{n_1 n_2}^{diag} IC_{n_1 n_2} \right)^2$. If $w_{n_1 n_2}^{diag} = 1$, the bias is zero, and $\sum \left(w_{n_1 n_2}^{diag} IC_{n_1 n_2} \right)^2$ is unknown, but it is constant. Therefore, we do not need to evaluate piecewise integrated covariance $IC_{n_1 n_2}$, which is difficult to estimate.

We still need the variance of noise σ_i^2 and the piecewise integrated volatility IV_{n_i} to evaluate the MSE (3.2). It is difficult to estimate the piecewise integrated volatility IV_{n_i} as well as $IC_{n_1 n_2}$. To avoid evaluating IV_{n_i}, we impose the assumption: "Volatility does not change so much over $[0, T]$." This assumption is described in Bandi and Russell (2006). Under this assumption, the following approximation is valid.[4]

$$IV_{n_i} \approx \frac{IV_i \Delta t_{n_i}}{T}, \tag{4.1}$$

where $\Delta t_{n_i} = t_{n_i} - t_{n_i - 1}$. For estimations of σ_i^2 and IV_i, several established methods exist, see e.g. Bandi and Russell (2005a), Zhang et al. (2005). The estimation methods of those parameters are out of scope of this paper, therefore we treat them as known parameters.

Now the minimization of the finite sample MSE of WRC reduces to

$$\min_{\theta}(A' + B' + C' + D), \tag{4.2}$$

where

$$\begin{aligned}
A' &= T^{-2} IV_1 IV_2 \sum w_{n_1 n_2}^2 \Delta t_{n_1} \Delta t_{n_2}, \\
B' &= 2T^{-1} IV_1 \sigma_2^2 \sum w_{n_1 n_2}(w_{n_1 n_2} - w_{n_1 n_2 - 1}) \Delta t_{n_1}, \\
C' &= 2T^{-1} IV_2 \sigma_1^2 \sum w_{n_1 n_2}(w_{n_1 n_2} - w_{n_1 - 1 n_2}) \Delta t_{n_2}, \\
w_{n_1 n_2} &= \begin{cases} 1 & \text{if } (t_{n_1 - 1}, t_{n_1}] \cap (t_{n_2 - 1}, t_{n_2}] \neq \emptyset, \\ f(t_{n_1}, t_{n_2}; \theta) & \text{otherwise.} \end{cases}
\end{aligned}$$

In the next section, we see concrete examples of the weight function.

5 Examples of weight function

The Fourier estimator was proposed originally by Malliavin and Mancino (2002) in a different form of WRC. However, Kanatani (2004) shows that

[4] Bandi and Russell (2006) use the approximation $IV_{n_i} \approx IV_i / N_i$ to derive the optimal frequency based on a finite sample MSE of the subsampling estimator. However, such approximation implies the assumption that the "time difference does not change so much." Therefore, we use a less-restricted approximation (4.1) because we do not need to derive the optimal frequency explicitly.

this estimator is proved to be written in the form of WRC with the following weight:

$$w_{n_1n_2} = \begin{cases} 1 & \text{if } t_{n_1} = t_{n_2}, \\ \frac{\sin\frac{(n+1)(t_{n_1}-t_{n_2})}{2}\cos\frac{n(t_{n_1}-t_{n_2})}{2}}{n\sin\frac{(t_{n_1}-t_{n_2})}{2}} & \text{otherwise,} \end{cases} \tag{5.1}$$

where n is the number of Fourier coefficients. Note that (3.1) and (5.1) imply that this estimator is biased in finite samples. In this paper, we examine the bias corrected version of the Fourier estimator, which has the weight matrix:

$$w_{n_1n_2} = \begin{cases} 1 & \text{if } (t_{n_1-1}, t_{n_1}] \cap (t_{n_2-1}, t_{n_2}] \neq \emptyset, \\ \frac{\sin\frac{(n+1)(t_{n_1}-t_{n_2})}{2}\cos\frac{n(t_{n_1}-t_{n_2})}{2}}{n\sin\frac{(t_{n_1}-t_{n_2})}{2}} & \text{otherwise.} \end{cases}$$

We call the WRC with this weight the *Modified Fourier Estimator* (MFE). Now we can select an optimal number of Fourier coefficients of MFE based on finite sample MSE.

