Constructing a Realized DCC Model with Measurement Errors

Hideto Shigemoto ,Takayuki Morimoto [1]

Abstract

We propose novel conditional autoregressive Wishart models for high-dimensional realized covariance matrices of asset returns. We incorporate measurement errors into realized variance dynamics of Realized Dynamic Conditional Correlation model. It is well known that the measurement errors make the realized volatility less persistent than the latent volatility process. Therefore, by introducing measurement errors into realized volatility, the persistence of realized volatility based on the magnitude of the corresponding measurement errors can be incorporated into the multivariate model. Our empirical analysis performs out-of-sample evaluations for 100 stocks on the Tokyo Stock Exchange from January 1, 2014, through December 31, 2020. Our model based on logarithmic realized volatility shows the best forecast performance across test periods and loss functions.

Keywords:
Realized covariance; Measurement error; Dynamic conditional correlation model; Covariance forecasting.

1 Introduction

Forecasting and modeling covariance matrices of asset returns are fundamental elements in financial practices. The multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) models are used to estimate conditional covariance matrices from asset returns (Bauwens et al., 2006). MGARCH models treat covariance matrices as latent, but a recent alternative uses high-frequency data to estimate a realized covariance (RC) measures for low-frequency returns precisely (Andersen et al., 2003; Barndorff-Nielsen and Shephard, 2004). These estimators which are estimated by high-frequency data can be modeled directly. Forecasting models using RC measures can obtain more accurate predictions than the MGARCH models (Golosnoy et al., 2012).

The BEKK-type conditional autoregressive Wishart (CAW) model proposed by Golosnoy et al. (2012) guarantees the positive definiteness of forecast covariances without constraining parameters by assuming a conditional Wishart distribution across an entire RC matrix. BEKK-type models are difficult to apply to high-dimensional covariance matrices because of their greater number of parameters. Bauwens et al. (2012) adopted the assumptions of the CAW model and proposed realized dynamic conditional correlation model (Re-DCC). Then Bauwens et al. (2012) provided an empirical analysis for 50 assets.

A realized volatility (RV) measure estimated from high-frequency data is a consistent estimator of a latent volatility under a high-frequency setting, but it is subject to measurement error for any finite sample. Therefore, RV equals the sum of two elements: the latent volatility and a measurement error. Attenuation bias arises because the directly observable realized measure process has less persistence than has the latent volatility (Bollerslev et al., 2016). To address that problem, Bollerslev et al. (2016) proposed HARQ model and Hizmeri et al. (2020) proposed a dilution bias correction (DBC) HAR model. Both introduce time-varying autoregressive parameters that are adjusted, so parameters are large (small) when the variance of the realized measure error is small (large).

In this paper, we propose the novel CAW models that incorporates measurement errors. We extend the the Re-DCC model to simplify the introduction of measurement errors and make predictions even in high-dimensional matrices. The Re-DCC model decomposes the scale matrix of the Wishart distribution into variance and correlation. We consider measurement errors within variance dynamics. By incorporating measurement errors into the estimation of variance dynamics of the Re-DCC model, the persistence of RV improvement and fast mean reversion considered in the HARQ and DBC HAR models can be incorporated into the multivariate model. We use 100 individual stocks in the Nikkei 225. The out-of-sample evaluation also showed improved predictive accuracy in many situations. Our proposed model based on log-realized volatility performs best.

2 Realized measures

We consider the n-dimensional log-price process:

$$X(t) = \int_0^t \mu(u)du + \int_0^t \sigma(u)dW(u), \quad (1)$$

[1] Email addresses: h-shigemoto@kwansei.ac.jp (H, Shigemoto), morimot@kwansei.ac.jp (T, Morimoto)

where, $\mu(u)$ and $\sigma(u)$ are instantaneous drift and volatility processes. $W(u)$ denotes the n-dimensional vector of independent Brownian motion. $\Sigma(u) = \sigma(u)\sigma(u)'$ is a spot covariance matrix. We exclude jumps for simplicity.

