Linear time-varying regression with Copula–DCC–GARCH models for volatility

Jong-Min Kim [a], Hojin Jung [b,*]

[a] Statistics Discipline, Division of Science and Mathematics, University of Minnesota at Morris, Morris, MN 56267, United States
[b] School of Economics, Henan University, Ming Lun Street, Kaifeng, Henan, 475001, China

HIGHLIGHTS

- Providing a new linear time-varying regression model with dynamic conditional correlation.
- Comparing our proposed models with the model currently used in literature in the ex-post volatility forecast evaluations.
- Finding evidence that our proposed model provides superior forecasts for volatility over the model currently used.

ARTICLE INFO

ABSTRACT

This paper provides a new linear time-varying regression with dynamic conditional correlation (DCC) estimated by Gaussian and Student-t copulas for forecasting financial volatility. Time-varying parameters will be estimated for nonparametric dependence by using copula functions with United States stock market data. We compare our model with Kim et al.'s (2016) linear time-varying regression (LTVR) with DCC–GARCH in the ex-post volatility forecast evaluations. Empirical study shows that our proposed volatility models are more efficient than the LTVR model. We also use the superior predictive ability and the reality check for data snooping. Evidence can be found supporting that our proposed model with copula functions provides superior forecasts for volatility over the LTVR model.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Understanding the volatility of financial asset returns is crucial for hedging, risk management, and portfolio optimization. The family of GARCH has been widely used to model and forecast volatility (see Bollerslev et al., 1992; Engle, 2004, for a comprehensive review). In particular, multivariate GARCH models are popular for taking into account financial volatilities comovements by estimating a conditional correlation matrix. One method to estimate the conditional correlation is the constant conditional correlation (CCC) model introduced by Bollerslev (1990). But, this model assumes that a constant correlation is constant over time, which is not realistic in most of cases. One popular approach for considering this problem in the literature is the DCC–GARCH model introduced by Engle (2002).

The aim of this article is to introduce an alternative tool for forecasting stock return volatility. In particular, we develop the LTVR with DCC–GARCH (Kim et al., 2016), by estimating the DCC with copulas. Since Patton (2006) introduced copulas with time-varying parameters to model exchange rate dependence, a number of studies have employed the time-varying copulas. For example, Jondeau and Rockinger (2006) employ the skewed Student-t copula in the GARCH framework for the daily return markets. Bartram et al. (2007) use Patton's (2006) time-varying copula model in order to investigate the dependence between seventeen European stock markets. See Manner and Reznikova (2012) for an overview and comparison of time-varying copula models.

We compare the performance of volatility forecasting of our proposed model to Kim et al.'s (2016) LTVR model. A number of

* Corresponding author.
  E-mail addresses: jongmink@morris.umn.edu (J.-M. Kim),
2hojin.jung@gmail.com (H. Jung).

the pairwise comparisons, such as the mean square error (MSE) and mean absolute error (MAE) are employed to evaluate the forecast accuracy of our model. This study finds empirical evidence that the linear time-varying regression model with the Copula–DCC–GARCH statistically outperforms the linear time-varying regression model with the DCC–GARCH. The remainder of this paper is organized as follows. In the next section, we discuss the linear state space regression model with time-varying parameters proposed by Kim et al. (2016). Various copula functions and their theoretical backgrounds are discussed in Section 3. Section 4 discusses empirical analysis. Concluding remarks are presented in Section 5.

2. Volatility by using a linear state space regression model with time-varying parameters

Kim et al. (2016) consider a linear state space regression model combined with a DCC–GARCH model in order to predict volatility. In a linear regression of Y given X, the estimate of the regression slope coefficient can be expressed as:

$$\hat{\beta}_1 = \hat{\rho} \times \frac{\text{sample standard deviation of } Y}{\text{sample standard deviation of } X} \tag{1}$$

where $\hat{\rho}$ is the sample correlation between X and Y. Thus, Eq. (1) can be easily rewritten as follows:

$$\hat{\sigma}_Y = \hat{\beta}_1 \times \frac{\hat{\sigma}_X}{\hat{\rho}}.$$

Given the simple expression, Kim et al. (2016) propose the simple but powerful method for estimating financial asset return volatility:

$$\hat{\sigma}_{Y_t} = \hat{\beta}_t \times \frac{\hat{\sigma}_{X_t}}{\hat{\rho}_t}, \tag{2}$$

where $\hat{\sigma}_{X_t}$ is the $\sqrt{r_t^2}$ of an individual firm stock index and $\hat{\beta}_t$ is the parameter estimates from a linear state space regression model. $\hat{\rho}_t$ is dynamic conditional correlation estimates between the S&P 500 index and the individual firm stock index. To obtain the estimates of the time-varying regression slopes, we follow Kim et al.'s (2016) LTVR model, which is defined by

$$\begin{aligned} \text{Measurement equation}: \ & y_t = \alpha_t + \beta_t x_t + v_t, \\ & v_t \sim N(0, \sigma_v^2), \\ \text{Transition equation}: \ & \alpha_t = \alpha_{t-1} + \omega_{\alpha,t}, \quad \omega_{\alpha,t} \sim N(0, \sigma_\alpha^2), \\ \text{Transition equation}: \ & \beta_t = \beta_{t-1} + \omega_{\beta,t}, \quad \omega_{\beta,t} \sim N(0, \sigma_\beta^2), \end{aligned} \tag{3}$$

where σ_v^2 indicates the variance of the observation noise. The first diagonal element of the variance matrix of the system noise is σ_α^2, and σ_β^2 is placed in the second diagonal element.[1] The measurement equation has the time-varying coefficients (β_t). The variance of disturbance σ_v^2 is also assumed to be time-variant. In this study, the parameters β_t follow the first-order random walk process. In our study, the parameter estimation is accomplished via Kalman filtering functions, **dlm.MLE**, in the R package *dlm*.

We make an extension of their LTVR model with a regular DCC–GARCH model by using a Copula–DCC–GARCH model. We improve Kim et al. (2016) by considering time-varying nonparametric dependence by copula functions. After the parameters estimation, we are interested in the conditional correlation estimates from the time-varying normal copulas.

---

[1] Nakajima (2011) employs the time-varying parameter vector autoregressive (TVP-VAR) model with stochastic volatility for the Japanese macroeconomic variables.

3. The GARCH–copula model

Sklar (1959) introduced copula functions as a general tool to construct joint multivariate distributions. They have been introduced recently in the finance literature by Frey et al. (2001) and Li (1999). The copula models we consider here are elliptical copulas. In particular, Gaussian and Student copulas are used to estimate the time-varying correlation matrix using the DCC model.

Copula functions allow us to obtain the univariate marginal distribution function from the dependence structure of a set of random variables. A copula is an efficient way to characterize and model correlated multivariate random variables. Let X_i be a random variable with a marginal distribution function F_i for $i = 1, 2, \ldots, n$. Sklar (1973) shows that each multivariate distribution function $F(x_1, \ldots, x_n)$ can be represented as its marginal distribution function by using a copula such as

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)).$$

An n-dimensional copula C determined in $[0, 1]^n$ for distributions F can be defined by

$$C(u_1, \ldots, u_n) = F\left(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)\right)$$
$$\text{for } \forall u_i \in [0, 1], \ i = 1, \ldots, n.$$

Then, the density functions of F and C are given by

$$f(x_1, \ldots, x_n) = c(F_1(x_1), \ldots, F_n(x_n)) \prod_{i=1}^n f_i(x_i)$$

$$c(u_1, \ldots, u_n) = \frac{f(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))}{\prod_{i=1}^n f_i\left(F_i^{-1}(u_i)\right)}$$

where f_i are the marginal densities and F_i^{-1} are the quantile function of the margins.