The next candidate of the unbiased weight function is

$$w_{n_1n_2} = \begin{cases} 1 & \text{if } (t_{n_1-1}, t_{n_1}] \cap (t_{n_2-1}, t_{n_2}] \neq \emptyset, \\ \exp\left(-\left(\frac{t_{n_1}-t_{n_2}}{h}\right)^2\right) & \text{otherwise,} \end{cases} \tag{5.2}$$

where $h > 0$. We name the WRC with this weight the *Error Function weight estimator* (EF). The h functions as a bandwidth to control how the estimator should account for the noise. In extreme cases, when h goes to zero, for given $\{t_{n_i}\}_{n_i=1}^{N_i}$, all elements of $w_{n_1n_2}^{off}$ go to zero, then WRC reduces to a Hayashi-Yoshida estimator. On the other hand, when h goes to infinity, all elements go to unity, and WRC reduces to $(p_1^o(T) - p_1^o(0))(p_2^o(T) - p_2^o(0))$. This estimation means that all data $\{p_i(t_i)\}_{n_i=1}^{N_i-1}$ are discarded. In moderate cases, through the minimization (4.2), we can select a moderate value of the optimal h.

Our framework of the minimization (4.2) is also applicable to the kernels that are used in Barndorff-Nielsen et al. (2006). Although situations are different between variance and covariance, in other words, for synchronicity

Table 1: Kernels

	$k(x)$
Bartlett	$1-x$
Epanechnikov	$1-x^2$
Parzen	$1-6x^2+6x^3$ $(0 \leq x \leq 1/2)$
	$2(1-x)^3$ $(1/2 < x \leq 1)$
Tukey-Hanning	$(1+\cos(\pi x))/2$
Mod. Tukey-Hanning	$(1-\cos \pi(1-x)^2)/2$

and for nonsynchronicity, we select and apply several kernels listed in Table 1. Barndorff-Nielsen et al. (2006) use those kernels as functions of the lead and lag numbers, although we slightly modify the kernels as functions of time difference $|t_{n_1} - t_{n_2}|$.

$$w_{n_1 n_2} = \begin{cases} 1 & \text{if } (t_{n_1-1}, t_{n_1}] \cap (t_{n_2-1}, t_{n_2}] \neq \emptyset, \\ k\left(\frac{|t_{n_1} - t_{n_2}|}{H}\right) & \text{if } (t_{n_1-1}, t_{n_1}] \cap (t_{n_2-1}, t_{n_2}] = \emptyset \text{ and } |t_{n_1} - t_{n_2}| < H, \\ 0 & \text{otherwise.} \end{cases} \tag{5.3}$$

Therein, $H > 0$. Unlike the error function weight, these kernels have compact supports. Not only can we select each optimal parameter of each weight function; we can also decide which function is the best among them by comparing $A' + B' + C' + D$ in (4.2).

As described above, Hayashi and Yoshida (2005) proposed an unbiased estimator for nonsynchronous true observations; it has the following weight:

$$w_{n_1 n_2} = \begin{cases} 1 & \text{if } (t_{n_1-1}, t_{n_1}] \cap (t_{n_2-1}, t_{n_2}] \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

To mitigate the effect of the noise, Griffin and Oomen (2006) proposed a lower-frequency version of the Hayashi-Yoshida estimator (LHY) with weight:

$$w_{n_1n_2} = \begin{cases} 1 & \text{if } (t_{k(n_1-1)}, t_{kn_1}] \cap (t_{k(n_2-1)}, t_{kn_2}] \neq \emptyset, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a positive integer. Griffin and Oomen (2006) calculate the MSE to optimize k under the condition of constant volatilities and Poisson random sampling. Now we can also select optimal k by minimizing the MSE (3.2) under more general settings. However, we must unfortunately evaluate piecewise integrated covariance $IV_{kn_1kn_2}$ because $\sum(IV_{kn_1kn_2}w^{diag}_{kn_1kn_2})^2$ is not constant for different k. We cannot evaluate the MSE of LHY and compare it with that of other previously mentioned estimators, unless we evaluate the covariance process itself.

