Center for the Study of Finance and Insurance
Osaka University

Discussion Paper Series 2012-02

Volatility Forecast Comparison with Biased Proxy

Shuichi Nagata and Kosuke Oya

Center for the Study of Finance and Insurance
Osaka University

Volatility Forecast Comparison with Biased Proxy

Shuichi Nagata
Department of Mathematical Sciences,
Kwansei Gakuin University

Kosuke Oya
CSFI & Graduate School of Economics,
Osaka University

October, 2012

Abstract

Various loss functions are employed in literature to evaluate competing volatility forecasting models. The loss function for an evaluation requires true volatility, which is unobservable. Patton (2011) provides a new class of loss functions that guarantees the consistency of the ranking (asymptotically) when the unbiased volatility proxy is used instead of true volatility. However, while realized variance (RV) is commonly used as the proxy in practice, it is natural to consider that RV does not satisfy the unbiasedness condition due to market microstructure noise.

In this paper, we show that such bias in the volatility proxy can cause misspecified rankings of competing models. We also introduce a new notion for the robustness of loss functions to incorporate the effect of the biased volatility proxy and propose a proper method to evaluate the competing forecasting models if the volatility proxy is biased. We conduct a series of Monte Carlo simulation to access the performance of our method and confirm that the proposed method behaves well.

keywords: forecasting evaluation, loss function, realized variance, volatility.

1 Introduction

Since volatility (conditional variance of return distribution) plays a crucial role in financial research and practice, a number of studies have been conducted that aim to forecast future volatility accurately. However, the evaluation of the accuracy of these comparative forecasting models is not easy, because true volatility is unobserved and even ex-post. That is, it is common practice to evaluate the performance of forecasting using a loss function and substituting the volatility proxy for it. Traditionally, squared return is employed as a proxy for true volatility.

Various loss functions are proposed in previous studies to evaluate the performances of forecasting. The problem is that these loss functions sometimes provide different results for ranking volatility forecasts in empirical research. Therefore, one would want to select the better loss function from these proposed loss functions in some criteria.

To provide a solution to this question, Hansen and Lunde (2006a) who explicitly consider the error of proxy show the sufficient conditions for loss functions that give a consistent ranking of volatility forecasting even proxy is imperfect (has error). They also show mean squared error (MSE) is robust in that sense. Patton (2011) shows the class of robust loss function and it includes major loss functions, MSE and quasi likelifood (QLIKE). Therefore, one can employ MSE or QLIKE to compare the accuracy of the forecasting candidates, and the consistency of ranking them is approved when all their assumptions are satisfied. Furthermore, considering the simulation results of Patton and Sheppard (2009) that compare the power of the test proposed by Diebold and Mariano (1995) and West (1996) (DMW test), QLIKE is the best loss function for volatility forecasting comparison.

However, we should carefully review and verify the relation of their assumptions to a real situation. Recently, realized variance (RV) is employed as a volatility proxy due to it greater efficiency than squared daily returns. While RV is commonly used as the proxy in practice, it is natural to consider that RV does not satisfy the unbiasedness condition due to market microstructure noise. The treatment of market microstructure noise is discussed in econometric literature, such as Bandi and Rusell (2008) and Barndorff-Nielsen et al. (2008). We give a brief explanation for market microstructure noise and its effect on RV estimation in the latter section.

We should note that the results of Hansen and Lunde (2006a) and Patton (2011) are derived with the assumption that volatility proxy is unbiased. Therefore, it is important to know whether robustness of the loss functions holds when biased volatility proxy is employed.

Here, we consider consistency for ranking volatility forecasting when the volatility proxy is biased. First, we reconsider the result in Hansen and Lunde (2006a) by relaxing their condition, which is the unbiasedness of the volatility proxy. As a result of this reconsideration, we introduce a new notion of robustness to take on the effect of the bias of volatility proxy on the robustness of the loss function. We show that using our method, one can evaluate the robustness of loss function when volatility proxy is biased. The feature of our method is that it is free from estimation of the bias of proxy, and this property is useful in practice.