Finally, the density of the Gaussian copula is defined by

$$c(u; \mathbf{R}) = \frac{1}{|\mathbf{R}|^{1/2}} e^{-\frac{1}{2} u'(\mathbf{R}'-\mathbf{I})u} \tag{4}$$

where $\mathbf{R}$ is the correlation matrix implied by the covariance matrix, $u_i = F(x_i)$, and $\mathbf{I}$ is the identity matrix. The density of the Student-t copula with shape parameter τ is defined by

$$c(\mathbf{u}; \mathbf{R}, \tau) = \frac{\Gamma\left(\frac{\tau+n}{2}\right) \left(\Gamma\left(\frac{\tau}{2}\right)\right)^n \left(1 + \tau^{-1} \mathbf{u}' \mathbf{R}^{-1} \mathbf{u}\right)^{-(\tau+n)/2}}{|\mathbf{R}|^{1/2} \left(\Gamma\left(\frac{\tau+n}{2}\right)\right)^n \Gamma\left(\frac{\tau}{2}\right) \prod_{i=1}^n \left(1 + \frac{\mathbf{u}_i^2}{\tau}\right)^{-(\tau+1)/2}} \tag{5}$$

where $\mathbf{u}_i = t_\tau^{-1}(F(x_i; \tau))$. Note that t_τ^{-1} in the density functions is the quantile function. See Ghalanos (2015) for additional details. In order to investigate the time-varying correlations of stock returns, we adopt the DCC model. In the DCC model of Engle (2002) and Tse and Tsui (2002), the correlation matrix is time varying, and the covariance matrix can be decomposed into:

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R} \mathbf{D}_t = \rho_{ij} \sqrt{h_{iit} h_{jjt}}, \tag{6}$$

where $\mathbf{D}_t = diag(\sqrt{h_{11,t}}, \ldots, \sqrt{h_{nn,t}})$ containing the time-varying standard deviations is obtained from GARCH models. The DCC model in Engle (2002) has the following structure:

$$\mathbf{R}_t = diag(\mathbf{Q}_t)^{-1/2} \mathbf{Q}_t \, diag(\mathbf{Q}_t)^{-1/2} \tag{7}$$

where

$$\begin{aligned} \mathbf{Q}_t &= \bar{\mathbf{Q}} + a\left(z_{t-1} z'_{t-1} - \bar{\mathbf{Q}}\right) + b\left(\mathbf{Q}_{t-1} - \bar{\mathbf{Q}}\right) \\ &= (1 - a - b)\bar{\mathbf{Q}} + a z_{t-1} z'_{t-1} + b \mathbf{Q}_{t-1} \end{aligned} \tag{8}$$

where $a, b \geq 0$ such that $a + b < 1$ to ensure stationarity and positive definiteness of $\mathbf{Q}_t$. $\bar{\mathbf{Q}}$ is the unconditional variance–covariance matrix of the standardized errors z_t.

The time-varying conditional correlation in the copula framework with the elliptical copulas is an extension of the DCC model. Let $\mathbf{r}_t = (r_{it}, \ldots, r_{nt})$ be a $n \times 1$ vector of asset returns and it follows a copula–GARCH model with joint distribution given by:

$$F(\mathbf{r}_t | \boldsymbol{\mu}_t, \mathbf{h}_t) = C(F_1(r_{1t} | \mu_{1t}, h_{1t}), \ldots, F_n(r_{nt} | \mu_{nt}, h_{nt})) \quad (9)$$

where F_i and C are the conditional distribution and the copula function, respectively.

The conditional mean $E[r_{it} | \mathfrak{F}_{t-1}] = \mu_{it}$ is a linear function of its one-lag past returns, and it follows an ARMA(1,1) process. The conditional variance h_{it} follows a GARCH(1,1) process based on model selection criteria, such as the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). They are defined as:

$$r_{it} = \mu_{it} + \theta_1(r_{it-1} - \mu_{it}) + \theta_2 \epsilon_{it-1}^2 + \epsilon_{it}, \quad \epsilon_{it} = \sqrt{h_{it}} z_{it} \quad (10)$$

$$h_{it} = \omega + \alpha_1 \epsilon_{it-1}^2 + \beta h_{it-1} \quad (11)$$

where z_{it} are i.i.d. random variables, which conditionally follow Johnson's reparametrized SU distribution, $z_{it} \sim JSU(\mu, \sigma, \nu, \tau)$ (Ghalanos, 2015). In the four parameters, μ, σ are mean and standard deviation for all values of the skew and shape parameters ν and τ respectively. The dependence structure is modeled using elliptical copulas with conditional correlation $\mathbf{R}_t$ and constant shape parameter τ. The conditional density with a Gaussian copula and a Student-t copula, respectively, are given by:

$$c_t(u_{it}, \ldots, u_{nt} | \mathbf{R}_t) = \frac{f_t\left(F_i^{-1}(u_{it}), \ldots, F_i^{-1}(u_{nt}) | \mathbf{R}_t\right)}{\prod_{i=1}^n f_i\left(F_i^{-1}(u_{it})\right)} \quad (12)$$

$$c_t(u_{it}, \ldots, u_{nt} | \mathbf{R}_t, \tau) = \frac{f_t\left(F_i^{-1}(u_{it} | \tau), \ldots, F_i^{-1}(u_{nt} | \tau) | \mathbf{R}_t, \tau\right)}{\prod_{i=1}^n f_i\left(F_i^{-1}(u_{it} | \tau) | \tau\right)} \quad (13)$$

where $u_{it} = F_{it}(r_{it} | \mu_{it}, h_{it}, \nu_t, \tau_i)$ is the probability integral transformed values by F_{it} estimated via the GARCH process and $F_i^{-1}(u_{it} | \tau)$ represents the quantile transformation. Finally, the joint density of the estimation is defined by:

$$f(\mathbf{r}_t | \boldsymbol{\mu}_t, \mathbf{h}_t, \mathbf{R}_t, \tau)$$
$$= c_t(u_{it}, \ldots, u_{nt} | \mathbf{R}_t, \tau) \prod_{i=1}^n \frac{1}{\sqrt{h_{it}}} f_{it}(z_{it} | \nu_i, \tau_i). \quad (14)$$

In this study, we estimate each conditional correlation via the function **cgarchspec** command in the R package *rmgarch* implementing the Gaussian and Student-t copulas.

We propose a linear state space regression model combined with a Copula–DCC–GARCH model in order to predict volatility. We use copula functions to estimate the DCC ($\hat{\rho}_t$) described in Eq. (2). In particular, our model employs Gaussian and Student-t copulas in order to estimate the conditional covariance matrix. The time-varying regression model with a Copula–DCC–GARCH based on the Gaussian copula is called hereafter "GCTVR" while the model with the Student-t is called "TCTVR".

4. Forecast evaluation

4.1. An illustrated example with forecast measures

To compare the predictive power of our suggested model with Kim et al.'s (2016), we use the same return data as employed in

Table 1
Forecast evaluation.

		MSE_1	MSE_2	MAE_1	MAE_2
Apple, Inc.	LTVR	0.6887	25.3849	0.5639	1.5705
	GCTVR	**0.6597**	25.0924	**0.5500**	1.5173
	TCTVR	0.6627	**22.7786**	0.5537	**1.5133**
Amazon.com	LTVR	0.9914	105.0816	0.6250	2.3004
	GCTVR	**0.7050**	37.9482	**0.5514**	**1.6451**
	TCTVR	0.7096	**29.8331**	0.5600	1.6454
IBM	LTVR	0.4617	20.9633	0.4684	1.3737
	GCTVR	**0.4125**	**13.8545**	**0.4475**	**1.2149**
	TCTVR	0.4399	14.5030	0.4582	1.2596
GE	LTVR	0.3219	15.2375	0.3942	1.0528
	GCTVR	**0.3124**	**13.3275**	**0.3898**	**1.0158**
	TCTVR	0.3250	14.4241	0.3976	1.0492
Walmart	LTVR	**0.6779**	28.2098	**0.5583**	1.5949
	GCTVR	0.6788	28.6299	0.5586	**1.5948**
	TCTVR	0.6888	**28.0843**	0.5623	1.6067

their study. We work with daily returns on the S&P 500 index. In addition, the daily stock returns dataset contains actively traded US firms such as Apple Inc., Amazon.com, IBM, General Electric, and Walmart from 2 January 2002 to 3 June 2015. All stock index prices are retrieved from http://quote.yahoo.com. The total number of observations is 3378. We estimate the time-varying conditional correlation coefficients by implementing copula functions.