6 Monte Carlo study

We performed a Monte Carlo simulation to confirm our theory. In our simulation, the efficient price process is generated by

$$\begin{pmatrix} dp_1(t) \\ dp_2(t) \end{pmatrix} = \begin{pmatrix} \sigma_{11}(t) & 0 \\ \sigma_{21}(t) & \sigma_{22}(t) \end{pmatrix} \begin{pmatrix} dz_1(t) \\ dz_2(t) \end{pmatrix}, \ 0 \leq t \leq T,$$
$$d\sigma_{ij}(t) = \kappa(\theta - \sigma_{ij}(t))\,dt + \gamma dz_{ij}(t), \ i, j = 1, 2,$$

where $\kappa = 0.1$, $\theta = 1$, $\gamma = 0.1$, $T = 1$(day). However, we generate a proxy of the process with a time-step of $\Delta = 1/60 \times 60 \times 4.5$ (one second precision for Japanese stock exchanges). Time differences are drawn from an exponential distribution:

$$F(t_{n_i} - t_{n_i-1}) = 1 - \exp\{-\lambda_i(t_{n_i} - t_{n_i-1})\}, \ i = 1, 2,$$

in which $F(\cdot)$ denotes a cumulative distribution function, $\lambda_i = 1/60\Delta$. Therefore, the average time difference is 60 s for each asset. At each time point, the efficient price is observed with independent noise: $e_1(t_{n_1}) \sim NID(0, 0.025)$, $e_2(t_{n_2}) \sim NID(0, 0.05)$.

We performed 500 daily replications and computed the sample MSE:

$$MSE = \frac{1}{500} \sum_{r=1}^{500} \left(estimate^{(r)} - IC^{(r)} \right)^2 .$$

We compared performances of the 10 different estimators shown in Table 2.

The optimal parameters were selected by solving the minimization (4.2) with true values of σ_i^2 and IV_i. As described above, we need piecewise integrated covariance $IC_{n_1 n_2}$ for minimizing the MSE of low-frequency HY estimator. In this simulation, using true IC, we approximated the piecewise covariance as $IC_{n_1 n_2} \approx IC/N_{12}$ where $N_{12} = N_1 + N_2 - \sum I(\{t_{n_1} = t_{n_2}\})$.

The MSEs of estimators with weight functions (5.2) and (5.3) are not so different, especially among EF, Bartlett, Parzen, and Mod. Tukey-Hanning. This implies that the detailed form of function does not have a crucial role; it is more important that we select reasonable bandwidth H or h through optimization.