We conduct a simulation to evaluate the effect of biased volatility proxy on the robustness of loss functions. First, we assume the bias direction is upward, and it is a natural and acceptable assumption with respect to the empirical results in the literature. we also consider the situation that the sign of the bias is unknown. We employ major robust loss functions, and compute the accuracy of their ranking when the volatility proxy is biased. We also report the performance using our method of choosing loss function with consideration for robustness of bias.

The remainder of this paper is organized as follows. In Section 2, we introduce our theoretical framework and show the main results. We analyze a simulation to assess our theoretical results in Section 3. Finally, we conclude the paper in section 4.

2 Robust loss function and realized variance

First, we establish notation. Let r_t be a financial asset return at time t, usually daily or monthly. $\mathcal{F}_{t-1}$ is the information set at time $t-1$, which includes a higher frequency than r_t. We assume $E[r_t \mid \mathcal{F}_{t-1}] = 0$, thus the variable of interest $\sigma_t^2 = V[r_t \mid \mathcal{F}_{t-1}] = E[r_t^2 \mid \mathcal{F}_{t-1}]$. The loss function for forecasting evaluation is $L : \mathbb{R}_+ \times \mathcal{H} \rightarrow \mathbb{R}_+$ where $\mathbb{R}_+$ denotes a nonnegative part of the real line. The first argument of L is σ_t^2 or its proxy $\hat{\sigma}_t^2$ and the second is h_t which is a forecast of σ_t^2. $\mathcal{H}$ is compact subset of $\mathbb{R}_{++}$, which denotes a positive part of the real line.

We consider the situation where one wants to rank comparative forecasting from models A and B in terms of the accuracy of forecasting. The key difference between volatility forecasting and other forecasts is that true volatility is not observable even ex-post. Therefore, one must use volatility proxy $\hat{\sigma}_t^2$ instead of true volatility for the forecast evaluation.

Andersen and Bollerslev (1998), Hansen and Lunde (2006a), and Patton (2011) focus on volatility proxy as beeing imperfect and discuss the desirable properties of loss function in such a situation. The following notion, "robustness of loss function" is a key idea in literature.

Recalling Patton (2011), we define the loss function $L(\sigma_t^2, h_t)$ is robust, if the ranking of the two volatility forecasts h_A and h_B using expected loss of $L(\sigma_t^2, h_t)$ with true volatility and with the proxy are the same. More precisely, loss function $L(\sigma_t^2, h_t)$ is robust if $L(\sigma_t^2, h_t)$ is satisfied using the following relational explanation:

$$E[\ L(\sigma_t^2, h_{t,A})\ ] \gtrless E[\ L(\sigma_t^2, h_{t,B})\ ] \ \Leftrightarrow\ E[\ L(\hat{\sigma}_t^2, h_{t,A})\ ] \gtrless E[\ L(\hat{\sigma}_t^2, h_{t,B})\ ], \tag{1}$$

for any s.t. $E[\hat{\sigma}_t^2 | \mathcal{F}_{t-1}] = \sigma_t^2$.

Hansen and Lunde (2006a) show that the sufficient condition for a loss function to be robust is that $\frac{\partial^2 L(\sigma^2, h)}{\partial(\sigma^2)^2}$ does not depend on h. Following Hansen and Lunde (2006a), Patton (2011) derives necessary sufficient condition for robust loss function and provides a new class of loss functions that guarantee the consistency of the ranking (asymptotically) when the unbiased volatility proxy is used instead of true volatility.

$$R(\hat{\sigma}, h; b) = \begin{cases} \frac{1}{(b+1)(b+2)}(\hat{\sigma}^{2b+4} - h^{b+2}) - \frac{1}{b+1}h^{b+1}(\hat{\sigma}^2 - h) & (b \notin \{-1, -2\}) \\ \hat{\sigma}^2 \log \frac{\hat{\sigma}^2}{h} - (\hat{\sigma}^2 - h) & (b = -1) \\ \frac{\hat{\sigma}^2}{h} - \log \frac{\hat{\sigma}^2}{h} - 1 & (b = -2) \end{cases} \tag{2}$$

where b is scalar parameter. $R(\hat{\sigma}, h; b)$ with $b = -2$ and $b = 0$ corresponding to QLIKE and MSE, respectively.