Finally, we use four statistical loss functions to evaluate the forecast accuracy, such as

$$MSE_1 \equiv n^{-1} \sum_{t=1}^n (\sigma_t - h_t)^2 \qquad MSE_2 \equiv n^{-1} \sum_{t=1}^n (\sigma_t^2 - h_t^2)^2$$

$$MAE_1 \equiv n^{-1} \sum_{t=1}^n |\sigma_t - h_t| \qquad MAE_2 \equiv n^{-1} \sum_{t=1}^n |\sigma_t^2 - h_t^2|.$$

These functions with mathematical simplicity are the popular measures to evaluate forecasting performance in the literature (e.g., Hansen and Lunde, 2005 and West and Cho, 1995). In the loss functions, as implemented in Kim et al. (2016), we still use the squared returns r_t^2 as a proxy for the latent volatility σ_t^2. They justify the use of the squared return as a proxy for volatility. In particular, to test the forecasting power of our proposed approach, we include the LTVR model by Kim et al. (2016) in the competing candidate. Therefore, the corresponding volatility forecast h_t is obtained from three different volatility models.

Table 1 represents the results from the accuracy comparison of three different models. From the selected firms' return data, Table 1 shows that the GCTVR model provides the most accurate forecasts in terms of all loss functions. Our proposed models with Gaussian and Student-t copula functions increase the accuracy in forecasting the Apple stock return volatility. The GCTVR model has the smallest values in terms of MSE_1 and MAE_1 compared to TCTVR and LTVR while the TCTVR model has the smallest MSE_2 and MAE_2 given alternatives. Our empirical results also show that for Amazon.com the GCTVR model has the smallest values across the accuracy criteria, except for MSE_2. In particular, In the cases of IBM and GE, the GCTVR outperforms existing alternatives in terms of all four loss functions. It is worth mentioning that the LTVR model shows better accuracy for the Walmart return series in terms of the MSE_1 and the MAE_1 loss functions. However, the differences in the accuracy between LTVR and our proposed models are negligible.

4.2. Superiority prediction tests

In this subsection, we employ the more robust tests such as Hansen's (2005) superior predictive ability (SPA) and White's (2000) reality check (RC). These statistical tests are useful for

Table 2
Tests for reality check and superior predictive ability.

Benchmark	Competing	p-value	Apple	Amazon	IBM	GE	Walmart
GCTVR	LTVR	RC	1.000	1.000	1.000	1.000	1.000
		SPA	1.000	1.000	1.000	1.000	1.000
TCTVR	LTVR	RC	1.000	1.000	1.000	1.000	1.000
		SPA	1.000	1.000	1.000	1.000	1.000
GCTVR	TCTVR	RC	0.087	0.131	1.000	1.000	0.103
		SPA	0.077	0.112	1.000	1.000	0.103

Notes: The table presents p-values of the RC and SPA test.

selecting a better forecasting model than the model currently being used in forecasting. Our goal here is to test whether our proposed GCTVR and TCTVR models are outperformed by an alternative forecast, LTVR, proposed by Kim et al. (2016). We conduct a robustness check in order to evaluate whether a presented superiority in a forecasting performance of our copula models in the previous subsection is statistically significant or could have occurred by chance. To do this, we consider our proposed model with copula functions as the benchmark model individually against alternative forecast of the LTVR model of Kim et al. (2016), which is noted as a. The performance measure relative to a benchmark model is defined as:

$$D_{a,t} \equiv \phi(\psi_t, \delta_{0,t-1}) - \phi(\psi_t, \delta_{a,t-1}),$$

where $\phi(\psi_t, \delta_{a,t-1}) = (\psi_t - \delta_{a,t-1})^2$; ψ_t represents squared returns as a volatility proxy, r_t^2; $\delta_{a,t-1}$ is a volatility forecast of the alternative forecasting model based on $t-1$ information. Assuming that $D_a \equiv E(D_{a,t})$ is well defined, the null hypothesis in this analysis is that:

$$H_0 : D_a \leq 0.$$

Following Hansen (2005), we impose all the same assumptions in this test framework.

