Copula-Based Vector Autogressive Models for Bivariate Cointegrated Data

Hideaki Taima and Ana Maria L. Tabunda, Ph.D.
University of the Philippines Diliman

The copula method is well applied in finance and actuarial science but its application in economic studies is limited and its use in the cointegration framework virtually nil. This paper explores the use of copula method to analyze the remaining dependence after a cointegration relationship is modeled. Specifically, simulated data is used to characterize the behavior of the dependence parameter estimates of several copulas fitted to the distribution of the residuals after cointegrated Vector Autoregressive (VAR) and Vector Error-Correction Mechanism (VECM) models are fitted, as well as evaluate the forecasting ability of the copula-based models. The Clayton, Frank, Gaussian, Gumbel and Plackett copulas are used and are compared on the basis of bias, root mean square error (RMSE) and maximum likelihood. The density forecasting ability of the copula-based VAR and VECM is then compared with that of standard models via conditional Kullback-Leibler Information Criterion (KLIC) divergence measure using simulated and empirical data. The simulation results indicate that the copula-based models generally have better density forecasting ability than standard VAR and VECM models, a finding that is supported in the application of a copula-based VAR to empirical data.

Keywords: Copula, Cointegration, VAR, VECM

1. Introduction

It is well known that economic variables have certain relationships among themselves, for example comovements. Cointegration is one of the breakthrough approaches to characterizing several series through the imposition of long-run co-movement restriction. Substantial work has since been introduced to model multivariate time series following the approach of cointegration. However, satisfying the assumption of joint normality of the series is usually difficult in applications. In this paper we explore the use of copulas to model the dependence between series that are not necessarily normal.

The copula is a function that links multivariate distribution functions to their one-dimensional marginal distribution functions (Nelson, 2006). Specifically, the joint distribution function of a random vector can be represented via Sklar's theorem in terms of the copula function with the marginal distribution functions of the components of the vector as the arguments of the copula function. The application of the copula helps separate the parameters of the marginal distributions from their intrinsic association as captured by the dependence parameters. An attractive feature of this approach is that the copula and the association parameter are invariant under continuous and monotonically increasing transformations of the marginal variables. Hence copulas have an advantage when the interest centers on the intrinsic association among the marginals (Joe, 1997; Kim et al., 2007).

The copula method has been used to analyze multivariate data with flexible functional forms rather extensively in finance and in actuarial science. For example, the use of the copula in conjunction with the generalized autoregressive conditional heteroskedastic model (GARCH), or the copula-based GARCH, facilitates the analysis of several series of returns that are non-normal (Jondeau and Rockinger, 2006) and allows for flexibility in modeling the conditional dependence structure between the Deutsche mark and the Japanese yen relative to the US dollar (Patton, 2005). Various analyses in finance show better results from the use of copulas in multivariate GARCH.

The application of the copula method to economic studies, however, is limited and its use in the cointegration framework is virtually nil. The copula-based studies with applications in econometrics include Granger et al. (2006), which looks into the multivariate GARCH model, and Mitchell (2007), which models dependence between the survey of professional forecasters (SPF)'s inflation and output growth density forecasts of the US economy. Bianchi et al. (2009), on the other hand, uses the copula-based vector autoregression (VAR) approach to forecast industrial production series in the core European Monetary Union (EMU) countries and provide evidence that the copula-VAR model outperforms or, at worst, is at par with standard VAR models.

This study explores the use of copula-based vector autoregressive models under a cointegration framework to yield insights on the effect of any remaining dependence on forecast models for bivariate cointegrated variables. Specifically, we use simulated data to evaluate the performance of copula-based cointegrated VAR and vector error correction mechanism model (VECM) against standard models in the presence of dependence. Following Vuong (1989) and Bao et al. (2007), we employ the Kullback-

Leibler information criterion (KLIC) divergence measure between two conditional densities, described in Section 3.3, to evaluate the density forecasting ability of the copula-based models.

One limitation of this study is that it considers only two symmetric distributions for the error terms used in the data generating model in the simulations – the bivariate normal distribution and the bivariate student's t distribution. Thus the conclusions drawn in the paper will be applicable only for such situations.

This paper is organized as follows: Section 2 introduces the framework of the model and reviews the definition of copula. Section 3 presents simulation results under different scenarios depending on the assumptions on the form of the error distribution, the strength of the dependence between the component error terms, and the model fitting method. In particular, the simulation section compares five forms of copula that were used in estimating the distribution of the residuals obtained from fitting bivariate VAR and VECM models under three dependence scenarios for the model error terms. The copulas are compared on the basis of accuracy, precision as well as fit to the distribution of the data, while the copula-based VECM and VAR models are compared with standard VECM and VAR models in terms of density forecasting ability. The copula-based models are then fitted to Japan consumption and income and compared with standard VAR models. Section 4 gives a summary of the results and presents recommendations.

2. Copula-Based Bivariate Cointegrated Vector Autoregressive Model

The VAR and VECM models are useful in studying linear dynamic relationships among several time series variables. The VAR is the generalization of the univariate autoregression model and can be used whether the vector variables are cointegrated or not.