It should be noted that the results of Patton (2011) and Hansen and Lunde (2006a) were derived under the assumption that volatility proxy is unbiased. As explained in a later section, it is possible for volatility proxy to be biased in practice. Therefore, it is important to know the robustness of the loss function when biased volatility proxy is employed.

In the next subsection, while introducing the key idea of Bandi and Rusell (2008), we discuss the possible existence of the RV bias the common volatility proxy in practice. We also review the discussion of robustness of loss function with biased proxy assumption in 2.2.

2.1 Realized variance as a volatility proxy

Assume that the logarithmic price process $p(t)$ is determined by

$$dp(t) = \sigma(t)dW(t), \tag{3}$$

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 major volatility models.

Then σ_t^2 is defined as follows:

$$\sigma_t^2 = \int_{t-1}^{t} \sigma^2(s)ds.$$

We now consider the estimation of σ_t^2 from a discrete price set by (2), $p_{t,1}, p_{t,2}, ..., p_{t,M}$ in $[t-1, t]$. However, we can only obtain the data set $\tilde{p}_{t,i}$ that contaminated by market microstructure noise.

$$\tilde{p}_{t,i} = p_{t,i} + u_{t,i}, \quad u_{t,i} \sim i.i.d.(0, \omega^2) \qquad \text{for } i = 1, ..., M, \tag{4}$$

where $u_{t,i}$ is market microstructure noise and its mean and variance are 0 and ω^2, respectively.

Using observable price set $\tilde{p}_1, ..., \tilde{p}_{t,M}$, RV is defined as

$$RV_{t,M} = \sum_{i=1}^{M} \tilde{r}_{t,i}^2, \tag{5}$$

where $\tilde{r}_{t,i} = \tilde{p}_{i,t} - \tilde{p}_{i,t-1}$.

As mentioned, $RV_{t,M}$ is an inconsistent estimator of σ_t^2 in the presence of market microstructure noise. More precisely, if $M \to \infty$ $RV_{t,M}$ diverges to infinity. The conditional expectation and variance with assuming the independence noise of price,

$$E[RV_{t,M}|\sigma_t] = M\omega^2 + o(M),$$
$$Var[RV_{t,M}|\sigma_t] = 2IQ_t/M + o(1/M),$$

where $IQ_t = \int_{t-1}^{t} \sigma^4(s)ds$. Using above result, Bandi and Rusell (2008) propose the optimal sampling frequency of RV in terms of MSE. Optimal number of observations M^* is determined as

$$M^* \simeq (IQ_t/\omega^4)^{1/3}.$$

We can estimate RV effectively in terms of MSE, using above rule. However, it should be noted that if we use it to determine the number of observations, RV estimates include not only sampling errors but also upward bias.

After the study conducted by Bandi and Rusell (2008), more sophisticated estimators were proposed for integrated volatility estimation such as those of Zhang et. al. (2005) and Barndorff-Nielsen et al. (2008). However, optimal sampling RV (or sparse sampling RV) is still common in practice, because of its simplicity.

2.2 The effect of biased volatility proxy

In this subsection, we reconsider the result of Hansen and Lunde (2006a). To be specific, we relax the condition of unbiasedness of volatility proxy $\hat{\sigma}_t^2$. Here, we introduce the constant bias q as $E[\hat{\sigma}_t^2] = \sigma_t^2 + q$, where $q \in \mathbb{R}$.

First, we recall Assumption 2 in Hansen and Lunde (2006a), making a small modification.

Assumption 1. An extension of Assumption 2 in Hansen and Lunde (2006a) σ_t^2 and $\hat{\sigma}_t^2$ denote two (possibly random) variables.

