18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Bias Corrected Realized Variance with Dependent Microstructure Noise

Oya, K.

Graduate School of Economics, Osaka University, Osaka 560-0043 Japan
Email: oya@econ.osaka-u.ac.jp

Abstract

The estimation of the daily integrated variance of the returns of financial assets is important task for pricing the derivatives of financial asset and risk management. It is well known that a realized variance (RV) is the simplest estimator of the daily integrated variance (IV). It is important that RV is badly biased estimator where the equilibrium price process is contaminated with the market microstructure noise. The microstructure noise is induced by various market frictions such as bid-ask bounces and the discreteness of price changes. There are three approaches to cope with the noise contamination: (i) use of the returns on the proper length of intervals based on optimal sampling frequency proposed by Bandi and Russell(2008a), (ii) subsampling and bias correction proposed by Zhang et al.(2005) and (iii) kernel estimation by Barndorff-Nielsen, et al.(2008). McAleer and Medeiros(2008) extensively review the recent RV literature. The key which ensures unbiasedness and consistency of IV estimator in the presence of market microstructure noise is the time dependence structure of the noise. All three approaches mentioned above eventually require the knowledge of the dependence structure. To identify the dependence structure, Ubukata and Oya(2009) have proposed consistent cross and autocovariance estimators and test statistics for the statistical significance. In this study, we propose the selection procedure of two time scales for TSRV by Aït-Sahalia et al.(2006) under general noise dependence structure applying the statistical inference proposed by Ubukata and Oya(2009). Further an alternative bias corrected IV estimator is also proposed.

We conduct a series of Monte Carlo simulation to compare the bias and root mean squared error of the proposed estimator with TSRV and confirm that the proposed estimator has relatively small MSE and the proposed selection method of two time scales works well.

Denote the extended TSRV and its bias adjusted one with the selected lags $(\hat{J}, \hat{K})$ through the procedure proposed by Ubukata and Oya (2009) as $RV_{\hat{J},\hat{K}}$ and $RV^{(adj)}_{\hat{J},\hat{K}}$, respectively. $RV^{(bc)}_{\hat{K}}$ is the proposed bias corrected estimator with the selected lag $\hat{K}$. In the AR(1) noise dependence case, the empirical distributions of $RV_{\hat{J},\hat{K}}$ and $RV^{(adj)}_{\hat{J},\hat{K}}$ are skew to the right. On the other hand, the skewnesses of the empirical distribution for $i.i.d.$ and MA noise dependence cases are not severe. The empirical distribution of the proposed estimator $RV^{(bc)}_{\hat{K}}$ is closer to symmetric than others when the noise dependence is strong.

The bias and RMSE of estimators are reported in Table 1. The bias of $RV^{(bc)}_{\hat{K}}$ is generally smaller than those of $RV_{\hat{J},\hat{K}}$ and $RV^{(adj)}_{\hat{J},\hat{K}}$. The RMSE is almost same for all cases except the strong noise dependence case. These simulation result suggests that the extended TSRV and its bias adjusted one with selected $(\hat{J},\ \hat{K})$ and the proposed estimator $RV^{(bc)}_{\hat{K}}$ are robust to the dependence of microstructure noise.

Table 1. Relative bias and RMSE of estimators

noise	Bias $RV_{\hat{J},\hat{K}}$	Bias $RV^{(adj)}_{\hat{J},\hat{K}}$	Bias $RV^{(bc)}_{\hat{K}}$	RMSE $RV_{\hat{J},\hat{K}}$	RMSE $RV^{(adj)}_{\hat{J},\hat{K}}$	RMSE $RV^{(bc)}_{\hat{K}}$
AR: $\rho = -0.8$	-0.131	0.035	0.008	0.326	0.366	0.259
AR: $\rho = -0.4$	-0.064	0.009	-0.006	0.208	0.218	0.214
AR: $\rho = 0.4$	0.093	0.176	0.035	0.337	0.373	0.290
AR: $\rho = 0.8$	0.123	0.346	0.078	0.548	0.694	0.390
$i.i.d.$	-0.077	-0.020	-0.009	0.163	0.141	0.212
MA(1)	-0.107	-0.036	-0.008	0.180	0.166	0.200
MA(2)	-0.167	-0.090	-0.011	0.221	0.201	0.204
MA(3)	-0.194	-0.122	-0.033	0.235	0.212	0.217

