Received: 17 August 2021 | Revised: 8 June 2022 | Accepted: 10 August 2022

DOI: 10.1002/for.2902

RESEARCH ARTICLE

WILEY

Structural and predictive analyses with a mixed copula-based vector autoregression model

Woraphon Yamaka[1] | **Rangan Gupta**[2] | **Sukrit Thongkairat**[1] | **Paravee Maneejuk**[1]

[1]Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand

[2]Department of Economics, University of Pretoria, Pretoria, South Africa

Correspondence
Woraphon Yamaka, Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand.
Email: woraphon.econ@gmail.com

Funding information
This research is supported by Fundamental Fund 2022, Chiang Mai University, Thailand.

Abstract

In this study, we introduce a mixed copula-based vector autoregressive (VAR) model for investigating the relationship between random variables. The one-step maximum likelihood estimation is used to obtain point estimates of the autoregressive parameters and mixed copula parameters. More specifically, we combine the likelihoods of the marginal and mixed copula to construct the full likelihood function. The simulation study is used to confirm the accuracy of the estimation as well as the reliability of the proposed model. Various mixed copula forms from a combination of Gaussian, Student's *t*, Clayton, Frank, Gumbel, and Joe copulas are introduced. The proposed model is compared to the traditional VAR model and single copula-based VAR models to assess its performance. Furthermore, the real data study is also conducted to validate our proposed method. As a result, it is found that the one-step maximum likelihood provides accurate and reliable results. Also, we show that if we ignore the complex and nonlinear correlation between the errors, it causes significant efficiency loss in the parameter estimation in terms of $|Bias|$ and MSE. In the application study, the mixed copula-based VAR is the best fitting copula for our application study.

KEYWORDS

forecasting, mixed copula, predictive power, vector autoregressive

1 | INTRODUCTION

Many studies have tried various approaches to understand the economy's working processes to allow policymakers to determine and implement the appropriate economic policy for solving the uncertainty of aggregate economic variability. In the traditional approaches, the relationship between the variables is mostly estimated by the linear correlation assumption. However, the linear correlation cannot adequately describe the true and complicated relationship among the macroeconomic variables. So, the vector autoregression (VAR) model was developed for describing the dynamic behavior of macroeconomic variables more entirely than the linear model can do (Hamilton, 1994; Lutkepohl, 2005; Sims, 1980; Tsay, 2005). The VAR model has an excellent forecasting capability, easy to estimate, and is easy to test for relationships and causality between the variables as it treats the variables to be endogenous variables.

By the properties of the VAR model, it is a system of equations to be estimated simultaneously. As it can combine the information from the different equations by allowing error terms of each equation to be correlated, the model becomes more consistent. However, the problem is that the model's multivariate normal distribution might curse the parameter estimates as the real data might exhibit a non-normal distribution. Thus, the multivariate normal distribution may fail to join the error term of each VAR equation. To relax this strong assumption, a copula approach has been proposed to deal with this complicated dependence (see Hu, 2006; Liu et al., 2020; Maneejuk et al., 2016; Mohti et al., 2019; Pastpipatkul et al., 2016, 2017; Xu & Gao, 2019). These studies confirm the superiority of the copula-based multivariate model over the conventional model. Specifically, copula allows us to construct a joint multivariate distribution of all error terms in the model (Maneejuk et al., 2016). In addition, Joe and Xu (1996) and Fan and Patton (2014) mentioned copula is a good alternative method for modeling the dependence of multivariate data when the multivariate normality of the multivariate data is doubtful. This is because it can capture the nonlinear dependence and tail dependence, and there are no constraints regarding the marginal distributions of random variables. Therefore, the copula approach is introduced to the VAR model to improve efficiency by allowing the model to have different marginal distributions of error terms and not need to be normally distributed.

In addition, our study aims to go beyond those previous copulas-based multivariate models by considering a new class of the copula approach, that is, the mixed copula. Recent studies have shown that this copula class is more flexible than a single copula class (see Maneejuk et al., 2018; Nguyen et al., 2016; Tansuchat & Maneejuk, 2018). The mixed copula is a combination of different copula families, and it can capture both symmetric and asymmetric and other complicated dependence structures. To our knowledge, the copula-based VAR model has already been proposed by Brechmann and Czado (2015) and Yamaka and Thongkairat (2020); however, the estimation and computation aspects of mixed copula-based VAR have never been proposed yet. For this reason, this study attempts to fill the gap of knowledge by applying various combinations of mixed copula to allow better flexibility of capturing almost all possible dependence structures between error terms in the VAR framework. This study also introduces the way to a one-step estimation technique to estimate the mixed Copula-based VAR model. To confirm our model's performance and the accuracy of the estimation, experimental studies and a real application study are introduced.

The contribution of this paper is three-fold. Firstly, it proposes a mixed copula-based VAR model and introduces various combinations of mixed copulas to the VAR model. This is the novelty of our model development, which relaxes the limitations of the conventional VAR model. Secondly, we verify the reliability and accuracy of relaxing the strong assumption of a linear relationship and multivariate normal distribution of the error terms in the VAR model by providing simulation studies. Third, we show the flexibility and validity using the mixed copula-based VAR to investigate a real data relationship.

The paper's remainder proceeds as follows: The methodologies used in this study are present in Section 2, comprising mixed copulas function, VAR model, and estimation. Then, experimental study is provided in Section 3. Section 4 provides the application study. Last but not least, a conclusion is provided in Section 5.