(i) Define $\eta_t = \hat{\sigma}_t^2 - \sigma_t^2$ and $\{\mathcal{F}_t\}$ a filtration it holds that h_t and θ_t are $\{\mathcal{F}_{t-1}\}$ measurable.

(ii) Either, (a) $L'(\sigma^2, h) = \partial L(\sigma^2, h)/\partial\sigma^2$ exists and does not depend on h; or (b) $L''(\sigma^2, h) = \partial^2 L(\sigma^2, h)/\partial\sigma^2\partial\sigma^{2\prime}$ exists, and does not depend on h, and $\{\eta_t - q, \mathcal{F}_t\}$ is a martingale difference sequence.

We only relax the condition of unbiasedness of volatility proxy $\hat{\sigma}_t^2$, then we can obtain following proposition.

Proposition 1. Assumption 1 is satisfied. Then, the robustness indicator of Loss function

$L(\theta, X)$ with two forecasts $H_t = h_{t,A}, h_{t,B}$ is determined by $Z_L(H_t)$, where

$$Z_L(H_t) = \frac{E[\ L(\sigma_t^2, h_{t,A})\ ] - E[\ L(\sigma_t^2, h_{t,B})\ ]}{E[\ L(\hat{\sigma}_t^2, h_{t,A})\ ] - E[\ L(\hat{\sigma}_t^2, h_{t,B})\ ]} = 1 \ - \ q\, w_L(H_t),$$
$$w_L(H_t) = \frac{E[L'(\sigma_t^2, h_{t,A}) - L'(\sigma_t^2, h_{t,B})]}{E[L(\hat{\sigma}_t^2, h_{t,A})] - E[L(\hat{\sigma}_t^2, h_{t,B})]}.$$

If $Z_L(H_t) > 0$, then $L(\hat{\sigma}^2, h)$ is robust.

Proof of Proposition 1. The Proof is given in Appendix.

If $Z_L(H_t) = 1$, that is, either $q = 0$ or $q \neq 0$ and $w_L(H_t) = 0$, then obviously $L(\hat{\sigma}^2, h)$ is robust. Therefore, this proposition includes Theorem 2 in Hansen and Lunde (2006a) as the $q = 0$ case.

We now focus on the $q \neq 0$ case. Then, $L(\hat{\sigma}^2, h)$ presents an inverse ranking when $q w_L(H_t) > 1$. Thus, robustness of Loss function L does not hold when volatility proxy is biased. As shown in Proposition 1, the sign of $Z_L(H_t)$ shows the robustness of L. we want to know about $Z_L(H_t)$ using data H_t and $\hat{\sigma}^2$ for $t = 1, ..., n$. However, it is infeasible to evaluate $Z_L(H_t)$ in a practical case, because of the numerator of $Z_L(H_t)$ includes σ_t^2.

On the other hand, if the numerator of $w_L(H_t)$ does not include σ_t^2, then, it is possible to evaluate robustness of L via $w_L(H_t)$. Therefore, by calculating $w_L(H_t)$ for all candidate loss functions, we can select a better loss function to the obtain correct ranking of models. In next section, we demonstrate how to use our proposed indicator in practice.

Our indicator can be defined if Loss function satisfies assumption 1. That is, we can define our indicator for robust loss functions proposed by Patton (2011).