The integrated covariance for the log-price process in Eq. (1) on day t is

$$\Sigma_t = \int_{t-1}^{t} \Sigma(u)du. \tag{2}$$

The integrated covariance cannot be observed directly, but we can estimate it consistently based on intraday high-frequency data. In this paper, we use the multivariate realized kernel estimator proposed by Barndorff-Nielsen et al. (2011) as a consistent estimator of the integrated covariance. We estimate the variance of the multivariate realized kernel (Π_t) using Theorem 2 in Barndorff-Nielsen et al. (2011).

3 Model

Let C_t be a consistent estimator for Σ_t in Eq. (2). We model the n-dimensional stochastic positive-definite realized covariance measures $\{C_t\}_{t=1}^{T}$. Given the filtration $F_{t-1} = \{C_{t-1}, C_{t-2}, \dots\}$, C_t is assumed to follow the central Wishart distribution

$$C_t|F_{t-1} \sim W_n(\nu, S_t/\nu), \tag{3}$$

where $\nu > n-1$ is the degree of freedom and S_t/ν is the symmetric, positive definite $n \times n$ scale matrix. In addition, $E(C_t|F_{t-1}) = S_t$ is obtained from characteristics of the Wishart distribution. The scale matrix is decomposed into RV and realized correlation as $S_t = D_t R_t D_t$, where R_t is a conditional correlation matrix and $D_t = \{\text{diag}(S_t)\}^{1/2}$ denotes the diagonal matrix in which the conditional standard deviation $\sqrt{S_{jj,t}}$ of asset j aligns diagonally.

Bauwens et al. (2012) estimate RV dynamics using a GARCH type model and a HAR model. We use the specification of realized correlation dynamics in Bauwens et al. (2012):

$$R_t = (1-a-b)\overline{R} + \sum_{i=1}^{p} a_i P_{t-i} + \sum_{j=1}^{q} b_j R_{t-j},$$
$$P_t = \{\text{diag}(C_t)\}^{-1/2} C_t \{\text{diag}(C_t)\}^{-1/2},$$

where $0 \leq a, 0 \leq b, 0 \leq a+b \leq 1$. $\overline{R}$ is estimated by the sample mean of the realized correlation matrix (correlation targeting).

3.1 HARQ Re-DCC model

As a proxy for latent volatility, RV is subject to measurement error for a finite sample. Bollerslev et al. (2016) show that variance of the measurement error of RV is a function of IQ. They proposed the HARQ model, given that the larger the variance of the measurement error, the smaller the persistence of the observed process.

We propose a HARQ Re-DCC model which apply a HARQ model to conditional variance.

$$RV_t = \beta_c + \underbrace{(\beta_d + \alpha \Pi_{t-1}^{1/2})}_{\beta_{d,t}} RV_{t-1} + \beta_w RV_{t-5|t-1} + \beta_m RV_{t-22|t-1},$$

where $RV_{t-h|t-1} = \frac{1}{h}\sum_{i=1}^{h} RV_{t-i}$. Also $\beta_d, \beta_w, \beta_m$ are positive. We expect α will be negative. The coefficient of the daily parameter diminishes when $\alpha < 0$ and the value of $\Pi^{1/2}$ is large. We can construct the HAR model, which has a time-varying daily parameter, because values of the $\beta_{d,t}$ parameter vary with the estimated measurement error variance.

3.2 DBC Re-DCC model

Hizmeri et al. (2020) show that daily realized volatility always has large measurement errors. However, monthly RV is less affected by the measurement errors than daily RV, and the distribution is centered to zero. That is, RV generally overestimates integrated volatility in the presence of measurement errors. Thus, they propose DBC HAR model which takes the absolute value of the difference between the daily RV and the monthly RV as the relative magnitude of the measurement errors of the daily RV.