2 | METHODOLOGY

2.1 | VAR model

The VAR model is the multivariate model providing useful information on the relationship between the set of variables. This model is put forward by Sims (1980). In practice, it is generally used to describe the dynamic behavior of time series variables and forecast future values. The basic VAR model with lag p and n dimensional vector of variables measured at time t, $\mathbf{y}_t = (y_{1t}, y_{2t}, ...y_{nt})'$, can be written as

$$\begin{aligned} y_{1t} &= \alpha_1 + \beta_{11} y_{1t-1} + ... + \beta_{1p} y_{nt-p} + \varepsilon_{1t} \\ &\vdots \\ y_{nt} &= \alpha_n + \beta_{n1} y_{nt-1} + ... + \beta_{np} y_{nt-p} + \varepsilon_{nt} \end{aligned}, \quad (1)$$

where $\beta_{ij},\ i = 1, ..., n;\ j = 1, ..., p$ are the autoregressive coefficients, $\boldsymbol{\alpha} = (\alpha_i, ..., \alpha_n)$ is a vector of the intercept term, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{nt})$ is the vector of the noise process with the distribution $N(0, \Sigma)$, whereas Σ is the variance–covariance matrix. As we suggest to use a copula to join the error terms of the VAR model, the use of the copula allows us to model the nonlinear correlation of the standardized errors of VAR.

2.2 | Concept to copula

The copula is introduced in Sklar's theorem (Sklar, 1959). It was proposed to explain a multivariate distribution function of the joint univariate marginal distributions.

The linkage among these marginal distributions can be described by the copula function so that

$$\mathrm{H}(z_1,\ldots,z_n)=C(F_1(z_1),\ldots,F_n(z_n))=C(u_1,\ldots,u_n), \quad (2)$$

where C denotes the copula, $\mathrm{H}(z_1,\ldots,z_n)$ is a joint distribution, and $\mathbf{z}=z_1,\ldots,z_n$ is a $n\times 1$ the realization of the standardized residuals $\left(\varepsilon_i/\sigma_i^2\right)$, where σ_i^2 is the variance of equation i. $u=u_1,\ldots,u_n$ is a $n\times 1$ uniformly distributed marginals. We can write the joint density of the random variables as

$$\begin{aligned} h(z_1,\ldots,z_n)&=\frac{\partial^n C(u_1,\ldots,u_n)}{\partial u_1,\ldots,\partial u_1}\cdot\frac{\partial F_1(z_1)}{\partial z_1}\cdot\ldots\cdot\frac{\partial F_n(z_n)}{\partial z_n} \\ &=c(u_1,\ldots,u_n)\cdot f(z_1)\cdot\ldots\cdot f(z_n). \end{aligned} \quad (3)$$

The three main classes of the copula function are Elliptical Copulas, Archimedean Copulas, and mixed Copula. Many studies in the last decade have intensively conducted the first two classes of the copula. After that, Nelsen (2006) suggested using a convex combination for mixing different copulas, thus allowing for flexibility in the copula dependence structure. The recent development of copula can be found in Fermanian (2017).

2.2.1 | Elliptical copulas

Elliptical copulas are elliptically contoured in their distributions class. The advantage of this copula is that it can explain different correlations between the marginals, where the value of correlation is restricted to be $[-1,1]$. The disadvantages of this class are that they are restricted to have radial symmetry and do not have explicit expressions. There are two parametric copula families in this class, namely, Gaussian and Student's t Copula.

2.2.2 | Archimedean copulas

Archimedean class is another class of the copula function where it can capture various dependence structures, for example, concordance and tail dependence, and have explicit expressions. Comparing with Elliptical copulas, Archimedean copulas are not derived from multivariate distributions using Sklar's theorem. There are many families of the copula in this class, but in this study, we consider only the Clayton copula, Frank copula, Joe copula, and Gumbel copula. Clayton copula only has lower tail dependence, whereas Gumbel copula and Joe copula have only upper tail dependence. For Frank copula, there is no tail dependence. The advantage of these copulas is that they take into account the asymmetric dependence structure. For further detail, we refer to Nelsen (2006) and Hofert et al. (2012).

2.2.3 | Mixed copulas

The Elliptical and Archimedean copulas may provide an unreliable dependency measure, as some copula families' dependence parameter is restricted in specific ranges. Nelsen (2006) introduced an idea to mix the copula function through the convex combination method to improve the ability to capture a wide range of complicated dependence structures. Several additional advantages are obtained when the mixed copula is used to measure the dependence of the random variables. First, as the mixed copula is constructed from various copula functions, it becomes more flexible to join any form of the dependence structure. Second, the dependence structures captured by mixed copulas are not changed, even though the data are transformed into several types. The mixed copula can be derived by

$$c_{Mix}(u_1,\ldots,u_n|\theta_1,\theta_2)=wc_{\theta_1}(u_1,\ldots,u_n|\theta_1)+(1-w)c_{\theta_2}(u_1,\ldots,u_n|\theta_2), \quad (4)$$

where w is the weight parameter with the value ranges between zero and one. θ_1 and θ_2 are an exchangeable copula parameter. One of the advantages of the convex combination approach is the flexibility in assigning weights for calculating appropriate value between two copula functions. There are various copula functions proposed to join the marginal distributions, and the selection of copula type is essential. In this study, we consider five types of copulas to capture different patterns of dependence between the random variables. Copula functions commonly used in the research are Gaussian (no tail dependence), Student's t (symmetric tail dependence), Gumbel (right tail dependence), Clayton (left tail dependence), and Frank (no tail dependence).