Collorary 1. Empirical robustness of the robust loss function $R(\hat{\sigma}^2, h; b)$ by comparing two forecasts $H_t = h_{t,A}, h_{t,B}$, is determined by $w_{R(b)}(H_t)$ and is as follows;

$$w_{R(b)}(H_t) = \frac{s(H_t; b)}{s(H_t; b)\hat{\sigma}_t^2 + v(H_t; b)},$$

where

$$s(H_t; b) = \begin{cases} h_{t,A}^{b+1} - h_{t,B}^{b+1} \\ \log(h_{t,A}/h_{t,B}) \\ 1/h_{t,A} - 1/h_{t,B} \end{cases} \quad \text{and} \quad v(H_t; b) = \begin{cases} -\frac{b+1}{b+2}(h_{t,A}^{b+2} - h_{t,B}^{b+2}) & (b \notin \{-1, -2\}) \\ (h_{t,A} - h_{t,B}) & (b = -1) \\ -(\log(h_{t,A}) - \log(h_{t,B})), & (b = -2). \end{cases}$$

3 Monte Carlo simulation

in this section, we conduct a Monte Carlo simulation. The purposes of the simulation are; (1) to evaluate the effect of biased proxy on empirical ranking of volatility models; (2) to evaluate the method for loss function selection using the proposed indicator. We generate artificial data by following a standard GARCH(1,1) model, which is employed in Patton and Sheppard (2009),

$$r_t = \sigma_t \varepsilon_t, \quad t = 1, 2...T. \tag{6}$$

$$\sigma_t^2 = \mu + \alpha \sigma_{t-1}^2 + \beta r_{t-1}^2. \tag{7}$$

Following Patton and Sheppard (2009), we set the parameter values of (7) and (8) as $\mu = 0.05$ $\alpha = 0.85$, $\beta = 0.10$ $T = 1060$.

We consider the situation of employing the RV as a volatility proxy. Following Patton and Sheppard (2009) we allow the error for RV, and additionally allow the bias for RV. First, we consider upward bias for RV with reference to Bandi and Rusell (2008). We also consider the sign of the bias is unknown case in later of this section. Here, RV with M samples are computed as follows.

$$RV_t(M, q) = \sigma_t^2 \sum_{i=1}^{m} \left(\sum_{j=\gamma(i-1)+1}^{\gamma i} \xi_{t-i} \right)^2 + q \tag{8}$$

where $\xi_i \sim i.i.d.N(0, 1/390)$ and $\gamma = M/m$. We calculate RV with $M = 26, 78$ and 390, that corresponds to using fifteen minute, five minute, and one minute returns, respectively. q is a bias of volatility proxy. We set q as 0. 0.1, 0.5 and 1.

In this simulation, the signal to noise ratio $\xi^2 := w^2/IV$ is $\xi^2 = w^2$ because $E[\sigma_t^2] = 1$. The relation of the parameters are as $\xi^2 = q/M$. We employ $RV_t(M, q)$ as a proxy with $0.1 \leq q \leq 1$ and $26 \leq M \leq 390$. That is, it is equal to set $0.0002 \leq \xi^2 \leq 0.038$.

However, the range of ξ^2 obtained by empirical analysis of U. S. stocks in Hansen and Lunde (2006b) is $0.0002 \leq \hat{\xi}^2 \leq 0.006$. Hence, we only report the simulation results with the settings of ξ^2 (determined by M and q) that fall above range.

Following Patton (2011), we employ (1) Rolling window and (2) Risk Metrics for volatility forecasting.

$$h_{t,RW} = \frac{1}{60} \sum_{i=1}^{60} r_{t-i}^2 \tag{9}$$

$$h_{t,RM} = \lambda h_{t-1,RM} + (1 - \lambda) r_{t-1}^2 \tag{10}$$

where $\lambda = 0.75$.

Then, we compare the above two candidates using three loss functions. We calculate the average loss by loss functions for three different choices of parameters $b = -2, -1, 0$.* Patton's Robust loss function with $b = -2$ and $b = 0$ corresponds to QLIKE and MSE, respectively. We set $n = 100, 800$ to calculate the average of loss function values.

Here, we propose a new method using the results from Section 2. Our propose methods are as follows:

*First we also calculate loss for $b = 1, 2$ cases. However, the performances of these settings are not good, therefore, we skip its discussion.