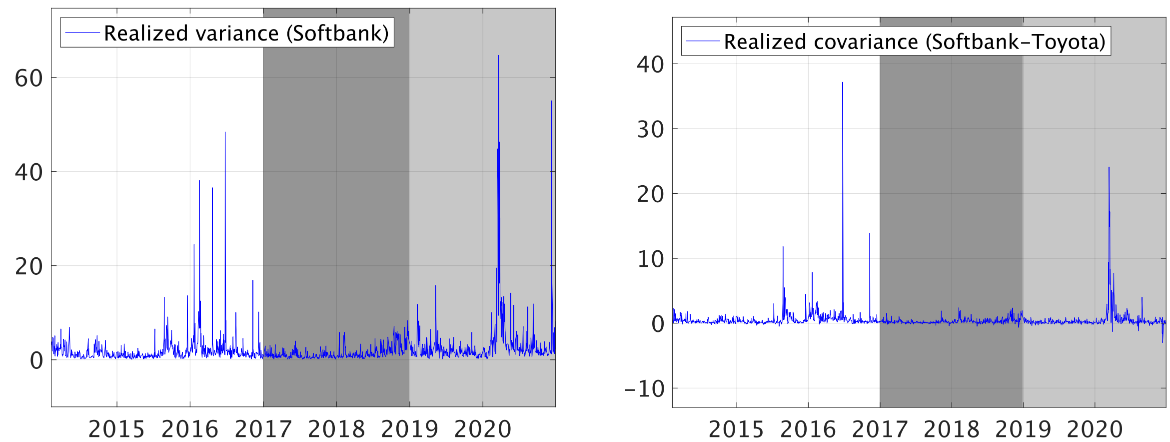

Figure 1: RV and RC between Softbank and Toyota. Gray shaded areas are the out-of-sample period.

We propose a DBC Re-DCC model which apply it to a DBC HAR model for conditional variance:

$$RV_t = \beta_c + \underbrace{\{\beta_d + \alpha|RV_{t-1} - RV_{t-22|t-1}|\}}_{\theta_{d,t}} RV_{t-1} + \beta_w RV_{t-5|t-1} + \beta_m RV_{t-22|t-1},$$

where like in the HARQ Re-DCC model, $\beta_d, \beta_w, \beta_m$ are positive and α again is expected to be negative.

3.2.1 Log type model

To guarantee RV is non-negative, we also propose DBC Re-DCC (log), a logarithmic model for RV.

$$\log(RV_t) = \beta_c + \underbrace{\{\beta_d + \alpha|\log(RV_{t-1}) - \log(RV_{t-22|t-1})|\}}_{\theta_{d,t}} \log(RV_{t-1}) + \beta_w \log(RV_{t-5|t-1}) + \beta_m \log(RV_{t-22|t-1}).$$

4 Empirical analysis

Our empirical analysis is based on daily realized covariance matrices of 100 individual stocks traded on the Nikkei 225. According to Barndorff-Nielsen et al. (2011), we estimate the multivariate realized kernel.

The sample period starts at January 1, 2014, and ends on December 31, 2020, covering 1624 trading days. Figure 1 shows realized volatility, realized covariance and realized correlation between Softbank Group and Toyota Motor. The out-of-sample data is from January 1, 2017, and ends on December 31, 2020.

4.1 Out-of-sample evaluation

We evaluate forecasting performance of the Re-DCC and HAR Re-DCC, and our proposed HARQ Re-DCC and DBC Re-DCC models. We apply an "insanity filter" to ensure that all forecast values are not less (larger) than the minimum (maximum) of historical observations (Bollerslev et al., 2018).

To select loss functions, we use the noise robust functions suggested by Patton (2011). We also use the global minimum-variance portfolio, an economically important indicator. We use the average of loss functions and the model confidence set (MCS) proposed by Hansen et al. (2011). For the MCS process, we set the significance level to 20% and use block-bootstrapping with length 25.

Table 1 shows the result of one-ahead forecasting. For QLIKE and GMVP, the DBC Re-DCC (log) model is selected via MCS process. Also, the DBC Re-DCC model is selected for MSE (Var) and MSE (Corr). However, for MSE (Cov), MSE (Var) and Frobenius loss, the existing model which are HAR Re-DCC and Re-DCC models are selected. These model has more accurate in volatile period like light gray area in Figure1. When volatility is high (inaccurate volatility), the sustainability of daily RV declines in the proposed models after considering measurement errors. Therefore, when the volatility is continuously high, forecasting accuracy fades, and differences from the RV increase.