2.3 | Estimation of the mixed copula model

In general, the copula-based model can be estimated by the two-step estimation called Inference Function for Margins (IFM) of Joe and Xu (1996). In the first step, the marginal distribution is estimated, and, in the second step, the copula parameter is estimated, given the estimated parameters from the first step. Joe and Xu (1996) suggested several advantages of this method. First, it is easy to estimate the multivariate models; second, the multivariate model is still robust even though a misspecification copula is used to model the marginal distributions' dependence. Also, it is more robust against outliers or perturbations of the data, and third, the IFM method can deal with the large parameter estimates in the multivariate models as it estimates the marginal parameters and copula parameters separately. Therefore, this method could reduce the cost of computation in the multivariate model. Although there are several advantages of this IFM method, Louzada and Ferreira (2016) revealed that the IFM method still produces a biased estimate. Thus, we consider using a one-step maximum likelihood estimator to obtain our proposed model's parameter estimates. We note that the maximum likelihood estimation is a flexible estimation for estimating the point parameter estimates. As the VAR model is the multivariate equation model; thus, a multivariate copula with continuous marginal distribution is required in the estimation. Let $\Theta=\{\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_n\}$, $\boldsymbol{\theta}=\{\theta_1,\theta_2\}$, and similar to Equation (3), the joint density function of mixed copula-based VAR model is

$$h(z_1,...,z_n;\Theta,\boldsymbol{\theta},w)= f_1(z_1;\boldsymbol{\beta}_1,\sigma_1^2),...,f_n(z_n;\boldsymbol{\beta}_n,\sigma_n^2)c(u_1,...,u_n;\theta_1,\theta_2,w), \tag{5}$$

where $\boldsymbol{\beta}_i$ is the vector of parameter in equation i, $i=1,...,n$, and $f_i(z_i;\boldsymbol{\beta}_i,\sigma_i^2)$ is the density function of equation i, which is assumed to have a normal distribution. $c(u_1,...,u_n;\theta_1,\theta_2,w)$ is the mixed copula density. Note that we are following copula families, that is, Gaussian, Student's t, Clayton, Gumbel, Frank, Joe, and mixed copula families such as Gaussian–Student's t, Gaussian–Clayton, Gaussian–Gumbel, Student's t–Clayton, Student's t, and Frank. Hence, log-likelihood analysis implies

$$\ln L(\Phi)=\ln f_1(z_1;\boldsymbol{\beta}_1,\sigma_1^2)+...+\ln f_n(z_n;\boldsymbol{\beta}_n,\sigma_n^2) + \ln c(u_1,...,u_n;\theta_1,\theta_2,w) \tag{6}$$

where Φ is all the parameter estimates. Then, the log-likelihood Equation (6) is maximized to estimate the marginal distribution parameters and copula parameters. In other words, the maximum likelihood estimator maximizes the log-likelihood and is given by

$$\widehat{\Phi}=\underset{\Phi}{argmax}\sum_{t=1}^{T}\ln L(\Phi), \tag{7}$$

where $\widehat{\Phi}$ is the optimally estimated parameters. We set the score function to zero (this function is defined as the first-order partial derivative of the full likelihood function's logarithm Φ). However, it is not always possible to find closed-form expressions for these estimators. Therefore, the use of iterative methods is often needed. In this estimation, we employ the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.

3 | SIMULATION STUDY

To evaluate the accuracy and reliability of the mixed copula-based VAR model, the Monte Carlo simulation study is employed. This helps us to compare the performance of our proposed model and other conventional competitive models. We consider two experiments for checking the accuracy and performance of our proposed model. In the first experiment, the accuracy of the maximum likelihood estimation for mixed copula-based VAR models is investigated. Second, our model's performance is compared with two conventional models, namely, Gaussian copula-based VAR and the Independent copula-based VAR model. We note that the Independent copula-based VAR model ignores the correlation among VAR equations. Specifically, it is assumed that the error terms of VAR models are independent. Thus, the copula density is equal to the product of the marginal distribution of the standardized residuals. That is, independent copula density is equal to one. We note that the estimation of independent copula-based VAR is the separate estimation, where each equation is separately estimated using maximum likelihood. Specifically, the error terms of this model are jointly independent.