Assuming the y_t's are I(1), a VAR process could be modeled as follows:

$$\Delta y_{1,t} = \sum_{i=1}^{p} A_{11i} \Delta y_{1,t-i} + \sum_{i=1}^{p} A_{12i} \Delta y_{2,t-i} + \sqrt{v_{1,t}}\, e_{1,t}$$

$$\Delta y_{2,t} = \sum_{i=1}^{p} A_{21i} \Delta y_{1,t-i} + \sum_{i=1}^{p} A_{22i} \Delta y_{2,t-i} + \sqrt{v_{2,t}}\, e_{2,t} \qquad (1)$$

where $e_{1,t}$ and $e_{1,t}$ have mean 0 and variance σ^2_{1et} and σ^2_{2et}, respectively.

The bivariate Vector Error Correction Mechanism (VECM), on the other hand, is given by

$$\Delta y_{1,t} = \alpha_1 (y_{1,t-1} + \beta_1 y_{2,t-1}) + \sum_{i=1}^{p-1} \Gamma_{11i} \Delta y_{1,t-i} + \sum_{i=1}^{p-1} \Gamma_{12i} \Delta y_{2,t-i} + \sqrt{u_{1,t}}\, \varepsilon_{1,t}$$

$$\Delta y_{2,t} = \alpha_2 (y_{1,t-1} + \beta_1 y_{2,t-1}) + \sum_{i=1}^{p-1} \Gamma_{21i} \Delta y_{1,t-i} + \sum_{i=1}^{p-1} \Gamma_{22i} \Delta y_{2,t-i} + \sqrt{u_{2,t}}\, \varepsilon_{2,t} \qquad (2)$$

where $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ have 0 and mean variance σ_{1t}^2 and σ_{2t}^2 respectively.

The VECM can be viewed as a restricted VAR. It incorporates the cointegrating relationship(s) in a VAR model equation.

Now an expression for the conditional joint distribution of the error terms of VECM and VAR models can be obtained through Sklar's theorem (1959). For example, the conditional joint distribution of $U(\varepsilon_{1,t}, \varepsilon_{2,t})$ in equation (2) can be expressed in the following manner:

$$(\varepsilon_{1,t}, \varepsilon_{2,t}) \sim U(\varepsilon_{1,t}, \varepsilon_{2,t};\theta) = C_t\,(F_{1,t}\,(\varepsilon_{1,t};\delta_1), \mathrm{F}_{2,\mathrm{t}}\,(\varepsilon_{2,t};\delta_2);\ \varphi) \qquad (3)$$

where $\delta_{1,}\ \delta_2$ and θ are marginal parameters and copula, respectively.

The use of copulas allows one to model different types of dependence in a flexible way and allows for various marginal distributions. Through the use of copulas, it is possible for the marginal distributions to have different degrees of freedom. For instance distribution F_1 may have a student's t-distribution with v_1 degree of freedom, while distribution F_2 has a student's t-distribution with v_2 degree of freedom, or F_2 has a normal distribution.

2.1 Copula

The definition of a bivariate copula is given in Nelson (2006) as follows:

Definition. A 2-dimensional copula is a function C whose domain is $\mathbf{I}^2$, where the unit square $\mathbf{I}^2$ is the Cartesian product **I**×**I**, **I**=[0,1]. The function C has the following properties:

1. For every u, v in **I**, where **I**=[0,1],

$$C(u,0)=C(0,v)=0, \qquad C(u,1)=1 \text{ and } C(1,v)=v,$$

2. For every u_1,u_2, v_1,v_2 in **I** such that $u_1 \leq u_2$, *and* $v_1 \leq v_2$,

$$C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1) \geq 0.$$

The following theorem provides the framework for the application of copulas.

Theorem (Sklar's Theorem). Let H be a joint distribution function with margins F and G. Then there exists a copula C such that for all x, y in $\overline{R}$

$$H(x,y) = C\big(F(x), G(y)\big) \tag{4}$$

If F and G are continuous, then C is unique; otherwise C is uniquely determined on Range $F \times$ Range G. Conversely, if C is a copula and F and G are distribution functions, then the function H defined by (4) is a joint distribution function with margins F and G.

To obtain the best-fitting model, Granger et al. (2006) tested several forms of copulas. As mentioned in their paper, there is no guidance from economic theory regarding the choice of copula, thereby necessitating modeling and comparing several copulas and choosing the best in terms of maximum log-likelihood. In this paper we use the one-parameter copulas commonly used in finance and actuarial science, which are: Gaussian copula, Clayton copula, Gumbel copula, Frank copula and Plackett copula. Each copula allows for different dependence properties.

The *Gaussian copula* is the copula derived from the multivariate Gaussian distribution. The bivariate Gaussian Copula is obtained by the inversion method as follows:

$$\begin{aligned} C_G(u,v) &= \Phi_{pxy}\big(\Phi^{-1}(u), \Phi^{-1}(v)\big) \\ &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-p_{xy}^2)^{1/2}} \exp\left[\frac{-(s^2 - 2p_{xy}st + t^2)}{2(p_{xy}^2 - 1)}\right] dsdt \end{aligned} \tag{5}$$

where Φ_{pxy} is the CDF of standard bivariate distribution with linear correlation parameter *pxy* and Φ is CDF of standard normal distribution.

The *Clayton, Frank and Gumbel copulas* belong to the Archimedean family of copulas. The Archimedean copula with generator φ is given by

$$\text{C}\ (u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \tag{6}$$

where φ is a continuous, convex, strictly decreasing function from [0,1] to $[0,\infty]$ such that $\varphi(1)=0$, and $\varphi^{[-1]}$ is the pseudo-inverse function of given by

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \le t \le \varphi(0) \\ 0, & \phi(0) \le t \le +\infty \end{cases} \tag{7}$$

The specific copula functions and generators corresponding to the Clayton, Frank and Gumbel copulas are shown in Table 1.