Keywords: *Realized variance, Dependent market microstructure noise, Two time scales*

1. INTRODUCTION

The estimation of the daily integrated variance of the returns of financial assets is important task for pricing the derivatives of financial asset and risk management. It is well known that a realized variance (RV) is the simplest estimator of the daily integrated variance (IV). The most important fact about the properties of RV is that RV is badly biased estimator where the equilibrium price process is contaminated with the market microstructure noise. The microstructure noise is induced by various market frictions such as bid-ask bounces and the discreteness of price changes. There are three approaches to cope with the noise contamination: (i) use of the returns on the proper length of intervals based on optimal sampling frequency proposed by Bandi and Russell(2008a), (ii) subsampling and bias correction proposed by Zhang et al.(2005) and (iii) kernel estimation Barndorff-Nielsen, et al.(2008). McAleer and Medeiros(2008) extensively review the recent RV literature. The key which ensures unbiasedness and consistency of estimator of IV in the presence of microstructure noise is the time dependent structure of the noise. To identify the dependence structure, Ubukata and Oya(2009) have proposed consistent cross and autocovariance estimators and test statistics for the statistical significance.

In this study, the selection procedure of time scales for Two Scales RV (TSRV) by Aït-Sahalia et al.(2006) is proposed under general noise dependence structure. Further an alternative bias corrected estimator of IV is also proposed. The remainder of this paper is organized as follows. We present the framework of the price process and market microstructure noise in section 2. In section 3, realized variance and related estimator are given. After giving a brief review of the autocovariance estimator of microstructure noise proposed by Ubukata and Oya(2009), we provide the selection two scales for TSRV and an alternative estimator of IV in section 4. In section 5, we examine the finite sample performance of the proposed selection procedure and new estimator under the general noise dependence through Monte Carlo simulation. Some concluding remarks are given in section 6.

2. PRICE PROCESS AND MICROSTRUCTURE NOISE

The most basic assumption for the logarithmic equilibrium price process of an financial asset is that the price process follows a continuous semi-martingale process $\mathrm{d}P^*(t) = \mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t)$ where $P^*(t)$ is the logarithmic equilibrium continuously compounded intra-daily price, $W(t)$ is a standard Brownian motion, $\mu(t)$ and $\sigma(t) > 0$ are bounded measurable functions. The diffusion term $\sigma(t)$ is to be estimated as in the form of integrated variance over a fixed interval $[0, T]$

$$IV = \langle P^*, P^* \rangle_T = \int_0^T \sigma^2(t)\mathrm{d}t \tag{1}$$

using observed logarithmic price of the asset for $t \in [0, T]$. T represents the trading hours per day. During the trading hours on one day, it is usually that the trend of the price process is quite small and the drift term $\mu(t)$ is almost zero. Since our interest is estimating the daily integrated variance (1), we assume that $\mu(t)$ is set to be zero. Suppose that the price is observed at the discrete times $t_0 = 0 < t_1 < t_2 < \cdots < t_n = T$. t_i represents the i-th transaction time. The length of a interval between the $(i-1)$-th and i-th transaction times is defined as $\Delta t_i = t_i - t_{i-1}$. $\Delta t_i = T/n$ when we consider the regular sampling scheme, on the other hand, $\Delta t_i \neq \Delta t_j$, for $i \neq j$ for the non-regular sampling scheme.

To incorporate the effect of market microstructure noise, we assume that the observed logarithmic price process $P(t)$ consists of the equilibrium continuously compounded intra-daily price process $P^*(t)$ which is unobservable and the noise process $\eta(t)$ which is caused by the market microstructure effects as follows

$$P(t) = P^*(t) + \eta(t). \tag{2}$$

The i-th transaction price is $P(t_i)$ and the i-th intraday return is defined as $r_i = P(t_i) - P(t_{i-1})$.

It is natural to consider that the market microstructure noise is serially dependent random variable since the noise is related to bid-ask bounce, the clustering of order flows and other market imperfection. Thus we make the following assumptions for the microstructure noise.

Assumption 1 Market microstructure noise: *Suppose (a) $\{\eta(t)\}$ is a sequence of random variables with zero mean, (b) the noise process is covariance stationary with autocovariance function, which has finite dependence in the sense that:*

$$\gamma_\eta(\ell) = \mathrm{E}[\eta(t)\eta(t-\ell)] = 0, \quad for\ all\ \ |\ \ell\ | > m$$

where m is a finite positive integer, (c) there exists some positive number $\beta > 1$ that satisfies $\mathrm{E}\big|\eta(t)\eta(s)\big|^{4\beta} < \infty$ for all t, s and (d) the noise process is independent of the equilibrium price process.