3.1 | Investigating the accuracy of the one-step maximum likelihood estimation

This experiment's simulated data are generated from the simple bivariate and trivariate VAR model with lag one. These two models take the form of the following:

Model 1: The bivariate VAR(1) model

$$\begin{aligned} y_{1t}&=2+0.5y_{1t-1}+0.2y_{2t-1}+\varepsilon_{1t} \\ y_{2t}&=4+0.3y_{1t-1}+0.4y_{2t-1}+\varepsilon_{2t} \end{aligned}. \tag{8}$$

Model 2: The trivariate VAR(1) model

$$\begin{aligned} y_{1t}&=2+0.3y_{1t-1}+0.2y_{2t-1}+0.2y_{3t-1}+\varepsilon_{1t} \\ y_{2t}&=4+0.3y_{1t-1}+0.5y_{2t-1}+0.3y_{3t-1}+\varepsilon_{2t} \\ y_{3t}&=1+0.4y_{1t-1}+0.3y_{2t-1}+0.3y_{3t-1}+\varepsilon_{3t} \end{aligned}. \tag{9}$$

TABLE 1 Simulation results of the mixed copula-based VAR model

Model 1	**Gaussian–Student's *t***				**Clayton–Gumbel**				**Frank–Joe**			
Para	**True**	**$t = 200$**	**$t = 500$**	**$t = 1000$**	**True**	**$t = 200$**	**$t = 500$**	**$t = 1000$**	**True**	**$t = 200$**	**$t = 500$**	**$t = 1000$**
α_1	2	1.983 (0.010)	2.013 (0.013)	2.003 (0.002)	2	1.908 (0.104)	2.022 (0.021)	2.001 (0.001)	2	1.943 (0.050)	2.013 (0.013)	2.003 (0.002)
β_{11}	0.5	0.457 (0.023)	0.521 (0.020)	0.501 (0.001)	0.5	0.443 (0.053)	0.534 (0.030)	0.510 (0.010)	0.5	0.483 (0.030)	0.521 (0.020)	0.501 (0.001)
β_{12}	0.2	0.237 (0.042)	0.230 (0.021)	0.224 (0.020)	0.2	0.270 (0.041)	0.250 (0.051)	0.228 (0.017)	0.2	0.213 (0.022)	0.250 (0.051)	0.224 (0.020)
α_2	4	3.993 (0.009)	4.035 (0.042)	4.010 (0.010)	4	3.894 (0.020)	4.021 (0.032)	4.019 (0.024)	4	3.876 (0.302)	4.035 (0.042)	4.010 (0.010)
β_{21}	0.3	0.372 (0.067)	0.272 (0.021)	0.305 (0.003)	0.3	0.359 (0.040)	0.282 (0.012)	0.303 (0.003)	0.3	0.278 (0.038)	0.272 (0.021)	0.305 (0.003)
β_{22}	0.4	0.419 (0.077)	0.409 (0.010)	0.399 (0.002)	0.4	0.423 (0.015)	0.421 (0.033)	0.409 (0.004)	0.4	0.387 (0.024)	0.409 (0.010)	0.399 (0.002)
θ_1	0.5	0.594 (0.123)	0.524 (0.023)	0.501 (0.001)	3	3.320 (0.453)	3.112 (0.213)	3.012 (0.018)	3	3.183 (0.100)	3.059 (0.031)	3.032 (0.022)
θ_2	0.5	0.471 (0.020)	0.489 (0.012)	0.506 (0.004)	3	3.209 (0.220)	3.147 (0.100)	3.069 (0.025)	3	3.397 (0.462)	3.244 (0.142)	3.050 (0.013)
w	0.5	0.390 (0.120)	0.456 (0.063)	0.510 (0.010)	0.5	0.413 (0.194)	0.456 (0.044)	0.502 (0.006)	0.5	0.432 (0.130)	0.478 (0.102)	0.512 (0.021)
Model 2	**Gaussian–Student's *t***				**Clayton–Gumbel**				**Frank–Joe**			
Para	True	$t = 200$	$t = 500$	$t = 1000$	True	$t = 200$	$t = 500$	$t = 1000$	True	$t = 200$	$t = 500$	$t = 1000$
α_1	2	2.219 (0.382)	2.109 (0.113)	2.002 (0.004)	2	2.203 (0.221)	2.123 (0.109)	2.010 (0.021)	2	2.392 (0.345)	2.230 (0.211)	2.010 (0.011)
β_{11}	0.3	0.252 (0.066)	0.286 (0.052)	0.322 (0.024)	0.3	0.289 (0.023)	0.293 (0.010)	0.303 (0.004)	0.3	0.342 (0.044)	0.290 (0.015)	0.306 (0.011)
β_{12}	0.2	0.177 (0.043)	0.195 (0.028)	0.206 (0.002)	0.2	0.221 (0.032)	0.210 (0.020)	0.203 (0.005)	0.2	0.223 (0.034)	0.213 (0.015)	0.210 (0.009)
β_{13}	0.2	0.213 (0.011)	0.209 (0.009)	0.199 (0.001)	0.2	0.185 (0.026)	0.208 (0.007)	0.199 (0.001)	0.2	0.232 (0.035)	0.221 (0.018)	0.203 (0.005)
α_2	4	4.521 (0.892)	4.346 (0.509)	4.007 (0.005)	4	4.392 (0.286)	4.201 (0.182)	4.004 (0.005)	4	4.193 (0.290)	4.086 (0.056)	4.028 (0.021)
β_{21}	0.3	0.413 (0.108)	0.329 (0.027)	0.305 (0.010)	0.3	0.324 (0.030)	0.317 (0.022)	0.310 (0.011)	0.3	0.389 (0.100)	0.352 (0.047)	0.293 (0.005)
β_{22}	0.5	0.432 (0.056)	0.472 (0.032)	0.501 (0.001)	0.5	0.422 (0.070)	0.478 (0.029)	0.502 (0.002)	0.5	0.446 (0.058)	0.488 (0.029)	0.512 (0.012)

(Continues)

TABLE 1 (Continued)