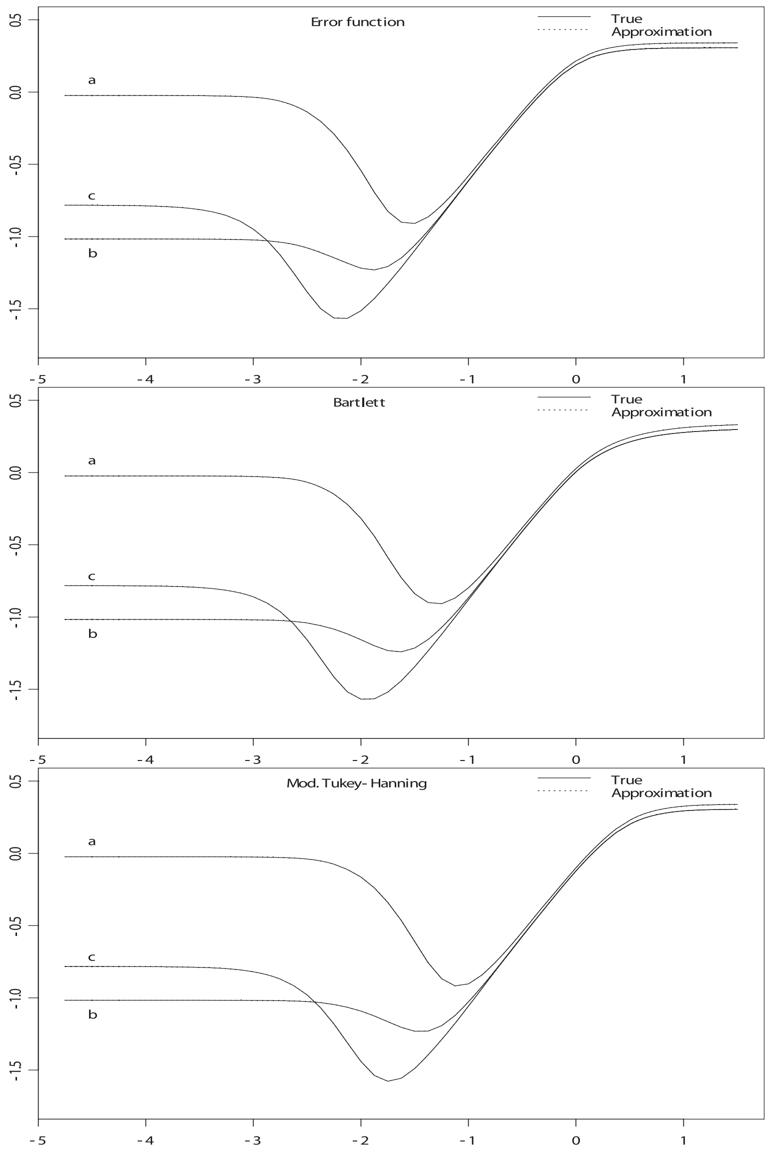
We also performed experiments under different parameter settings. As shown in Table 2, in the case of $\sigma_1^2 = 0.005$, $\sigma_2^2 = 0.01$ and $\lambda_i \Delta = 1/60$, the effect of the noise is sufficiently small that it can be ignored in the estimation by LHY. Consequently, $k = 1$ is selected as the optimal parameter of LHY in every replication. All other kernel methods slightly improve the MSEs compared to the HY estimator. For $\sigma_1^2 = 0.005$, $\sigma_2^2 = 0.01$ and $\lambda_i \Delta = 1/15$, even though the noise is small, the observations are numerous, so that the effect of the noise is accumulated; therefore, it is not negligible. The Bartlett kernel method is the best; however, it is not so different from EF, Parzen, and Mod. Tukey-Hanning.

Figure 1 shows the accuracy of the approximation (4.1). We draw the MSE minus constant of three different estimators for a realization. The true line is drawn by $\sum_{n_1,n_2} w_{n_1 n_2}^2 IV_{n_1} IV_{n_2} + B + C + D$ in (3.2) whereas the approximation by $A' + B' + C' + D$ in (4.2). The volatilities are modeled by mean-reverting diffusion. Therefore, the approximation does not seem to be harmful at all. Figure 2 presents the approximation accuracy in the case in which the assumption is considered to be violated much more. We set