For (d), even if the noise is correlated with the equilibrium price, the effect of the dependence is dominated by the variation of the noise as the number of high-frequency observations increases. Hansen and Lunde(2006) suggest that the independence assumption (d) does not statistically damage the analysis of asset prices with high trading intensities.

3. REALIZED VARIANCE AND RELATED ESTIMATORS

3.1. Realized Variance

The most popular estimator of (1) is a realized variance which is defined as the sum of squared returns given by

$$RV = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} (P(t_i) - P(t_{i-1}))^2. \tag{3}$$

Given intervals $I^i = (t_{i-1}, t_i]$ for all i, the expectation conditional on the stochastic arrival times is defined as $\mathrm{E}_I[\,\cdot\,]$. The conditional expectation of (3) is written as

$$\mathrm{E}_I[RV] = \sum_{i=1}^{n} \mathrm{E}_I[(r_i^* + \eta(t_i) - \eta(t_{i-1}))^2] = \sum_{i=1}^{n} \mathrm{E}_I[(r_i^*)^2] + 2n\gamma_\eta(0) - 2\sum_{i=1}^{n} \gamma_\eta(\Delta t_i) \tag{4}$$

where $r_i^* = P^*(t_i) - P^*(t_{i-1})$. It is straightforward to show that the variance $\gamma_\eta(0)$ and the sum of autocovariances $\sum_{i=1}^{n} \gamma_\eta(\Delta t_i)$ bring bias to (3). The total bias is $2n(\gamma_\eta(0) - \gamma_\eta(T/n))$ for the regular sampling scheme. In what follows, we assume that the sampling scheme is regular and $\Delta t_i = 1$, $i = 1, \dots, n$ for simplicity.

3.2. Two Scales Realized Variance

Two Scales Realized Variance (TSRV) by Zhang et al.(2005) which is unbiased when the noise is independent. Denote the original grid of observation times as $\mathcal{G} = \{t_0, t_1, \dots, t_n\}$. $\mathcal{G}$ is partitioned into K nonoverlapping subgrids, $\mathcal{G}_K^{(j)}$, $j = 1, \dots, K$, such that $\mathcal{G} = \cup_{j=1}^{K} \mathcal{G}_K^{(j)}$, where $\mathcal{G}_K^{(j)} \cap \mathcal{G}_K^{(\ell)} = \phi$ for $j \neq \ell$. Since we assume the sampling scheme is regular, the j-th nonoverlapping subgrid can be represented as $\mathcal{G}_K^{(j)} = \{t_{j-1}, t_{j-1+K}, t_{j-1+2K}, \dots, t_{j-1+n_jK}\}$ for $j = 1, \dots, K$ where n_j is the integer making t_{j-1+n_jK} the last element in the subgrid $\mathcal{G}_K^{(j)}$. Then the realized variance for the subgrid $\mathcal{G}_K^{(j)}$ is written as

$$RV_K^{(j)} = \sum_{i=1}^{n_j} (P(t_{(j-1)+iK}) - P(t_{(j-1)+(i-1)K}))^2. \tag{5}$$

Let $RV^{(all)}$ be the realized variance for the full grid $\mathcal{G}$. Then TSRV by Zhang et al.(2005) is represented as

$$RV_K = \frac{1}{K} \sum_{j=1}^{K} RV_K^{(j)} - \frac{\bar{n}}{n} RV^{(all)} \tag{6}$$

where $\bar{n} = \sum_{j=1}^{K} n_j/K = (n - K + 1)/K$. The first term of (6) is an average of $RV_K^{(j)}$ for the subgrid $\mathcal{G}_K^{(j)}$, $j = 1, \dots, K$ and is an biased estimator of IV. In the case of the independent noise, the second term of (6) is the bias correction term since the bias in the first term is $2\bar{n}\gamma_\eta(0)$ and $RV^{(all)}/(2n)$ is a consistent estimator of $\gamma_\eta(0)$. The refinement of (6) to remove the finite sample bias is also given in Zhang et al.(2005) as

$$RV_K^{(adj)} = \left(1 - \frac{\bar{n}}{n}\right)^{-1} RV_K. \tag{7}$$

Although these two scales realized variances are unbiased and consistent estimators of IV under the independent noise assumption, the favorable features are lost when the noise is not independent.