Model 2	Gaussian–Student's *t*				Clayton–Gumbel				Frank–Joe			
β_{23}	0.3	0.332 (0.049)	0.279 (0.021)	0.302 (0.004)	0.3	0.283 (0.019)	0.311 (0.012)	0.304 (0.010)	0.3	0.341 (0.056)	0.320 (0.019)	0.301 (0.001)
α_2	1	1.192 (0.136)	1.124 (0.132)	1.005 (0.007)	1	1.210 (0.311)	1.138 (0.237)	1.004 (0.006)	1	1.142 (0.102)	1.098 (0.077)	1.011 (0.013)
β_{31}	0.4	0.374 (0.029)	0.422 (0.028)	0.403 (0.009)	0.4	0.362 (0.033)	0.428 (0.025)	0.401 (0.001)	0.4	0.432 (0.038)	0.429 (0.025)	0.411 (0.014)
β_{32}	0.4	0.421 (0.022)	0.403 (0.013)	0.401 (0.001)	0.4	0.421 (0.022)	0.412 (0.010)	0.402 (0.003)	0.4	0.432 (0.043)	0.416 (0.015)	0.404 (0.005)
β_{33}	0.3	0.353 (0.671)	0.389 (0.021)	0.309 (0.010)	0.3	0.325 (0.018)	0.318 (0.010)	0.313 (0.007)	0.3	0.324 (0.043)	0.321 (0.011)	0.306 (0.005)
θ_1	0.5	0.439 (0.024)	0.486 (0.012)	0.505 (0.010)	3	2.762 (0.324)	2.833 (0.251)	3.042 (0.093)	3	3.453 (0.652)	3.189 (0.188)	3.072 (0.023)
θ_2	0.5	0.673 (0.201)	0.489 (0.014)	0.499 (0.001)	3	2.792 (0.391)	3.152 (0.117)	3.131 (0.122)	3	3.234 (0.211)	3.283 (0.167)	3.106 (0.110)
w	0,5	0.568 (0.048)	0.529 (0.021)	0.510 (0.019)	0.5	0.389 (0.110)	0.429 (0.072)	0.503 (0.005)	0.5	0.409 (0.103)	0.456 (0.443)	0.514 (0.014)

Note: The parentheses () denote the standard deviation of the parameter estimate.
Abbreviation: VAR, vector autoregression.

In each case under each model, 100 samples, each consisting of $t = 200$, 500, and 1000 data points, are generated using Equations (8) and (9). In this simulation study, three mixed copula functions, namely, Gaussian–Student's t, Clayton–Gumbel, and Frank–Joe, are used as an example. The data-generating mechanism consisted of the following steps:

1. We simulate the marginal u_i for equation i form the mixed copula model. The true parameters of the mixed copula are fixed as $w = 0.5$, $\theta_1 = \theta_2 = 0.5$ for Gaussian–Student's t copula (corresponding to a Kendall's tau association measure of 0.3), $\theta_1 = \theta_2 = 3$ for Clayton–Gumbel copula (corresponding to a Kendall's tau association measures of 0.600 and 0.666), and Frank–Joe copula (corresponding to a Kendall's tau association measures of 0.306 and 0.517).
2. The marginal u_i is then transformed to be error term for equation i, ε_{it}, using the inverse normal distribution with $N(0,1)$.
3. Finally, the y_{it} is generated from the VAR specification in Equations (8) and (9).

The mean and standard deviation of the parameter estimates from 100 datasets are reported in Table 1. The simulation results from Table 1 show that our proposed model and maximum likelihood estimation produce reliable parameter estimates. The average of the parameter estimates is close to the true values with reasonable standard deviations. It is also observed that maximum likelihood estimation's performance becomes better when the sample size increases from t going from 200 to 1000, indicating this estimator is unbiased and confirming the validity of asymptotic properties.

3.2 | Investigating the performance of mixed copula-based VAR model

In the second simulation study, we compare our mixed copula-based VAR model's performance with the two competing models, namely, Gaussian copula-based VAR and Independent copula-based VAR. We consider two criteria, absolute *Bias* and mean squared error (*MSE*). The $|Bias$ and *MSE* can be calculated by

$$|Bias| = \left| R^{-1} \sum_{r=1}^{R} \left(\widetilde{\Phi}_r - \Phi_r \right) \right|, \tag{10}$$

$$MSE = R^{-1} \sum_{r=1}^{R} \left(\widetilde{\Phi}_r - \Phi_r \right)^2, \tag{11}$$

where $R = 100$ is the number of Monte Carlo replications and $\widetilde{\Phi}_r$ and Φ_r are the estimated values and the true values, respectively. The sample sizes are fixed at 200 for each replication.

The data sets are generated in the same way as in the first experiment. However, we only consider the bivariate VAR specification in this second experiment (Equation 8). To make a fair simulation experiment, the data are generated from the three mixed copula-based VAR (Gaussian–Student's t, Clayton–Gumbel, and Frank–Joe) and Gaussain copula-based models. The purpose of this experiment is to investigate the performance of the mixed copula-based VAR model when the copula is correctly specified and when the copula is misspecified.