Table 2: Sample MSE and average of optimal parameter

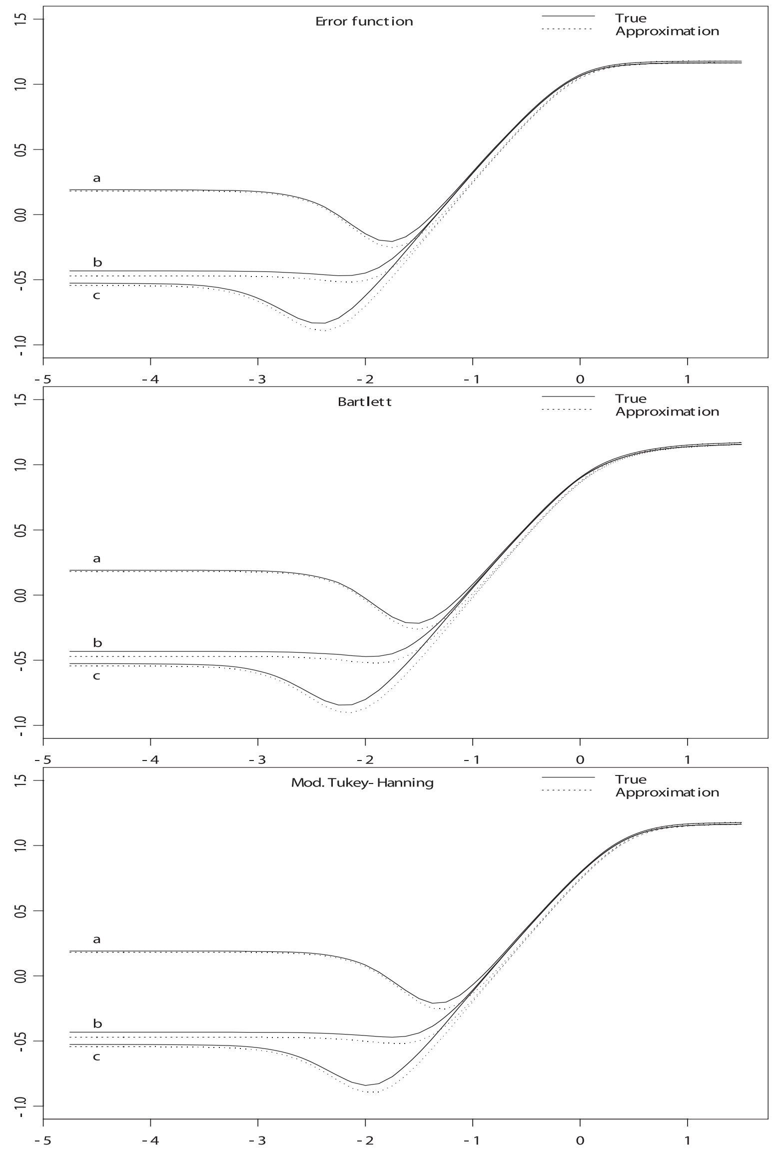
$\lambda_i\Delta$	1/60	1/60	1/15
σ_1^2	0.025	0.005	0.005
σ_2^2	0.05	0.01	0.01
Daily Return	4.55	3.46	2.41
Hayashi-Yoshida	0.845	0.118	0.168
Low Frequency HY	0.469	0.118	0.105
	$(\bar{k}^* = 7.37)$	$(\bar{k}^* = 1)$	$(\bar{k}^* = 6.05)$
Mod. Fourier Estimator	0.242	0.117	0.0485
	$(\bar{n}^* = 12.4)$	$(\bar{n}^* = 25.3)$	$(\bar{n}^* = 49.5)$
Error Function	0.146	0.0911	0.0358
	$(\bar{h}^* = 0.0293)$	$(\bar{h}^* = 0.0130)$	$(\bar{h}^* = 0.00651)$
Bartlett	0.145	0.0907	0.0348
	$(\bar{H}^* = 0.0509)$	$(\bar{H}^* = 0.0226)$	$(\bar{H}^* = 0.0117)$
Epanechnikov	0.185	0.0978	0.0439
	$(\bar{H}^* = 0.0405)$	$(\bar{H}^* = 0.0166)$	$(\bar{H}^* = 0.00918)$
Parzen	0.147	0.0920	0.0368
	$(\bar{H}^* = 0.0673)$	$(\bar{H}^* = 0.0314)$	$(\bar{H}^* = 0.0158)$
Tukey-Hanning	0.153	0.0949	0.0380
	$(\bar{H}^* = 0.0506)$	$(\bar{H}^* = 0.0231)$	$(\bar{H}^* = 0.0117)$
Mod. Tukey-Hanning	0.144	0.0924	0.0361
	$(\bar{H}^* = 0.081)$	$(\bar{H}^* = 0.0377)$	$(\bar{H}^* = 0.0194)$