Case 1: With upper bias assumption

According to the discussion of Bandi and Rusell (2008), it is natural to assume that $E[\hat{\sigma}_t^2] = \sigma_t^2 + q$ where $q > 0$. In this case, if $w_L(H_t) < 0$, $L(\hat{\sigma}^2, h)$ is robust. Now, we define $\hat{w}_{R(i)}(H) = \frac{1}{n}\sum_{t=1}^{n} w_{R(i)}(H_t)$. Then we choose loss function as follows:

Proposed method 1: (assuming upper bias)

1. Calculate $\hat{w}_{R(i)}(H)$ for each sample.

2. If $\max(\hat{w}_{R(i)}(H);\ \text{for } i = -2, -1, 0) \geq 0$, we choose the loss function $R(i)$ to rank forecasting models, whose $\hat{w}_{R(i)}(H)$ takes minimum value.

3. If $\max(\hat{w}_{R(i)}(H);\ \text{for } i = -2, -1, 0) < 0$, we choose the loss function $R(i)$ to rank forecasting models, whose $\hat{w}_{R(i)}(H)$ takes takes maximum value.

Case 2: Bias direction is unknown

Hansen and Lunde (2006b) show that RV estimates do not always diverge to the upper side that in practice. As shown in Section 2, the explanation of Bandi and Rusell (2008) assume that noise is i.i.d. Hansen and Lunde (2006b) show that the bias of RV has the possibility to be negative if the noise is dependent.

If we consider the sign of q as unknown, it is better to choose the loss function whose $|\hat{w}|$ is the least. Then we choose loss function as follows:

Proposed method 2: the sign of bias is unknown

1. Calculate $\hat{w}_{R(i)}(H)$ for $i = -2, -1$, and 0.

2. We choose the loss function $R(i)$ to rank forecasting models, whose $|\hat{w}_{R(i)}(H)|$ takes minimum value.

Finally, we calculate the failure rate α for the above two loss functions and our proposed method to evaluate the performance.

$$\alpha := \frac{1}{m}\sum_{j=1}^{m} 1_{(d^{(j)} - \hat{d}^{(j)} < 0)} \tag{11}$$

where $d^{(j)} = \frac{1}{n}\sum_{t=1}^{n} L(\sigma_t^{2(j)}, h_{t,A}^{(j)}) - L(\sigma_t^{2(j)}, h_{t,B}^{(j)})$, $\hat{d}^{(j)} = \frac{1}{n}\sum_{t=1}^{n} L(\hat{\sigma}_t^{2(j)}, h_{t,A}^{(j)}) - L(\hat{\sigma}_t^{2(j)}, h_{t,B}^{(j)})$, and m denote and the number of Monte Carlo replications. The sign of these quantities show which forecast is better. Clearly, it takes $0 \leq \alpha \leq 1$. If loss function is robust, then $\alpha = 0$ because $E[d^{(i)} - \hat{d}^{(i)}] = 0$. The results are summarized in Table 1.

Table 1: Monte Carlo comparison of α for several loss functions for the GARCH(1,1) model with upper biased proxy. The number of replications is 2,000.

$RV(q,m)$		$n=100$					$n=800$				
q	M	QL	$R(-1)$	MSE	Prop1	Prop2	QL	$R(-1)$	MSE	Prop1	Prop2
$q=0$	390	0.4	0.3	0.1	0.1	**0.0**	**0.0**	0.1	**0.0**	**0.0**	**0.0**
	72	0.6	0.8	0.4	0.2	**0.0**	**0.0**	0.3	**0.0**	**0.0**	**0.0**
	39	1.1	1.2	0.5	0.2	**0.0**	**0.0**	0.6	**0.0**	**0.0**	**0.0**
$q=0.1$	390	6.3	0.9	2.4	**0.0**	**0.0**	**0.0**	1.3	**0.0**	**0.0**	**0.0**
	72	6.0	1.1	2.5	0.1	**0.0**	**0.0**	1.3	**0.0**	**0.0**	**0.0**
	39	6.5	1.2	2.9	0.1	**0.0**	**0.0**	1.3	**0.0**	**0.0**	**0.0**
$q=0.5$	390	36.2	4.1	20.6	**0.0**	**0.0**	10.6	7.4	**0.0**	**0.0**	**0.0**
	72	36.4	4.4	20.2	**0.0**	**0.0**	10.8	7.2	**0.0**	**0.0**	**0.0**
$q=1.0$	390	51.0	7.2	31.7	**0.0**	0.4	71.8	19.6	0.2	**0.0**	0.1