The estimation results are presented in Table 2. The p-values of the RC and SPA tests are provided for each sample firm. The p-values clearly show that there is no evidence that the proposed models with copula functions are outperformed even if a moderate (1%) significance level is used. The p-values do not differ much across the sample firms, which verifies that our results are not sensitive. The results show that our proposed models are clearly preferred, which confirms the earlier findings from Table 1.

In Table 1, we found that the GCTVR model provides more accurate forecasts than TCTVR does. Finally, we take the GCTVR as the benchmark model and consider the TCTVR model as an alternative. The last two rows in Table 2 report the RC and SPA test results. The p-values clearly show that there is no statistical evidence that the GCTVR model is outperformed at any significance level. The empirical test provides evidence that the GCTVR model is clearly preferred among alternatives.

5. Conclusion

This paper extends the LTVR model (Kim et al., 2016) by implementing copula functions in order to estimate the DCC. We consider the most common copula functions such as Gaussian and Student-t copulas. In the empirical analysis, we evaluate its performance using the US daily stock returns against the model proposed by Kim et al. (2016). In particular, we compare forecast errors generated from each model under various assessment measures. The empirical results show that our proposed models significantly improve the predictive ability of Kim et al.'s (2016) model. In particular, the linear time-varying regression model with the DCC estimated by a Gaussian copula function outperforms the competing models in terms of its ability to forecast the conditional variance.

References

Bartram, S.M., Taylor, S.J., Wang, Y.H., 2007. The Euro and European financial market dependence. J. Banking Finance 31 (5), 1461–1481.

Bollerslev, T., 1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. Rev. Econ. Stat. 72 (3), 498–505.

Bollerslev, T., Chou, R.Y., Kroner, K.F., 1992. ARCH modeling in finance: A review of the theory and empirical evidence. J. Econometrics 52 (1–2), 5–59.

Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econom. Statist. 20 (3), 339–350.

Engle, R., 2004. Risk and volatility: Econometric models and financial practice. Amer. Econ. Rev. 94 (3), 405–420.

Frey, R., McNeil, A.J., McNeil, A.J., McNeil, A.J., 2001. Modelling dependent defaults, ETH, Eidgenössische Technische Hochschule Zürich, Department of Mathematics.

Ghalanos, A., 2015. The rmgarch Models: Background and Properties. (Version 1.2–8).

Hansen, P.R., 2005. A test for superior predictive ability. J. Bus. Econom. Statist. 23 (4).

Hansen, P.R., Lunde, A., 2005. A forecast comparison of volatility models: Does anything beat a GARCH (1,1)?. J. Appl. Econometrics 20 (7), 873–889.

Jondeau, E., Rockinger, M., 2006. The copula-GARCH model of conditional dependencies: An international stock market application. J. Int. Money Finance 25 (5), 827–853.

Kim, J.M., Jung, H., Qin, L., 2016. Linear time-varying regression with a DCC-GARCH model for volatility. Appl. Econ. 48 (17), 1573–1582.

Li, D.X., 1999. On default correlation: A copula function approach. J. Fixed Income 9 (4), 43–54.

Manner, H., Reznikova, O., 2012. A survey on time-varying copulas: Specification simulations, and applicationbel. Econ. Rev. 31 (6), 654–687.

Nakajima, J., 2011. Time-Varying Parameter VAR Model with Stochastic Volatility: An Overview of Methodology and Empirical Applications, Tech. Rep.. Institute for Monetary and Economic Studies, Bank of Japan.

Patton, A.J., 2006. Modelling asymmetric exchange rate dependence. Internat. Econom. Rev. 47 (2), 527–556.

Sklar, A., 1973. Random variables, joint distribution functions, and copulas. Kybernetika 9 (6), 449–460.

Sklar, M., 1959. Fonctions de Répartition À N Dimensions et Leurs Marges, Université Paris 8.

Tse, Y.K., Tsui, A.K.C., 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. J. Bus. Econom. Statist. 20 (3), 351–362.

West, K.D., Cho, D., 1995. The predictive ability of several models of exchange rate volatility. J. Econometrics 69 (2), 367–391.

White, H., 2000. A reality check for data snooping. Econometrica 68 (5), 1097–1126.