Because the mixed copula is proposed to improve the estimated parameters' accuracy in the VAR model, we thus focus here on the VAR parameter estimates. Table 2 presents the $|Bias|$ and *MSE* of each model measured assuming the dependence structure of the error terms follow Gaussian–Student's t, Clayton–Gumbel, and Frank–Joe mixed copulas. As we expected, the overall absolute *Bias* and *MSE* for autoregressive parameters from the mixed copula-based VAR are lower than the overall absolute *Bias* and *MSE* from the two conventional VAR: Gaussian copula-based VAR and Independent copula-based VAR models. In particular, the $|Bias|$ and *MSE* of autoregressive parameters from mixed Clayton–Gumbel copula-based VAR and mixed Frank–Joe copula-based VAR are obviously lower than the conventional VAR models. However, we observe interesting evidence that the $|Bias|$ and *MSE* of autoregressive parameters from Gaussian copula-based VAR seem to perform well in some cases. This result is consistent with Christopoulos et al. (2021) suggestion that Gaussian copula may capture the structure of nonlinear and complicated dependence of the marginals.

Next, we investigate the performance of our mixed copula-based VAR model when the data are generated from the wrong copula function. Therefore, we simulate the data from the Gaussian copula-based VAR. The performance of our mixed copula-based VAR is reported in Table 3. As expected, our Gaussian copula-based VAR model performs the best as the lowest $|Bias|$ and *MSE* are revealed. We also find that the misspecified copula functions bring a larger deviation of the approximated our mixed copula-based VAR models as the $|Bias|$ and *MSE* are large, especially mixed Clayton–Gumbel and mixed Frank–Joe Copulas. According to this experiment result, we can conclude that our proposed model is the robust model, and the incorrectly specified copula function will lead to the low accuracy of the model. Our finding is consistent with Kaewsompong et al. (2020).

TABLE 2 Comparing the absolute |*Bias*| and *MSE* of our proposed models and other two conventional when the data are generated from the mixed copula based model

Model 1 Para	Mixed Gaussian–Student's *t* copula-based VAR		Gaussian copula-based VAR		Independent copula-based VAR	
	|*Bias*|	*MSE*	|*Bias*|	*MSE*	|*Bias*|	*MSE*
α_1	0.015	0.001	0.637	0.404	1.024	1.047
β_{11}	0.040	0.002	0.088	0.007	0.129	0.016
β_{12}	0.042	0.002	0.113	0.013	0.332	0.110
α_2	0.067	0.004	0.514	0.263	1.829	3.349
β_{21}	0.071	0.005	0.233	0.054	0.316	0.099
β_{22}	0.021	0.001	0.142	0.020	0.221	0.048
Model 1 Para	**Mixed Clayton–Gumbel copula-based VAR**		**Gaussian copula-based VAR**		**Independent copula-based VAR**	
	|*Bias*|	*MSE*	|*Bias*|	*MSE*	|*Bias*|	*MSE*
α_1	0.129	0.016	0.934	0.874	1.239	1.538
β_{11}	0.062	0.004	0.129	0.016	0.242	0.057
β_{12}	0.039	0.002	0.102	0.010	0.211	0.044
α_2	0.024	0.002	0.911	0.830	2.034	3.137
β_{21}	0.051	0.003	0.102	0.010	0.298	0.088
β_{22}	0.014	0.001	0.234	0.054	0.239	0.056
Model 1 Para	**Mixed Frank–Joe copula-based VAR**		**Gaussian copula-based VAR**		**Independent copula-based VAR**	
	|*Bias*|	*MSE*	|*Bias*|	*MSE*	|*Bias*|	*MSE*
α_1	0.056	0.003	0.692	0.479	1.015	0.925
β_{11}	0.031	0.002	0.122	0.014	0.157	0.024
β_{12}	0.021	0.001	0.382	0.146	0.309	0.094
α_2	0.357	0.127	0.801	0.640	1.394	1.938
β_{21}	0.051	0.003	0.122	0.014	0.209	0.043
β_{22}	0.025	0.001	0.192	0.036	0.224	0.050

Abbreviation: VAR, vector autoregression.

TABLE 3 Comparing the absolute |*Bias*| and *MSE* of our proposed models and other two conventional models when the data are generated from the Gaussian copula-based model

Model 1 Para	Mixed Gaussian–Student's *t* copula-based VAR		Mixed Clayton–Gumbel copula-based VAR		Mixed Frank–Joe copula based VAR		Gaussian copula-based VAR		Independent copula-based VAR	
	|*Bias*|	*MSE*	|*Bias*|	*MSE*	|*Bias*|	*MSE*	|*Bias*|	*MSE*	|*Bias*|	*MSE*
α_1	0.013	0.001	0.789	0.619	0.350	0.122	0.010	0.001	0.583	0.341
β_{11}	0.040	0.002	0.214	0.045	0.243	0.058	0.042	0.002	0.087	0.007
β_{12}	0.045	0.002	0.139	0.019	0.203	0.041	0.039	0.002	0.091	0.008
α_2	0.087	0.009	0.897	0.803	0.923	0.853	0.071	0.005	0.910	0.813
β_{21}	0.098	0.010	0.237	0.057	0.179	0.003	0.049	0.003	0.109	0.012
β_{22}	0.050	0.003	0.284	0.080	0.209	0.043	0.033	0.001	0.232	0.053

Abbreviation: VAR, vector autoregression.