Table 1. Forms of Selected Archimedean Copulas

	$C_\theta(u,v)$	$\varphi_\theta(t)$	$\varphi_\theta(t)$
Clayton family	$\max\left(\left[u^{-\theta}+v^{-\theta}-1\right]^{-1/\theta},0\right)$	$\frac{1}{\theta}(t^{-\theta}-1)$	$[-1,0)\setminus(0,\infty)$
Frank family	$-\frac{1}{\theta}\left(1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right)$	$-\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$(-\infty,\infty)\setminus\{0\}$
Gumbel family	$\exp\left(-\left[(-\ln u)^{\theta}+(-\ln v)^{\theta}\right]^{-1/\theta}\right)$	$(-\ln t)^{\theta}$	$[1,\infty)$

The Placket copula, on the other hand, has the following form,

$$C(u,v)=\frac{\left[1+(\theta-1)(u+v)\right]-\sqrt{\left[1+(\theta-1)(u+v)\right]^2-4uv\theta(\theta-1)}}{2(\theta-1)} \qquad (8)$$

for $\theta>0$, $\theta\neq1$.

2.2 Estimation

Let (Z_{1t},Z_{2t}) denote a continuous bivariate random vector where t=1,…,T. Let $F_i(Z_{it})$ and $f_i(z_{it})$ denote the cumulative distribution function (cdf) and probability density function (pdf) of Z_{it}, respectively. Let $U_i= F_i(Z_{it})$ and let $C(U_{1t},U_{2t})$ denote the joint cdf of (U_{1t},U_{2t}), $c(u_{1t},u_{2t})$ denote the pdf corresponding to $C(u_{1t},u_{2t})$, and $H(Z_{it},Z_{2t})$ and $h(z_{1t},z_{2t})$ denote the cdf and pdf of (Z_1,Z_2), respectively. Then the joint density function $h(z_{1t},z_{2t})$of (Z_{1t},Z_{2t}) can be expressed via Sklar's theorem in the form of (4) as follows:

$$h(z_{1t}, z_{2t}) = c\{F_1(z_{1t}), F_2(z_{2t})\}\, f_1(z_{1t})\, f_2(z_{2t}) \qquad (9)$$

The log-likelihood function is given by:

$$L(\theta)=\sum_{t=1}^{T}\log\left[c\left(F_1(z_{1t}),F_2(z_{2t})\right)f_1(z_{1t})f_2(z_{2t})\right] \qquad (10)$$

where θ is the set of all parameters of both the marginal distributions and the copula.

Hence, given a set of marginal pdf's and copula, the log-likelihood may be written as in (10), and the maximum likelihood estimator obtained, where $\hat{\theta}_{MLE} = \max_{\theta \in \Theta} L(\theta)$.

In the simulation study, we employ multi-stage maximum log likelihood (MSML) because the computational burden is lighter especially when there are a lot of parameters to be estimated. MSML was proposed for estimating copula by Joe and Xu (1996) under the name inference functions for margins (IFM). The estimation procedure is performed in two steps. In the first step, parameter estimates are obtained separately by optimizing the univariate likelihoods based on the margins. This is then followed by optimizing the multivariate likelihood treated as a function of the copula parameter. In addition the canonical maximum likelihood (CML) method is employed in the empirical study. The CML method estimates each marginal distribution nonparametrically by the empirical distribution. This method is used to check the consistency of the dependence parameter with its IFM since it is difficult to determine the true distributions in empirical data.

3. Application

3.1.1 Simulation Design

The data employed in this study is generated from a first order cointegrated bivariate VAR model which has the form:

$$\Delta Y_t = \begin{bmatrix} -1 \\ -0.5 \end{bmatrix} \begin{bmatrix} 0.5 & 1 \end{bmatrix} Y_{t-1} + \varepsilon_t \qquad (11)$$

where $Y_t = (y_{1t}, y_{2t})'$ and error term $\varepsilon_t = (e_{1t}, e_{2t})$.

We consider two scenarios for the distribution of the error term, i.e., bivariate normal distribution with the marginal distributions having mean 0 and variance 1, and bivariate student's t- distribution with 3 degrees of freedom (d.f). In each scenario, three values of the dependence measure are considered corresponding to Kendall's tau of 0.3, 0.5 and 0.9, or equivalently, a linear correlation coefficient r of 0.454, 0.707 and 0.908, given the relation $\tau = \frac{2}{\pi} \arcsin \rho$ that holds for (essentially) all elliptical distributions (Lindskog, 2000). Thus estimates of the copula parameter of normally- and student's t-distributed variables can be obtained given the linear correlation coefficient (Kendall's tau).