3.3. Extended Two Scales Realized Variance

Aït-Sahalia et al.(2006) extend the TSRV to allow for dependent market microstructure noise as follows. The first term of (6) can be rewritten in the form of the average lag K realized variance as defined in Aït-Sahalia et al.(2006) for regular sampling scheme

$$RV_K^{(avg)} = \frac{1}{K} \sum_{j=1}^{K} RV_K^{(j)} = \frac{1}{K} \sum_{i=0}^{n-K} (P(t_{i+K}) - P(t_i))^2. \tag{8}$$

It is possible to use lag J instead of lag K for $1 \le J < K \le n$. Using two different lags J and K, the extended TSRV is defined as

$$RV_{J,K} = RV_K^{(avg)} - \frac{\bar{n}_K}{\bar{n}_J} RV_J^{(avg)} \tag{9}$$

where $\bar{n}_J = (n - J + 1)/J$, $\bar{n}_K = (n - K + 1)/K$, $1 \le J < K \le n$ and $K = o(n)$.

The main difference from TSRV appears in the second term of (9) which is a bias correction term. It is easy to see that the bias of $RV_{J,K}$ comes from the autocovariances $\gamma_\eta(J)$ and $\gamma_\eta(K)$ of the noise as follow

$$2\bar{n}_K(\gamma_\eta(t_{i+J} - t_i) - \gamma_\eta(t_{i+K} - t_i)) = 2\bar{n}_K(\gamma_\eta(J) - \gamma_\eta(K)). \tag{10}$$

These autocovariances $\gamma_\eta(J)$ and $\gamma_\eta(K)$ become negligible if lags J and K are selected large enough. Since we have assumed that the noise process is m-dependent in Assumption 1, we select $J = m + 1$ and $K = O(n^{2/3})$ as in Aït-Sahalia et al.(2006). A finite sample correction of the extended TSRV is given as

$$RV_{J,K}^{(adj)} = \left(1 - \frac{\bar{n}_K}{\bar{n}_J}\right)^{-1} RV_{J,K}. \tag{11}$$

4. SELECTION LAGS (J, K) AND ALTERNATIVE ESTIMATOR

The extended TSRV is suitable estimator of IV when the independent noise assumption is hardly acceptable. As described in previous section, however, we have to choose proper lags J and K. Although Aït-Sahalia et al.(2006) argue that the extended TSRV is robust for the selection of lags (J, K), the estimator is affected by the selection of J as we will see in section 5.2.

In this section, we propose the method which makes it possible to select a proper J using the testing procedure proposed in Ubukata and Oya(2009) and an alternative IV estimator which incorporates a different bias correction method from the extended TSRV.

4.1. Autocovariance Estimator of Microstructure Noise

Ubukata and Oya(2009) propose an unbiased and consistent estimator of autocovariance $\gamma_\eta(\ell)$ of the market microstructure noise and derive the asymptotic distribution. The test statistic of the null hypothesis $\gamma_\eta(\ell) = 0$ is applied to measure the dependence of the market microstructure noise process. Suppose the threshold value of noise dependence as m, that is, $\gamma_\eta(m+1) = 0$ and $\gamma_\eta(m) \neq 0$. The threshold value m can be determined through the test statistic given in Ubukata and Oya(2009, section 3.2).

To obtain unbiased estimator of $\gamma_\eta(\ell)$, we construct $Z_{\ell,ij}^{(\pm)}$ for all i, j such that $\ell = t_{j-1} - t_i$ using the selected threshold value m as follows

$$Z_{\ell,ij}^{(\pm)} = \underline{r}_i^{(-)} \bar{r}_j^{(+)} = (P(t_i) - P(\underline{t}_{i-1}^{(-)}))(P(\bar{t}_j^{(+)}) - P(t_{j-1})) \tag{12}$$

where $\bar{t}_j^{(+)}$ is the first transaction time, which follows t_j subject to $\bar{t}_j^{(+)} - t_i > m$, and $\underline{t}_{i-1}^{(-)}$ is the last transaction time, which is followed by t_{i-1} subject to $t_{j-1} - \underline{t}_{i-1}^{(-)} > m$. Then we have $\mathrm{E}[Z_{\ell,ij}^{(\pm)}] = -\gamma_\eta(\ell)$ for all i, j such that $\ell = t_{j-1} - t_i$ and given ℓ. Let $\{Z_{\ell,k}^{(\pm)}\}_{k=1}^{N_\ell}$ be a sequence that arranges $Z_{\ell,ij}^{(\pm)}$ satisfying $\ell = t_{j-1} - t_i$ in ascending order of index i. N_ℓ is the number of observations in the sequence. The unbiased autocovariance estimator of the microstructure noise is naturally constructed using the sample mean of $\{Z_{\ell,k}^{(\pm)}\}_{k=1}^{N_\ell}$.