In sum, the independent copula-based VAR (Separate Estimation) and Gaussian copula-based VAR model do not quite perform satisfactorily when the dependence structure or the error terms is mixed. However, its performance is further improved if the mixed copula-based VAR model is fitted. This indicates that when the dependence structure of the error terms has deviated from the multivariate elliptical distribution or independence, the elliptical copula-based VAR and the independent copula-based VAR models face higher Bias and variance. Thus, the slight gain in accuracy and precision in the parameter estimate can be obtained by the better fit afforded by a mixed copula. In addition, our simulation results also provide the accuracy improvement of one-step maximum likelihood estimation over the IFM method (two-step estimation) as shown with the lower $|Bias|$ and MSE of mixed copula-based VAR compared with the independent copula-based VAR. We would like to note that the results of the independent copula-based VAR model are obtained from the first step of IFM.

4 | THE EMPIRICAL APPLICATIONS

This study applies the proposed method to real-time-series data sets. Given the large literature on the impact of uncertainty on output that has emerged following the "Great Recession" (see Al-Thaqeb & Algharabali, 2019; Castelnuovo et al., 2017, Gupta et al., 2018, 2019, 2020; for detailed reviews of this literature), we investigate the relationship among a world uncertainty index (WUI), with outputs of the United States (US), other advanced economies (ADV), and emerging countries (EMs). The quarterly data on these variables cover 1990:1 through 2019:4. The real GDP data, capturing output, are obtained from the Global Economic Database maintained by the Federal Reserve Bank of Dallas, which is available for download from https://www.dallasfed.org/institute/dgei/gdp.aspx. The reader is referred to Grossman et al. (2014) for further details. Data on 18 advanced (excluding the United States, Japan, Germany, the United Kingdom [UK], France, Italy, Spain, Canada, South Korea, Australia, Taiwan, The Netherlands, Belgium, Sweden, Austria, Switzerland, Greece, Portugal, and the Czech Republic, in order of Purchasing Power Parity [PPP]-adjusted GDP shares in 2005) and 21 emerging (China, India, Russia, Brazil, Mexico, Turkey, Indonesia, Poland, Thailand, Argentina, South Africa, Colombia, Malaysia, Venezuela, Philippines, Nigeria, Chile, Peru, Hungary, Bulgaria, and Costa Rica, in order of PPP-adjusted GDP shares in 2005) countries are used to compile the aggregates for these blocs, by using trade weights with the United States in weighting the country-level data. At the same time, the WUI is based on the work of Ahir et al. (2018). These authors construct quarterly indices of economic uncertainty for 143 countries from 1990 onwards using frequency counts of "uncertainty" (and its variants) in the quarterly Economist Intelligence Unit (EIU) country reports. The EIU reports discuss major political and economic developments in each country and analyze and forecast political, policy, and economic conditions. To make the WUI comparable across countries, the raw counts are scaled by each report's total number of words. Globally, the WUI spikes near the 9/11 attacks, the SARS outbreak, Gulf War II, the Euro debt crisis, El Niño, Europe border-control crisis, the UK referendum vote in favor of Brexit, and the 2016 US presidential election. In general, the index is associated with greater economic policy uncertainty (EPU), stock market volatility, risk, and lower GDP growth. The data can be downloaded from https://worlduncertaintyindex.com/data/.

4.1 | Empirical model

We transform ADV, EM, and US into a growth rate and WUI to be logarithm for attaining the stationary property. In practice, we need to determine or find the VAR model's optimal lag length; thus, this issue is taken into account in this application. The basic statistics of the transformed data are provided in Table 4. Among the growth rate series, we observe the average of EM performs the highest values and followed by US and ADV, respectively. The variables EM, US and ADV show non-normality as there exhibit negative skewness and high kurtosis values (>3), whereas WUI shows a weak skewness and low kurtosis values (less than 3). Therefore, we confirm this empirical evidence by using the normality Jarque–Bera (J-B) test. According to the J-B test result in Table 3, the result provides decisive evidence of non-normal distribution for all variables, except WUI. Furthermore, as we consider the time-series data, the stationary property is tested for all series. The Augmented Dickey–Fuller (ADF) test is used here, and the result shows significant evidence of the stationary for our series.

4.2 | Model selection

We introduce many copula families for both single and mixed copula-based VAR models; we need to find the best copula-based model to investigate these four variables' relationships. To do this, we consider the Bayesain information criteria (BIC) value as a tool for selecting the best model in this study. Note that this criterion provides unbiased model selection (Shumway & Stoffer, 2011). The model selection result is provided in Table 5. We find that a lag order of $p = 3$ is suggested for most VAR models. When we compare our proposed model's performance, there is evidence that the mixed Clayton–Joe

TABLE 4 Data description

Growth	ADV	EM	US	WUI
Mean	2.040	4.520	2.370	9.640
Median	2.248	4.800	2.570	9.560
Maximum	−12.003	−9.680	−9.030	8.630
Minimum	5.050	8.240	5.300	10.900
SD	1.950	2.330	1.960	0.490
Skewness	−3.851	−2.320	−2.350	0.240
Kurtosis	23.010	10.900	9.940	−0.280
Jarque–Bera	2757.776***	657.424***	567.690***	1.655*
ADF test	−2.893**	−3.025***	−2.997***	−6.904***

***Represents significance at the 1% level.
**Represents significance at the 5% level.
*Represents significance at the 10% level.