The scenarios and cases considered in the simulations are as follows:

Scenario I: The error term has a bivariate normal distribution

Case (i): $\varepsilon_t = (e_{1t}, e_{2t}) \sim N(0,1)$ with $\tau=0.3$

Case (ii): $\varepsilon_t = (e_{1t}, e_{2t}) \sim N(0,1)$ with $\tau=0.5$

Case (iii): $\varepsilon_t = (e_{1t}, e_{2t}) \sim N(0,1)$ with $\tau=0.9$

Scenario II: The error term has a bivariate student's t-distribution

Case (iv): $\varepsilon_t = (e_{1t}, e_{2t}) \sim t_3$ with $\tau=0.3$

Case (v): $\varepsilon_t = (e_{1t}, e_{2t}) \sim t_3$ with $\tau=0.5$

Case (vi): $\varepsilon_t = (e_{1t}, e_{2t}) \sim t_3$ with $\tau=0.9$

In each case under each scenario, 200 samples each consisting of 200 data points are generated using equation (11) with the error terms following the aforementioned distributions and corresponding Kendall's tau. A sample size of 200 for each data set is considered with a view to applying the procedure to an actual data set that has less than 200 observations. Each generated sample is estimated via VECM and VAR model. Copula parameters are estimated from the error terms of marginal models to observe the performance of the copulas after filtering by VECM(1) and VAR(1).

For given marginals and method of copula estimation, let θ_i be the one-parameter copula estimator of q for the *i*th generated sample, i=1,…,200, and θ_0 represent the true value of the copula parameter given Kendall's tau and let N denote the number of generated samples. We consider two criteria, root mean squared error (RMSE) and estimated bias where $RMSE = \left[N^{-1}\sum\{\theta_i - \theta_0\}^2\right]^{1/2}$ and estimated bias = $N^{-1}\sum\theta_i - \theta_0$ in evaluating the five copulas tried.

3.1.2 Bias and RMSE and Log-Likelihood of Models of Copula

Tables 2 and 3 present the bias and RMSE of each copula model measured assuming the error terms follow a bivariate normal distribution and a bivariate student's t-distribution, respectively. It is seen from these tables that, with the exception of the Clayton copula in some cases and the Plackett copula in many cases, the biases obtained in estimating the dependence parameter are relatively small compared to the value of the true parameter irrespective of whether the true distribution of the error term is normal or Student's t-distribution.

In addition, irrespective of the distribution of the error term and copula used, the bias and RMSE of the copula parameter in the VECM model are lower than that produced by VAR when the dependence coefficient is relatively low($\tau=0.3$). But when the dependence coefficient is moderate ($\tau=0.5$), the VAR model produces lower bias and RMSE when the error distribution is normal for all copulas except the Clayton copula. However, the results are reversed when the error terms follow a Student's t-distribution and τ=0.5; the VECM model this time has generally lower bias and RMSE for all copulas except the Clayton copula. On the other hand, when the dependence coefficient is high (τ=0.9), in contrast to the results when the coefficient is low, the RMSE and bias of the estimates obtained from the VAR are generally lower compared to the corresponding VECM-based values.

Table 2. Bias and MSE of Copulas in Bivariate Normal Distribution*

Copula	Model		τ=0.3	τ=0.5	τ=0.9
Gaussian		**TRUE θ_0**	0.45	0.71	0.99
	VECM	**Bias**	0.09	0.05	0.03
		RMSE	0.14	0.07	0.03
	VAR	**Bias**	0.14	0.04	0.01
		RMSE	0.16	0.06	0.01
Gumbel		**TRUE θ_0**	1.43	2.00	10.00
	VECM	**Bias**	0.17	0.07	4.95
		RMSE	0.63	0.23	5.00
	VAR	**Bias**	0.30	0.04	3.22
		RMSE	0.02	0.17	3.30
Clayton		**TRUE θ_0**	0.86	2.00	18.00
	VECM	**Bias**	0.08	0.41	12.45
		RMSE	0.79	0.53	12.51
	VAR	**Bias**	0.25	0.46	10.06
		RMSE	1.19	0.53	10.13
Frank		**TRUE θ_0**	2.92	5.74	20.90
	VECM	**Bias**	1.10	0.73	1.68
		RMSE	3.01	1.27	3.23
	VAR	**Bias**	1.76	0.62	4.74
		RMSE	4.27	1.04	5.23
Plackett		**TRUE θ_0**	4.00	11.40	530.00
	VECM	**Bias**	3.77	1.39	431.24
		RMSE	15.76	3.99	432.66
	VAR	**Bias**	6.84	0.94	341.70
		RMSE	28.42	2.90	344.32

* *Bias and MSE were computed on 200 random samples of length 200 from bivariate cointegrated autoregressive process of order one.*

The results in Tables 2 and 3 suggest that, when the dependence between the error terms is high and therefore the departure from the long-run equilibrium relationship in the previous period does not sufficiently explain the variation in the two variables, simply fitting a copula-based VAR model on the data tends to yield more accurate and more precise dependence estimates. But when the dependence in the error terms not accounted for by the cointegrating relationship is low, then there is an advantage to incorporating the cointegrating relationship in the model.

However, when the dependence between the error terms is moderate, the choice of copula and the model fitting method is not as clear cut. Thus greater care has to be exercised in choosing the model fitting method as well as the copula when the dependence between the error terms is suspected to be moderate and accurate and precise estimates of the dependence parameter are desired.