Autocovariance Estimator (Ubukata and Oya, 2009): *The autocovariance estimator of the microstructure noise and its asymptotic distribution are given as*

$$\hat{\gamma}_\eta(\ell) = -\frac{1}{N_\ell}\sum_{k=1}^{N_\ell} Z_{\ell,k}^{(\pm)}, \quad N_\ell^{1/2}(\hat{\gamma}_\eta(\ell) - \gamma_\eta(\ell)) \xrightarrow{a} N(0, \omega_\ell^2) \tag{13}$$

where $\omega_\ell^2 = \lim_{N_\ell \to \infty} N_\ell \mathrm{E}[(\hat{\gamma}_\eta(\ell) - \gamma_\eta(\ell))^2]$.

The test statistic to find whether $\gamma_\eta(\ell) = 0$ is given in Ubukata and Oya(2009, corollary 2). We denote the test statistic as $\tau_\eta^*(\ell)$. To save space, we omit its explicit form. See Ubukata and Oya(2009) for details.

4.2. Selection Lags (J, K)

First, we test whether $\gamma_\eta(1) = 0$ through the test statistic $\tau^*_\eta(\ell)$. We conclude that the microstructure noise is uncorrelated when the null is not rejected. If the null is rejected, we test whether $\gamma_\eta(2) = 0$. We continue to test $\gamma_\eta(\ell) = 0$ until the null is not rejected. We denote the distance ℓ where the null hypothesis $\gamma_\eta(\ell) = 0$ is not rejected the first time as $\hat{J}$ for the extended TSRV.

The optimal choice of lag K is given in Zhang et al.(2005) under the $i.i.d.$ noise assumption, however it is still open question under the dependent noise assumption. We adopt a simple approach to settle this question. The original grid of observation times is $\mathcal{G} = \{t_0, t_1, \ldots, t_n\}$ as defined in previous section. Now we define a following new subgrid $\mathcal{G}_{\hat{J}} = \{t_0, t_{\hat{J}}, t_{2\hat{J}}, \ldots, t_{[n/\hat{J}]}\}$ where $\hat{J}$ is the selected lag which is the estimate of the threshold value of the noise dependence through the test statistic $\tau^*_\eta(\ell)$. It is reasonable to suppose that the microstructure noise $\{\ \eta(t_{i\hat{J}})\ \}_{i=0}^{[n/\hat{J}]}$ is an uncorrelated random sequence. By applying the method proposed in Zhang et al.(2005) to the sequence of the observed transaction price $\{\ P(t_{i\hat{J}})\ \}_{i=0}^{[n/\hat{J}]}$, we can find the optimal K. In what follows, the extended TSRV with selected lags ($\hat{J}$, $\hat{K}$) and its bias adjusted version are defined as $RV_{\hat{J},\hat{K}}$ and $RV^{(adj)}_{\hat{J},\hat{K}}$, respectively.

4.3. Alternative Bias Corrected Estimator

In this subsection, we provide an alternative IV estimator using the autocovariance estimator (13) of the market microstructure noise for the bias correction as follow

$$RV_K^{(bc)} = RV_K^{(avg)} - 2\bar{n}_K \hat{\gamma}_\eta(0). \tag{14}$$

The unbiasedness and the consistency of $RV_K^{(bc)}$ are immediately established from the unbiasedness and consistency of the autocovariance estimator (13) and the result given in Zhang et al.(2005).