Bold entries indicate that the best result of five candidates. The rows corresponding to $q=0$ show the result without bias. The α's are 0 in all loss functions (without $R(-1)$) when the sample size is large ($n=800$), thus it roughly meets the results in Patton (2011). The rows with $q \neq 0$ report the results with biased proxy. Comparing to the unbiased case, the values of α are higher, the bias of proxy affects the robustness of loss functions. Importantly, we find that the α of QLIKE and $R(-1)$ are relatively larger than MSE. We can observe that our proposed methods 1 and 2 provide better or equivalent to other loss functions, in terms of failure rate. That is, we can conclude that we can improve the performance of ranking of competing models by choosing the better loss function using our proposed method.

Table 2: Monte Carlo comparison of α of several loss functions for the GARCH(1,1) model with biased proxy. The number of replications is 2,000.

$RV(q,m)$		$n=100$					$n=800$				
$\lvert q\rvert$	M	QL	$R(-1)$	MSE	Prop1	Prop2	QL	$R(-1)$	MSE	Prop1	Prop2
$\lvert q\rvert=0$	390	0.4	0.3	0.1	0.1	**0.0**	**0.0**	0.1	**0.0**	**0.0**	**0.0**
	72	0.6	0.8	0.4	0.2	**0.0**	**0.0**	0.3	**0.0**	**0.0**	**0.0**
	39	1.1	1.2	0.5	0.2	**0.0**	**0.0**	0.6	**0.0**	**0.0**	**0.0**
$\lvert q\rvert=0.1$	390	3.9	0.8	1.5	0.5	**0.0**	**0.0**	1.4	**0.0**	**0.0**	**0.0**
	72	3.7	1.1	1.5	0.5	**0.0**	**0.0**	1.3	**0.0**	**0.0**	**0.0**
	39	3.7	1.2	1.9	0.7	**0.0**	**0.0**	1.2	**0.0**	**0.0**	**0.0**
$\lvert q\rvert=0.5$	390	24.5	4.8	18.8	9.2	**0.1**	4.9	5.5	**0.0**	**0.0**	**0.0**
	72	24.7	4.8	18.9	9.5	**0.0**	4.9	5.4	**0.0**	**0.0**	**0.0**
$\lvert q\rvert=1.0$	390	36.2	11.3	31.7	16.1	**1.6**	34.9	11.6	0.3	0.1	**0.0**

Table 2 show the results with the bias direction is unknown. From table 2 we find almost the same conclusion from table 1. If the proxy is unbiased ($q = 0$), the performances are equivalent in all loss functions when the sample size $n = 800$ except $R(-1)$. Interestingly, regardless of simulation settings, proposed method 2 uniformly shows the best performance in Table 2.

4 Conclusion

In this paper, we considered the situation where the volatility proxy is biased, and thus analyzed its effect on the robustness of loss functions. The intuition of our results is that bias in volatility proxy can causes misspecified rankings of competing models. Specifically, our simulation results indicate that the bias effect on the robustness of QLIKE is higher than those of MSE.

We introduced the robustness indicator and proposed the methods to choose a better loss function using the indicator. Our simulation results indicate that our proposed method is useful to avoid obtaining misspecified rankings with biased volatility proxy.

References

Andersen, T.G. and T. Bollerslev, (1998), "Answering the skeptics: yes, standard volatility models do provide accurate forecasts," *International Economic Review* vol. 39, 885-–905.

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

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., and N. Shephard (2008), "Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise," *Econometrica* vol. 76-6, 1481–1536.

Diebold, F. X. and R. S. Mariano (1995), "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, vol. 13, 3, 253–263.