Table 3. Bias and MSE of Copulas in Bivariate Student's-t Distribution*

Copula	Model		τ=0.3	τ=0.5	τ=0.9
		TRUE θ_0	0.09	0.04	0.03
	VECM	**Bias**	0.09	0.04	0.03
Gaussian		**RMSE**	0.13	0.08	0.03
	VAR	**Bias**	0.15	0.05	0.01
		RMSE	0.16	0.07	0.02
		TRUE θ_0	1.43	2	10
	VECM	**Bias**	0.17	0.21	4.97
Gumbel		**RMSE**	0.25	0.32	5.02
	VAR	**Bias**	0.28	0.23	3.28
		RMSE	0.33	0.33	3.38
		TRUE θ_0	0.06	0.14	1.27
	VECM	**Bias**	0.09	0.21	12.47
Clayton		**RMSE**	0.27	0.45	12.52
	VAR	**Bias**	0.24	0.21	10.31
		RMSE	0.34	0.41	10.38
		TRUE θ_0	2.92	5.74	20.9
	VECM	**Bias**	1.05	1.19	2.12
Frank		**RMSE**	1.38	1.65	3.41
	VAR	**Bias**	1.66	1.31	3.45
		RMSE	1.84	1.69	4.40
		TRUE θ_0	4	11.4	530
	VECM	**Bias**	2.75	5.31	430.96
Plackett		**RMSE**	3.50	7.05	432.13
	VAR	**Bias**	4.22	5.72	337.99
		RMSE	4.72	7.46	341.13

* *Bias and MSE were computed on 200 random samples of length 200 from bivariate cointegrated autoregressive processes of order one*

Tables 4 and 5 present the number and percentage of models which attained the maximum log-likelihood among the 200 samples of 200 observations, under assumptions of a normal distribution and a student's t distribution, respectively.

As may be expected, the Gaussian copula fits the marginal model better when the error term is normally distributed irrespective of whether the marginal models are estimated by VECM or VAR (Table 4). Nevertheless when τ=0.3, there is still about 30% to one-third chance that other copulas fit better than the Gaussian copula whether VAR or VECM is fitted. However, when the dependence coefficient is higher (τ=0.5 and τ=0.9), the Gaussian copula is far superior to the other copulas.

Table 4. Number and Percentage which Attained Maximum Log-likelihood in Bivariate Normal Distribution*

Copula	Model		τ=0.3	τ=0.5	τ=0.9
Gaussian	VECM	Number	134	162	177
		percentage	67%	81%	89%
	VAR	Number	144	160	161
		percentage	72%	80%	81%
Gumbel	VECM	Number	20	21	18
		percentage	10%	11%	9%
	VAR	Number	20	21	17
		percentage	10%	11%	9%
Clayton	VECM	Number	16	3	0
		percentage	8%	2%	0%
	VAR	Number	12	5	0
		percentage	6%	3%	0%
Frank	VECM	Number	23	9	5
		percentage	12%	5%	3%
	VAR	Number	19	11	20
		percentage	10%	6%	10%
Plackett	VECM	Number	7	5	0
		percentage	4%	3%	0%
	VAR	Number	5	3	2
		percentage	3%	2%	1%

* *log-likelihood were computed on 200 random samples of length 200 from bivariate cointegrated autoregressive processes of order one*

When the error term has a student's t distribution, the Gumbel copula and the Plackett copula appear to be good candidates for estimating the marginal distributions when τ=0.3, irrespective of whether VECM or VAR models are used (Table 5). This outcome corroborates the results of Granger et al. (2006) where the Gumbel copula fitted best with parameter coefficient 1.0977 even as the said study used the skewed t-distribution for the marginal density.[1] When τ=0.5 the Gaussian copula is added to these two candidates. When τ=0.9, the Gaussian copula unexpectedly outperforms the other copulas, yielding the highest number of resamples with maximum log-likelihood, especially when used in conjunction with VECM.

Thus, the results in Tables 4 and 5 indicate that the Gaussian copula is, as may be expected, a very good choice when the error terms can be assumed to follow a normal distribution, and its performance is even better the error terms are moderately or highly correlated. But when the correlation between the error terms is low, however, there is a 30% to 33% chance that another copula can model the distribution of the data better than the Gaussian copula. But in the absence of guidance as to which copula will perform better, the Gaussian copula is a safe choice.

Table 5. Number and Percentage which Attained Maximum Log-likelihood in Bivariate Student's-t Distribution*

Copula	Model		τ=0.3	τ=0.5	τ=0.9
Gaussian	VECM	Number	18	53	156
		percentage	9%	27%	78%
	VAR	Number	24	48	97
		percentage	12%	24%	49%
Gumbel	VECM	Number	57	59	36
		percentage	29%	30%	18%
	VAR	Number	60	60	58
		percentage	30%	30%	29%
Clayton	VECM	Number	25	7	0
		percentage	13%	4%	0%
	VAR	Number	12	2	1
		percentage	6%	1%	1%
Frank	VECM	Number	3	5	5
		percentage	2%	3%	3%
	VAR	Number	1	4	13
		percentage	1%	2%	7%
Plackett	VECM	Number	87	76	3
		percentage	44%	38%	2%
	VAR	Number	103	86	3
		percentage	52%	43%	16%

* *log-likelihood were computed on 200 random samples of length 200 from bivariate cointegrated autoregressive processes of order one*

Interestingly, the Gaussian copula also performs satisfactorily when the error distribution is student's t, provided the dependence parameter is high and its performance is further improved if the copula-based VECM model is fitted. Thus, the slight loss in accuracy and precision in the dependence estimate can be offset by the better fit afforded by a Gaussian copula even if a VECM model is used in the presence of high correlation among the error terms. For low to moderate dependence, however, the Plackett and Gumbel copulas perform better when the error distribution is student's t.