Figure 1: MSE minus constant ($\kappa = 0.1$,$\gamma = 0.1$)



Note: a: $\lambda_i = 1/60$, $\sigma_1^2 = 0.025$, $\sigma_2^2 = 0.05$; b: $\lambda_i = 1/60$, $\sigma_1^2 = 0.005$, $\sigma_2^2 = 0.01$; c: $\lambda_i = 1/15$, $\sigma_1^2 = 0.005$, $\sigma_2^2 = 0.01$ Both axes are log_{10}-scaled.

Figure 2: MSE minus constant ($\kappa = 0.01$,$\gamma = 1$)



Note: a: $\lambda_i = 1/60$, $\sigma_1^2 = 0.025$, $\sigma_2^2 = 0.05$; b: $\lambda_i = 1/60$, $\sigma_1^2 = 0.005$, $\sigma_2^2 = 0.01$; c: $\lambda_i = 1/15$, $\sigma_1^2 = 0.005$, $\sigma_2^2 = 0.01$ Both axes are log_{10}-scaled.

$\kappa = 0.01$,$\gamma = 1$, that is to say, volatilities are more volatile and persistent. Although the approximation is worse than that in the case of Fig. 1 ($\kappa = 0.1$,$\gamma = 0.1$), the shape of approximated line is similar to that of the true one, therefore the approximation is not so harmful for minimization of the MSE.

7 Concluding remarks

In this paper, we examined the finite sample MSE of weighted realized covariance in the presence of microstructure (independent) noise. Evaluating the MSE of WRC enables us to select not only the optimal parameter, but also the form of the unbiased weighting function. In this paper, as the first-step of application of WRC, we limited our discussion to the weight functions that are unbiased and have only one parameter. Studying more general weight functions is an important remaining task that is now under development.

A MSE of WRC

For A see Kanatani (2004).

Since

$$
\begin{aligned}
&E(r_1(t_{n_1})u_2(t_{n_2})w_{n_1n_2}r_1'Wu_2) = E(r_1(t_{n_1})u_2(t_{n_2})w_{n_1n_2})^2 \\
&+E(r_1(t_{n_1})^2u_2(t_{n_2})u_2(t_{n_2-1})w_{n_1n_2}w_{n_1n_2-1}) + E(r_1(t_{n_1})^2u_2(t_{n_2})u_2(t_{n_2+1})w_{n_1n_2}w_{n_1n_2+1}) \\
&=2IV_{n_1}\sigma_2^2w_{n_1n_2}^2 - IV_{n_1}\sigma_2^2w_{n_1n_2}w_{n_1n_2-1} - IV_{n_1}\sigma_2^2w_{n_1n_2}w_{n_1n_2+1},
\end{aligned}
$$

and

$$
\sum_{n_1,n_2} w_{n_1n_2}w_{n_1n_2-1}IV_{n_1} = \sum_{n_1,n_2} w_{n_1n_2}w_{n_1n_2+1}IV_{n_1},
$$

we obtain B. By symmetry we get C.