Table 1: Average daily one-period-ahead forecasting losses : January 1, 2017 - December 31, 2020.

All period : 1/1/2017 - 31/12/2020							
Model	Order	QLIKE	GMVP	MSE(Cov)	MSE(Var)	MSE(Corr)	Frobenius
HARQ Re-DCC	(1,1)	170.2	0.3637	5901	423.6	271.0	11379
	(2,2)	163.8	0.3546	6152		259.5	11880
	(3,3)	162.6	0.3568	6445		270.4	12467
DBC Re-DCC	(1,1)	147.0	0.3282	5985	416.1	254.2	11554
	(2,2)	157.7	0.3362	6369		262.0	12322
	(3,3)	161.3	0.3376	6073		255.8	11730
DBC Re-DCC (log)	(1,1)	103.6	0.2866	6294	416.3	261.1	12171
	(2,2)	112.6	0.2964	6603		269.1	12790
	(3,3)	110.4	0.2938	6557		266.9	12698
HAR Re-DCC	(1,1)	110.7	0.295	6085	404.6	268.0	11766
	(2,2)	112.4	0.2956	6077		266.9	11749
	(3,3)	109.5	0.2941	6065		267.7	11725
Re-DCC	(1,1)	136.2	0.3214	5741	463.1	268.2	11019
	(2,2)	122.0	0.3031	4995	417.8	256.2	9571
	(3,3)	123.9	0.3057	5375	406.0	266.3	10344

Notes : All values for MSE_{var} are the all the same in the HAR-type model because the dynamics of realized variance are regardless of order. Red indicates minimum and the gray shaded values indicate that the 80% MCS process includes the respective model.

5 Conclusion

In this paper, we propose a novel Re-DCC model that incorporates measurement errors for a time series of high-dimensional realized covariance matrices. We estimated the Re-DCC model by decomposing the scale matrix of a Wishart distribution into variances and correlation. Our proposed model used HARQ and DBC HAR models to estimate variance in the Re-DCC model. Doing so made it possible to introduce a time-varying autoregressive parameter that is adjusted so that the parameter is large when variance of the realized volatility error is small and large when the variance is small. Incorporating measurement errors into the Re-DCC model makes it multivariate model with improved persistence and fast mean reversion corresponding to the magnitude of realized volatility error, as considered in the HARQ and DBC HAR models.

We provided an empirical application to realized covariance measures for 100 stocks in the Nikkei 225. Our proposed model, particularly the DBC Re-DCC (log) model, provides forecasts superior to benchmarked models.

References

[1] Andersen, T. G., Bollerslev, T., Diebold, F.X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, **71(2)**, 579-625.

[2] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate realised kernels : Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, **162**, 149-169.

[3] Barndorff-Nielsen, O.E., and Shephard, N. (2004). Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. *Econometrica*, **72(3)**, 885-925.

[4] Bauwens, L., Laurent, S., and Rombouts, J. V. K., (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, **21**, 79-109.

[5] Bauwens, L., Storti, G., and Violante, F. (2012). Dynamic conditional correlation models for realized covariance matrices. CORE Working paper.

[6] Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2016). Exploiting the errors: a simple approach for improved volatility forecasting. *Journal of Econometrics*, **192(1)**, 1-18.

[7] Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2018). Modeling and forecasting (un)reliable realized covariances for more reliable financial decisions. *Journal of Econometrics*, **207**, 71-91.

[8] Golosnoy, V., Gribisch, B., and Liesenfeld, R. (2012). The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics*, **167**, 211-223.

[9] Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, **79**, 453-497.

[10] Hizmeri, R., Izzeldin, M., and Tsionas, M. (2020). A simple model correction for modeling and forecasting (un)reliable realized volatility. SSRN Working Paper.

[11] Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, **160(1)**, 246-256.