3.1.3 Comparison of Models

The standard VAR and VECM models are compared with the copula-based models on the basis of the multivariate conditional KLIC divergence measure. This measure is defined as the distance between the true multivariate conditional density $g_t\left(x_{1,t},\ldots,x_{m,t}\middle|F_{t-1};\lambda_m\right)$ and the model based density $f_t\left(x_{1,t},\ldots,x_{m,t}\middle|F_{t-1}\right)$:

$$L^m\left(f_t : g_t, \lambda_m\right) = E\left[\ln f_t(x_{1,t},\ldots,x_{m,t}) - \ln g_t(x_{1,t},\ldots,x_{m,t};\lambda_m)\right]$$

Low values of the KLIC indicate that the model is close to the true density. In this paper, we compare the density forecasts ability of the models by measuring the distances between the measured distribution density and the true distribution density through conditional KLIC divergence measure (Bao et al., 2007). In addition, we approximate the true multivariate distribution density using the non-parametric multivariate product kernel estimator suggested by Scott (1992) and Li and Racine (2007) which is given by:

$$\hat{f} = \frac{1}{nh_1 \cdots h_m} \sum_{t=1}^{n} \left\{ \prod_{j=1}^{m} K\left(\frac{x_j - x_{jt}}{h_j} \right) \right\}$$

where $K(\cdot)$ denotes the Gaussian kernel and h_j denotes kernel bandwidth. The kernel bandwidth is calculated as $\hat{h} = [4/(m+2)]^{1/(m+4)} \cdot \hat{\sigma}_j \cdot n^{1/(m+4)}$ where $\hat{\sigma}_j$ is the standard deviation of x_{jt}.

Table 6 below presents the number of models which attained the lowest KLIC and the total KLIC of models when the error terms are jointly normally distributed. In general, at least one copula-based model attains the most number of fitted models with lowest KLIC in any given dependence scenario, indicating better density forecasting ability compared to standard models.

Table 6. Number of Models which Attained the Lowest KLIC out of 200 Samples and Total KLIC of the Models when the Error Term is Normally Distributed

Models		τ=0.3		τ=0.5		τ=0.9	
		No. with. min KLIC	Total KLIC	No. with min.KLIC	Total KLIC	No. with min. KLIC	Total KLIC
Standard	VECM	11	14.35	29	13.4	8	24.7
Copula-Based	Gauss	24	4.83	15	5.69	0	23.11
VECM	Gumbel	20	5.12	18	5.42	0	22.6
	Clayton	2	17.15	0	26.37	0	37.32
	Frank	29	5.4	64	4.48	0	20.82
	Plackett	11	5.37	13	6.51	0	24.65
Standard	VAR	3	72.17	1	52.28	15	19.35
Copula-Based	Gauss	30	3.27	4	5.68	13	2.96
VAR	Gumbel	16	3.39	12	5.32	60	2.68
	Clayton	0	14.78	0	20.07	0	9.62
	Frank	30	3.04	41	4.07	98	2.51
	Plackett	24	3.58	3	6.15	6	3.8

The simulation results indicate that the Frank copula-based models perform rather well across the three dependence scenarios considered (low, moderate, and high) when the error term is normally distributed, and not the Gaussian-based models as one might expect. When the correlation between the error terms is low, the Gaussian copula-based model is nearly as good as the Frank-based model irrespective of whether the VECM or VAR model is used. However, when the correlation is moderate, the Frank copula-based VECM is best. When the correlation is high, the Frank copula-based VAR is the best choice.

As in the case of a jointly normally distributed error term, at least one copula-based model outperforms the corresponding standard VECM or VAR model when the error term follows a student's t-distribution (Table 7). And as in the normal case, the results suggest that the use of copula-based VECM models is not advisable in the presence of high correlation. However, unlike in the normal case, no single copula is the best choice across the dependence scenarios. The Clayton-based VAR performs best under low correlation, the Gaussian-based VECM under moderate correlation and the Frank-based VAR under high correlation.

Table 7. Number of Models which Attained the Lowest KLIC out of 200 Samples and Total KLIC of the Models when the Error Term is Student's t-distribution

Models		τ=0.3		τ=0.5		τ=0.9	
		No. with. min KLIC	Total KLIC	No. with min.KLIC	Total KLIC	No. with min. KLIC	Total KLIC
Standard	VECM	5	41.2	15	30.51	10	12.81
Copula-Based	Gauss	41	6.34	61	4.25	0	6.72
VECM	Gumbel	8	8.45	7	5.83	3	5.92
	Clayton	41	5.27	21	9.35	0	16.18
	Frank	6	8.6	5	6.37	4	5.2
	Plackett	6	9.41	7	5.95	0	7.23
Standard	VAR	2	77.43	6	31.99	6	24.65
Copula- Based	Gauss	8	8.36	51	4.11	1	3.1
VAR0	Gumbel	2	11.05	6	6	68	2.02
	Clayton	79	3.69	10	9.46	0	10.15
	Frank	2	11.52	6	6.55	106	1.96
	Plackett	0	12.24	5	6.1	2	2.9

3.2 Empirical Study

3.2.1 Data

Copula-based models are tried on the quarterly time series data on the real consumption expenditure and gross national income of Japan from 1957Q1 to 2004Q1. The data, obtained from International Financial Statistics (IFS) published by the International Monetary Fund, is in constant prices in billions of yen.