Hansen, P. R. and A. Lunde (2006a), "Consistent ranking of volatility models," *Journal of Econometrics*, vol. 131, issue 1-2, 97–121.

Hansen, P. R. and A. Lunde (2006b), "Realized variance and market microstructure noise," *Journal of Econometrics*, vol. 24, issue 2, 127– -161.

Patton, A. J. (2011), "Volatility forecast comparison using imperfect volatility proxies," *Journal of Econometrics*, vol. 160, issue 1, 246–256.

Patton, A.J. and K. Sheppard (2009), "Evaluating volatility and correlation forecasts," *The Handbook of Financial Time Series*, Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, T. (eds.), Springer Verlag.

West, K.D. (1996), "Asymptotic inference about predictive ability," *Econometrica*, vol. 64, 1067–1084.

Zhang, L., Mykland, P.A., and Y. At-Sahalia (2005), "A tale of two time scales: determining integrated volatility with noisy high-frequency data," *Journal of the American Statistical Association* vol. 100, 1394–1411.

Appendix

Proof of Proposition 1. When the Assumption (ii.a) is satisfied, the numerator of the second term of right hand side equals 0, it is clear.

With Assumption (ii.b), we consider the second order Taylor expansion of loss function L. That is,

$$L(\hat{\sigma}_t^2, h_t) \;=\; L(\sigma_t^2, h_t) \;+\; L'(\sigma_t^2, h_t)\,\eta_t \;+\; \frac{1}{2}\eta_t'\, L''(\sigma_t^{2\,**}, h_t)\,\eta_t,$$

where, $\sigma_t^{2\,**}$ is present between $\hat{\sigma}_t^2$ and σ_t^2. Taking Expected Value,

$$E[\,L(\hat{\sigma}_t^2, h_t)\,] = E[\,L(\sigma_t^2, h_t)\,] \;+\; E[\,L'(\sigma_t^2, h_t)\,\eta_t\,] \;+\; \frac{1}{2}E[\,\eta_t'\, L''(\sigma_t^{2\,**}, h_t)\,\eta_t\,], \tag{12}$$

where the last term does not depend on h_t because of Assumption (ii.b), and the second term is

$$\begin{aligned} E[\,L'(\sigma_t^2, h_t)\,\eta_t \mid \mathcal{F}_{t-1}\,] &= E[\,L'(\sigma_t^2, h_t)\,(\eta_t - q) \mid \mathcal{F}_{t-1}] + q\,E[\,L'(\sigma_t^2, h_t) \mid \mathcal{F}_{t-1}] \\ &= q\,E[\,L'(\sigma_t^2, h_t) \mid \mathcal{F}_{t-1}]. \end{aligned}$$

Therefore, for all t,

$$\begin{aligned} E[\,L(\hat{\sigma}_t^2, h_{t,A})\,] - E[\,L(\hat{\sigma}_t^2, h_{t,B})\,] &= E[\,L(\sigma_t^2, h_{t,A})\,] - E[\,L(\sigma_t^2, h_{t,B})\,] \\ &\quad + q \times \left\{E[\,L'(\sigma_t^2, h_{t,A})\,] - E[\,L'(\sigma_t^2, h_{t,B})\,]\right\}. \end{aligned}$$

Hence, we have

$$\frac{E[\,L(\sigma_t^2, h_{t,A})\,] - E[\,L(\sigma_t^2, h_{t,B})\,]}{E[\,L(\hat{\sigma}_t^2, h_{t,A})\,] - E[\,L(\hat{\sigma}_t^2, h_{t,B})\,]} \;=\; 1 \;-\; q\,\frac{E[\,L'(\sigma_t^2, h_{t,A})\,] - E[\,L'(\sigma_t^2, h_{t,B})\,]}{E[\,L(\hat{\sigma}_t^2, h_{t,A})\,] - E[\,L(\hat{\sigma}_t^2, h_{t,B})\,]}.$$

□