5. MONTE CARLO SIMULATION

5.1. Simulation Design

We conduct a series of Monte Carlo simulation to see the effect of the selection lags (J, K) for the extended TSRV and to compare the properties of the estimator (14) with the extended TSRV where the microstructure noise is dependent. The data generating process of $P^*(t)$ is the same as in Zhang et al.(2005).

$$\begin{aligned} \mathrm{d}P^*(t) &= (0.05 - \nu(t)/2)\mathrm{d}t \ + \ \sigma(t)\mathrm{d}B(t) \\ \mathrm{d}\nu(t) &= 5(0.04 - \nu(t))\mathrm{d}t \ + \ 0.5\nu(t)^{1/2}\mathrm{d}W(t), \ \ \nu(t) = \sigma^2(t) \end{aligned}$$

where the correlation between the two Brownian motions B and W is set to be -0.5. We generate 10,000 sample paths of the process by Euler scheme at time interval $\Delta t = 1$ second. $T = 1$ day and a day consists of 6.5 hours of open trading = 23400 sec. We observe the price discretely with a market microstructure noise. In the simulation, the time interval of observation is set to be 5 seconds, that is, $t_0 = 0, t_1 = 5, \ldots, t_{4680} = 23400$. Since our interest is in the case of dependent microstructure noise, we consider a variety of different dependent patterns represented by following autoregressive model and moving average model for the noise process

$$\mathrm{AR}(1): \ \ \eta(t_i) = \rho\, \eta(t_{i-1}) \ + \ \varepsilon(t_i), \quad \mathrm{MA}(3): \ \ \eta(t_i) = \varepsilon(t_i) \ + \ \sum_{s=1}^{3} \theta_s \varepsilon(t_{i-s})$$

$\rho =$ -0.8, -0.4, 0.0, 0.4, 0.8 for AR(1), $(\theta_1, \theta_2, \theta_3) =$ (-0.6, 0, 0) for MA(1), (-0.6, -0.3, 0) for MA(2) and (-0.6, -0.3, -0.15) for MA(3). The value of the variance of the market microstructure noise $\mathrm{E}[\eta(t)^2]$ should be selected carefully. The effect of the microstructure noise is negligible when $\mathrm{E}[\eta(t)^2]$ selected too small. Hansen and Lunde(2006) report that the *Noise to Signal Ratio (NSR)* defined as $\mathrm{E}[\eta(t)^2]/IV$ of stocks they examine in NYSE and NASDAQ ranges from 0.0002 to 0.006. We set $\mathrm{E}[\eta(t)^2]$ making the sample average of NSR for simulated path is 0.004. The observed price is given as $P(t_i) = P^*(t_i) + \eta(t_i)$. In the following subsections, we examine the influence of lag J selection on the extended TSRV with $K = 50$, 100, 200 and see the statistical properties of the selected J using the test statistic $\tau^*_\eta(\ell)$. Further, we compare the extended TSRV with selected lags (J, K), its bias-adjusted version and the proposed estimator (14). These estimators are obtained on each simulated sample path.

5.2. Influence of Lag Selection

We estimate the integrated variance using the $RV_{J,K}$ and $RV_{J,K}^{(adj)}$ with $J = 1, \ldots, 40$ and K= 50, 100, 200. The relative bias of estimators which is the sample means of $(estimate - IV)/IV$ is given in the first and third columns in Figure 1. The sample root mean squared error (RMSE) of $estimate/IV$ is given in the second and fourth columns. The first and third rows of Figure 1 show the bias and RMSE for $RV_{J,K}$. The second and fourth rows show those for $RV_{J,K}^{(adj)}$. The horizontal axes are J=1, . . . , 40. The models used for the noise process are AR(1) with $\rho = -0.8$ and 0.8, $i.i.d$ and MA(1) with θ=-0.8. It is clear that the bias of $RV_{J,K}^{(adj)}$ becomes negligible after J exceeds the threshold value of noise dependence while the bias of $RV_{J,K}$ grows as J increases. Although the bias adjusted estimator $RV_{J,K}^{(adj)}$ works well, the variance is affected by the adjustment. This effect is captured by RMSE.