3.2.2 Tests for Statandionarity and Cointegration

The results of the Augmented Dickey Fuller (ADF) test, Phillip-Perron (PP) test the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test and the Johansen Cointegration test are presented in Tables 8 and 9. Results from the three unit roots tests indicate that the variables are integrated order of one.[2] Moreover, the results of the cointegration test indicate that there is a cointegration relationship between income and consumption.

Table 8. Results of Unit Root Tests

	Test	Level	1st differenced
Logged consumption	ADF	-2.7339	-4.1132**
	PP	-4.0699	319.8787**
	KPSS	0.3494**	0.264239
Loged income	ADF	-2.6091	-5.1188**
	PP	-2.2892	-245.2109
	KPSS	0.4338**	0.1102

** *stands for significant at one percent significance level*

Table 9. Johansen Cointegration Test Results

	Number of Cointegration	Maximal Eigenvalue	Trace
Test Statistics	At most one	0.89	0.89
	None	44.78**	45.66

** *stands for significant at one percent level*

3.2.3 Estimation of Econometric Models and Copulas

We employed Akaike information criterion (AIC) and Schwarz's information criterion (SIC) to determine the appropriate lag length of the VECM and VAR models. These measures indicated a lag length of four periods for both the VECM and VAR models.

Table 10 presents the parameter estimates for the copula-based VAR econometric model. Copula-based VECM models for the data were not pursued further as the error correction terms in the VECM model were not significant, indicating that both consumption and income are not responding to adjustments from disequilibrium. The VECM and VAR models return almost the same estimated coefficients. Thus only the copula-based VAR is compared to the standard VAR model.

Table 10. Estimates of Copula-based VECM and VAR

		VAR			
		Consumption		Income	
intercept		0.007**	(0.002)	0.003^	(0.001)
Lag1	Consumption	-0.466**	(0.062)	-0.0002	(0.047)
	Income	0.397**	(0.125)	-0.015	(0.081)
Lag2	Consumption	-0.250**	(0.077)	0.064	(0.050)
	Income	0.565**	(0.126)	0.224**	(0.082)
Lag3	Consumption	-0.321**	(0.075)	0.020	(0.049)
	Income	0.377**	(0.131)	0.225**	(0.085)
Lag4	Consumption	0.414**	(0.125)	0.027	(0.081)
	Income	0.109	(0.125)	0.138^	(0.081)
ECM			-		-
Residual s.e		0.024		0.015	
d.f (residual s.e.)		175		175	
Multiple R-squared		0.76		0.26	
Adjusted R-squared		0.75		0.23	
F-statistics		69.91		8.016	
d.f (F-statistics)		8;175		8;175	
p-value		0.000		0.000	

Numbers inside parenthesis are standard errors of estimates
*** and * stand for significant at 1 and 0.5 percent significant level, respectively, while ^ denotes significant at 10 percent significant level*

The same copulas used in the simulation study were used to model the dependence between the marginals, except for the Gumbel copula. Optimization problems were encountered when the Gumbel copula was tried for both VECM and VAR models and using both IFM and CML methods. The standard normal distribution was found to fit the data for the margins and the residual distribution passed the goodness-of-fit test.

Table 11 reports estimates of the dependence parameter based on the IFM and CML methods that were applied after fitting the VAR models. The estimated coefficients of the dependence parameter yielded by the IFM and CML methods show small differences for the model residuals. This implies that the marginal models are correctly specified (Kim, et al., 2007b). The dependence estimates obtained imply relatively low dependence. For instance, the estimated parameter coefficients of the Gaussian and Frank copulas imply τ values of 0.2 to 0.3 approximately, that of the Clayton copula indicates a τ value in the 0.1 to 0.3 range, while that of the Plackett copula yields a τ of approximately 0.2 to 0.4. However the calculated value of Kendall's tau on the empirical distribution of the error terms is 0.27, which is not inconsistent with the above results. Among the copulas employed, the Gaussian copula registers the best fit to the data, implying symmetry in the dependence structure.

Table 11. Estimates of the Dependence Parameter of Copula *

	VAR			
	IFM			
	Gauss	Clayton	Frank	Plackett
Parameter estimates	0.406	0.329	2.897	4.501
	(0.056)	(0.109)	(0.535)	(1.0196)
Maximized log-likelihood	16.663	11.284	13.920	16.142
	CML			
	Gauss	Clayton	Frank	Plackett
Parameter estimates	0.379	0.548	2.407	3.541
	(0.059)	(0.123)	(0.477)	(0.772)
Maximized log-likelihood	14.146	13.784	12.849	14.355

** All estimates are significant at a=0.01. Numbers inside parentheses are standard errors of estimates.*

3.2.4 Model Comparison and Evaluation

Table 12 presents the out-of-sample conditional KLIC. To calibrate the out-of-sample conditional KLIC, the time period is divided into the in-sample period t_1, where t_1 =1,…,R and the out-of-sample period t_2, where t_2 = R+1,…,T. In this study, R is set as half period of time T, i.e., R=T/2. The out-of-sample KLIC symmetry is calculated using the parameters (mean and variance of the empirical distribution and dependent parameter of copula) obtained from the in-sample period.

Although the Gaussian copula shows best fit to the data, the Frank copula model attains the lowest conditional KLIC in the out-of-sample period. The estimated Kendall's tau for the empirical data is 0.27 and is, therefore, relatively low. Thus the results of the empirical model are not inconsistent with the simulation results.