Since

$$
\begin{aligned}
&E(u_1(t_{n_1})u_2(t_{n_2})w_{n_1n_2}r_1'Wu_2) = E(u_1(t_{n_1})u_2(t_{n_2})w_{n_1n_2})^2\\
&+E(u_1(t_{n_1})u_1(t_{n_1-1})u_2(t_{n_2})u_2(t_{n_2-1})w_{n_1n_2}w_{n_1-1n_2-1})\\
&+E(u_1(t_{n_1})u_1(t_{n_1-1})u_2(t_{n_2})u_2(t_{n_2+1})w_{n_1n_2}w_{n_1-1n_2+1})\\
&+E(u_1(t_{n_1})u_1(t_{n_1+1})u_2(t_{n_2})u_2(t_{n_2-1})w_{n_1n_2}w_{n_1+1n_2-1})\\
&+E(u_1(t_{n_1})u_1(t_{n_1+1})u_2(t_{n_2})u_2(t_{n_2+1})w_{n_1n_2}w_{n_1+1n_2+1})\\
&+E(u_1(t_{n_1})^2u_2(t_{n_2})u_2(t_{n_2-1})w_{n_1n_2}w_{n_1n_2-1}) + E(u_1(t_{n_1})^2u_2(t_{n_2})u_2(t_{n_2+1})w_{n_1n_2}w_{n_1n_2+1})\\
&+E(u_1(t_{n_1})u_1(t_{n_1-1})u_2(t_{n_2})^2w_{n_1n_2}w_{n_1-1n_2}) + E(u_1(t_{n_1})u_1(t_{n_1+1})u_2(t_{n_2})^2w_{n_1n_2}w_{n_1+1n_2})\\
&=4\sigma_1^2\sigma_2^2w_{n_1n_2}^2\\
&+\sigma_1^2\sigma_2^2w_{n_1n_2}w_{n_1-1n_2-1} + \sigma_1^2\sigma_2^2w_{n_1n_2}w_{n_1-1n_2+1} + \sigma_1^2\sigma_2^2w_{n_1n_2}w_{n_1+1n_2-1} + \sigma_1^2\sigma_2^2w_{n_1n_2}w_{n_1+1n_2+1}\\
&-2\sigma_1^2\sigma_2^2w_{n_1n_2}w_{n_1n_2-1} - 2\sigma_1^2\sigma_2^2w_{n_1n_2}w_{n_1n_2+1} - 2\sigma_1^2\sigma_2^2w_{n_1n_2}w_{n_1-1n_2} - 2\sigma_1^2\sigma_2^2w_{n_1n_2}w_{n_1+1n_2},
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{n_1,n_2} w_{n_1n_2}w_{n_1-1n_2-1} = \sum_{n_1,n_2} w_{n_1n_2}w_{n_1+1n_2+1},\ \sum_{n_1,n_2} w_{n_1n_2}w_{n_11n_2-1} = \sum_{n_1,n_2} w_{n_1n_2}w_{n_1n_2+1},\\
&\sum_{n_1,n_2} w_{n_1n_2}w_{n_1-1n_2+1} = \sum_{n_1,n_2} w_{n_1n_2}w_{n_1+1n_2-1},\ \sum_{n_1,n_2} w_{n_1n_2}w_{n_1-1n_2} = \sum_{n_1,n_2} w_{n_1n_2}w_{n_1+1n_2},
\end{aligned}
$$

then we get D.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Bandi, F. M. and J. R. Russell (2005a). Microstructure noise, realized variance, and optimal sampling. *mimeo*.

Bandi, F. M. and J. R. Russell (2005b). Realized covariation, realized beta and microstructure noise. *mimeo*.

Bandi, F. M. and J. R. Russell (2006). Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. *mimeo*.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2006). Designing realized kernels to measure the ex-post variation of equity price in the presence of noise. *mimeo*.

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

Griffin, J. E. and R. C. A. Oomen (2006). Covariance measurement in the presence of nonsynchronous trading and market microstructure noise. *mimeo*.

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

Kanatani, T. (2004). Optimally weighted realized volatility. *mimeo*.

Kanatani, T. and R. Renò (2007). Unbiased covariance estimation with interpolated data. *mimeo*.

Malliavin, P. and M. E. Mancino (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics 6*, 49–61.

Ubukata, M. and K. Oya (2007). Test of unbiasedness of the integrated covariance estimation in the presence of noise. *mimeo*.

Voev, V. and A. Lunde (2007). Integrated covariance estimation using high-frequency data in the presence of noise. *Journal of Financial Econometrics 5*(1), 68–104.

Zhang, L. (2006). Estimating covariation: Epps effect, microstructure noise. *mimeo*.

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.