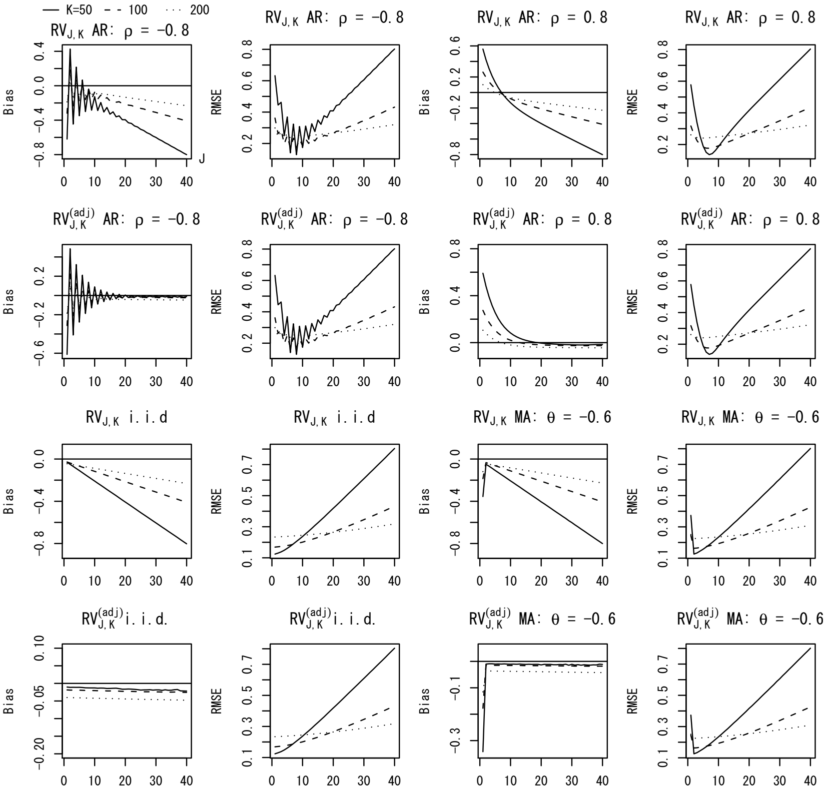

Figure 1. Effect of selection J on Extended TSRV with different K

It is important to select the lag J for both $RV_{J,K}$ and $RV_{J,K}^{(adj)}$ appropriately because their RMSEs are strongly depend on the lag J even in the independent noise case. From the second and fourth columns of Figure 1, we find that the influence of the selection K is not severe on RMSE of the estimator once the lag J is selected properly.

5.3. Performance of Lag Selection Procedure

The selection of the lag J is conducted by testing whether the null hypothesis $\gamma_\eta(\ell) = 0$ for $\ell > 0$ as described in section 4.2. We examine how the test statistic $\tau_\eta^*(\ell)$ works for selection of the lag J. Figure 2 shows the empirical distribution of selected J denoted as $\hat{J}$ for the representative cases. The mode of $\hat{J}$ for AR(1) with ρ = -0.8 and 0.8 are 8 and 6 when the serial dependence of the noise is strong. On the other hand, those for $i.i.d.$, MA(1) and MA(2) cases are 1, 2 and 3 respectively. However, the mode of $\hat{J}$ for MA(3) case is not 4. The selected number for $\hat{J}$=4 is the second largest. One possible explanation for this is that $\gamma_\eta(2)$ and $\gamma_\eta(3)$ are small in MA(3) noise process used in this experiment.

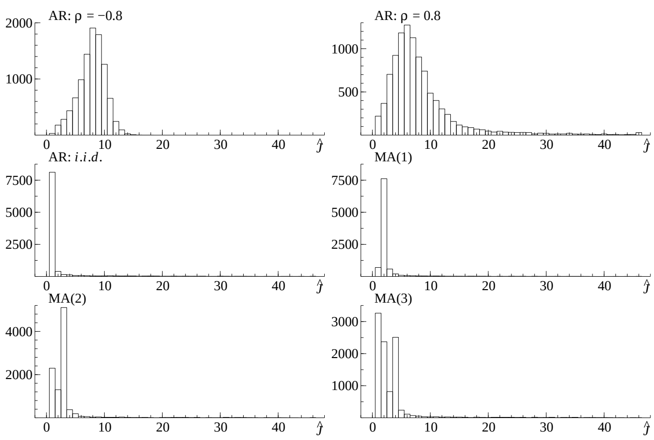

Figure 2. Empirical distribution of selected J using $\tau_\eta^*(\ell)$

It seems reasonable to conclude that the test statistic $\tau_\eta^*(\ell)$ works well as the criterion for the selection lag J except the small autocovariance case. Though we omit the result for the selection of K by the procedure described in the previous section, we find that the variance of the empirical distribution of $\hat{K}$ becomes larger as the noise dependence increases.