Table 12. Out-of-sample Conditional KLIC measures

		VAR
Standard		3.185
Copula-based models	Gauss	3.109
	Clayton	2.992
	Frank	2.947
	Plackett	2.986

4. Summary and Recommendations

The simulation results suggest that the choice of copula and model to be fitted, i.e., whether copula-based VECM or copula-based VAR, depends on the underlying distribution of the model error terms, the strength of the correlation between the error terms as well as focus of the study – i.e., whether estimation of the dependence parameter of the joint distribution is of interest, the fit of the copula-based models to the data, or the forecasting ability of the copula-based VAR or copula-based VECM model. Initially, this general finding may come as a surprise until it is recalled that different criteria are used to compare the competing copulas and models. A case in point is the ability of the multiple regression model to generate good forecasts in the presence of multicollinearity despite ill-behaved parameter estimates.

When the focus of the investigation is estimation of the dependence parameter, the simulations results suggest the use of a copula-based VECM when the correlation is low and the use of a copula-based VAR when the correlation is high irrespective of the underlying error distribution and choice of copula. However, when the correlation is moderate, the results suggest the use of copula-based VAR when the error term is normal, and the use of a copula-based VECM when the error term is student's t except, again, for the Clayton copula.

However, when the interest is in fitting the copula-based models to the data, the choice of copula is apparently more crucial than the choice between copula-based VAR and copula-based VECM models. The Gaussian copula expectedly outperforms the other copulas when the underlying error distribution is normal, particularly when the correlation between the error terms is moderate or high. It also performs satisfactorily when the error distribution is student's t, provided the dependence parameter is high and the copula-based VECM model is used. For low to moderate dependence between error components following a student's t distribution, however, the Plackett and Gumbel copulas outperform the Gaussian copula, with the copula-based VAR tending to outperform its VECM counterpart when the Plackett copula is used.

When forecasting ability is the main concern, copula-based models generally outperform standard models. The simulation results further indicate the superiority of copula-based VAR models over VECM-based models when the correlation in the error terms is high, irrespective of whether the error term follows a normal or a student's t distribution. When the error term is normally-distributed, the Frank copula, and not the Gaussian copula, is the copula of choice. When the error term follows a student's t-distribution, on the other hand, no single copula or model-fitting method is shown to be the best choice across the dependence scenarios. Results indicate the use of Clayton-based VAR, Gaussian-based VECM and the Frank-based VAR for the low, moderate and high correlation scenarios, respectively.

Results obtained from applying copula-based VAR on real consumption expenditure and real gross national income in Japan are consistent with the simulation findings. In particular, the copula-based VAR models display better density forecasting performance than the standard VAR model, with the Frank copula-based VAR providing the best forecasting performance.

Future simulation studies can focus on alternative data generating mechanisms as well as on a finer grid of values for the dependence parameter t (say, in increments of 0.1 from 0.1 to 0.9). Other types of copula such as the two-parameter copula can also be tried under different dependence scenarios such as asymmetrically distributed data.

NOTES

1 Granger et. al. did not employ cointegration analysis, but application of the Johansen test to the data used in the paper (which was taken from St. Louis Federal Reserve web page) indicated the existence of a cointegration relationship.

2 The null hypothesis of the KPSS test is that there is no unit root, while for the ADF and PP tests the null hypothesis is that there is a unit root.

REFERENCES

BAO, Y., LEE, T.H. and BURAK, S., 2007, Comparing Density Forecast Models, *Journal of Forecasting*, 26:203-225.

BIANCHI, C., CARTA, A., FANTAZZINI, D., DE GIULI, M.E. and MAGGI, M.A., 2009, Copula-VAR-X Approach for Industrial Production Modelling and Forecasting, *Applied Economics*, 24.

GRANGER, C.W.J., TERÄSVIRTA, T. and PATTON A.J., 2006, Common Factor in Conditional Distributions for Bivariate Time Series, *Journal of Econometrics*, 132: 43-57.

JOE H., 1997, *Multivariate Models and Dependence Concepts*, Chapman & Hall/CRC.

JONDEAU, E. and ROCKINGER, M., 2006, The Copula-GARCH model of Conditional Dependencies: An International Stock Market Application, *Journal of International Money and Finance*, 25:827-853.

KIM, G., SILVAPULLE, M.J. and SILVAPULLE, P., 2007, Estimating the Error Distribution in Multivariate Heteroscedastic Time Series Models, *Monash Econometrics and Business Statistics*, No 8/07.

LINDSKOG, F., 2000, Linear Correlation Estimation, Working Paper, Zürich.

MITCHELL, J., 2007, Constructing Bivariate Density Forecasts of Inflation and Output Growth Using Copulas: Modelling Dependence Using the Survey of Professional Forecasters, National Institute of Economic and Social Research.

NELSON, R.B., 2006, *An Introduction to Copulas*, Springer-Verlag New York.

PATTON, A.J., 2005, Modelling Time-Varying Exchange Rate Dependence, *Intentional Economic Review*, 47(2):527-556.

SKLAR, A., 1959, *Fonctions de repartition à n dimensions et leurs marges*, publ. Inst. Statist. Univ. Paris,8,229-231.

VUONG, Q.H., 1989, Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. *Econometrica*, 57: 307-334.