5.4. Comparison of Estimators

We have discussed three estimators of IV under the dependent microstructure noise assumption. In this subsection, we examine the properties of these estimators. $RV_{\hat{J},\hat{K}}$ and $RV_{\hat{J},\hat{K}}^{(adj)}$ are the extended TSRV and its bias adjusted one with the selected lags $(\hat{J}, \hat{K})$. $RV_{\hat{K}}^{(bc)}$ is the estimator proposed in the previous section with the selected lag $\hat{K}$. In the AR(1) noise dependence case, the empirical distributions of $RV_{\hat{J},\hat{K}}$ and $RV_{\hat{J},\hat{K}}^{(adj)}$ with $(\hat{J}, \hat{K})$ are skew to the right. On the other hand, the skewness of the empirical distribution for $i.i.d.$ and MA noise dependence cases is not severe. The empirical distribution of the proposed estimator $RV_{\hat{K}}^{(bc)}$ is closer to symmetric than others when the noise dependence is strong. The bias and RMSE of estimators are reported in Table 1. The bias of $RV_{\hat{K}}^{(bc)}$ is generally smaller than those of $RV_{\hat{J},\hat{K}}$ and $RV_{\hat{J},\hat{K}}^{(adj)}$. The RMSE is almost same for all cases except the strong noise dependence case. These simulation result suggests that the extended TSRV and its bias adjusted one with selected $(\hat{J},\ \hat{K})$ and the proposed estimator $RV_{\hat{K}}^{(bc)}$ are robust to the dependence of microstructure noise.

Table 1. Relative bias and RMSE of estimators

	Bias			RMSE		
noise	$RV_{\hat{J},\hat{K}}$	$RV_{\hat{J},\hat{K}}^{(adj)}$	$RV_{\hat{K}}^{(bc)}$	$RV_{\hat{J},\hat{K}}$	$RV_{\hat{J},\hat{K}}^{(adj)}$	$RV_{\hat{K}}^{(bc)}$
AR: $\rho = -0.8$	-0.131	0.035	0.008	0.326	0.366	0.259
AR: $\rho = -0.4$	-0.064	0.009	-0.006	0.208	0.218	0.214
AR: $\rho = 0.4$	0.093	0.176	0.035	0.337	0.373	0.290
AR: $\rho = 0.8$	0.123	0.346	0.078	0.548	0.694	0.390
$i.i.d.$	-0.077	-0.020	-0.009	0.163	0.141	0.212
MA(1)	-0.107	-0.036	-0.008	0.180	0.166	0.200
MA(2)	-0.167	-0.090	-0.011	0.221	0.201	0.204
MA(3)	-0.194	-0.122	-0.033	0.235	0.212	0.217

6. CONCLUDING REMARKS

This study provides a selection procedure of the lags $(J,\ K)$ for the two scales realized variance with dependent microstructure noise and an alternative bias corrected estimator using the variance estimator of microstructure noise proposed in Ubukata and Oya(2009). From Monte Carlo simulation result, we find that there is evidence that the proposed lag selection procedure works well and the proposed estimator is associated with relatively smaller bias and RMSE. From these viewpoint, one may say that the proposed lag selection procedure and the proposed estimator are useful for the analysis of actual transaction price data. In this study, we do not compare the kernel type estimator. The kernel type estimator is more efficient than the estimators we examine in this study. However, Bandi and Russell(2008b) show that this asymptotic property does not always provide satisfactory result for realistic sample size for empirical analysis. Thus, it is interesting to compare the estimators examined in this study with the kernel type estimator. We will take this matter in future study.

7. REFERENCES

Aït-Sahalia, Y., P. Mykland and L. Zhang (2006), "Ultra high frequency volatility estimation with dependent microstructure noise," NBER Working Paper No. 11380.

Bandi, F. M. and J. R. Russell (2008a), "Microstructure noise, realized variance, and optimal sampling," *Review of Economic Studies*, 75 (2), 339-369.

Bandi, F. M. and J. R. Russell (2008b), "Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations," Unpublished paper: Graduate School of Business, University of Chicago.

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

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

McAleer, M. and M. C. Medeiros (2008), "Realized Volatility: A Review," *Econometric Reviews*, 27, 10-45.

Ubukata, M. and K. Oya (2009), "Estimation and Testing for Dependence in Market Microstructure Noise," *Journal of Financial Econometrics*, doi:10.1093/jjfinec/nbn021.

Zhang, L., P. 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.