TABLE 5 Model selection

Model	Lag 1 BIC	Lag 2 BIC	Lag 3 BIC
Independent copula-based VAR	1048.348	1021.184	1055.202
Copula-based VAR			
Gaussian	1046.880	1011.624	1055.080
Student's *t*	1146.777	1150.697	1043.499
Gumbel	1222.085	1230.831	1077.457
Clayton	1103.821	1312.822	1018.175
Frank	1239.165	1237.442	1085.567
Joe	1219.107	1253.264	1089.245
Mixed copula-based VAR			
Gaussian–Student's *t*	1172.291	1241.679	1083.819
Gaussian–Clayton	1249.970	1490.354	1095.032
Gaussian–Frank	1233.371	1263.423	1112.558
Gaussian–Gumbel	1233.371	1476.242	1103.746
Gaussian–Joe	1197.421	1267.255	1096.779
Student's *t*–Clayton	1157.589	1310.263	1081.100
Student's *t*–Frank	1182.612	1239.030	1084.034
Student's *t*–Gumbel	1231.998	1238.152	1104.983
Student's *t*–Joe	1168.321	1210.010	1084.047
Clayton–Frank	1181.820	1324.939	1095.836
Clayton–Gumbel	1236.827	1247.129	1088.147
Clayton–Joe	1170.118	1257.368	**914.972**
Frank–Gumbel	1253.000	1246.949	1104.409
Frank–Joe	1212.855	1255.018	1101.157
Gumbel–Joe	1226.350	1253.551	1107.574

Note: Bold number indicated the lowest BIC.
Abbreviation: VAR, vector autoregression.

TABLE 6 Parameter estimates from mixed Clayton–Joe copula-based VAR model

	ADV_t	EM_t	US_t	WUI_t
α	3.7481*** (0.1821)	5.6732** (1.2432)	4.0490** (1.0241)	1.6386* (0.0923)
ADV_{t-1}	1.6457*** (0.3234)	0.6038*** (0.1029)	0.6037*** (0.2001)	−0.0171 (0.0209)
EM_{t-1}	0.3273 (0.4222)	1.4987*** (0.2039)	0.5126*** (0.1082)	−0.0832 (0.1032)
US_{t-1}	0.2882** (0.1233)	0.1954 (0.1232)	0.9752*** (0.2083)	0.0261 (0.3743)
WUI_{t-1}	−0.3152 (0.2863)	−0.3302 (0.2083)	−0.4543 (0.3982)	0.4529*** (0.1023)
ADV_{t-2}	−0.9988*** (0.1093)	−1.0229*** (0.1982)	−0.6599* (0.3593)	0.0217 (0.0234)
EM_{t-2}	−0.4013*** (0.1002)	−0.6312*** (0.2110)	−0.6375*** (0.1093)	−0.1089* (0.0511)
US_{t-2}	−0.4485** (0.2192)	−0.3001 (0.4083)	−0.2887 (0.1902)	−0.0355 (0.0233)
WUI_{t-2}	0.0327 (0.0823)	−0.2421 (0.1780)	−0.0432 (0.0322)	0.1625 (0.0800)
ADV_{t-3}	0.1950 (0.2729)	0.4655 (0.3782)	0.2347 (0.2330)	−0.0197 (0.0281)
EM_{t-3}	0.1361 (0.1023)	0.0717 (0.1002)	0.1784** (0.0523)	0.0489 (0.0663)
US_{t-3}	0.2169 (0.1672)	0.0351 (0.2823)	−0.0487 (0.0462)	0.0040 (0.0072)
WUI_{t-3}	−0.0778* (0.0389)	0.0401*** (0.0150)	0.1298*** (0.0491)	0.2206* (0.1128)
	Copula parameter	Kendall tau	Tail dependence	
θ_C	4.7323*** (1.082)	0.6623	Lower = 0.8637 Upper = 0.0000	
θ_J	4.9295*** (1.938)	0.6734	Lower = 0.0000 Upper = 0.8490	
w	0.1697*** (0.0930)			

Note: The parentheses () denote the standard error.
Abbreviation: VAR, vector autoregression.
***Represents significance at the 1% level.
**Represents significance at the 5% level.
*Represents significance at the 10% level.

copula-based VAR model is the best fit model in this application as it has the smallest value of BIC (914.972).

4.3 | Estimation results

The mixed copula dependence parameters are provided in the lower panel of Table 6. We know that Clayton and Joe copula is the best combination of copula families in this study. This indicates that the relationship among the WUI, the advanced stock market, emerging stock market, and the US gross domestic product is asymmetric. The weight parameter (w) is 0.1697, which indicates a mixing copula family in this model, and the weight seems to deviate to the Clayton copula rather than the Joe copula. We also observe that the copula dependence of Clayton and Joe copula are 4.7323 and 4.9295, indicating that residuals of these four variables exhibit a positive dependence. To measure the strength of the dependency of the error distributions, Kendall's taus and tail dependence coefficients for Clayton and Joe copulas are computed, and we found that Kendall's taus for Clayton and Joe copulas are 0.6623 and 0.6734, respectively, whereas the upper tail dependence of Joe copula is 0.8490, and the lower tail dependence of Clayton copula is 0.8637. These dependence measures indicate a high correlation among the error distributions in both normal and extreme events. The impulse response function (IRF) is an essential implication of the VAR model, however, due to the page and word limitations, we report the IRF in the online appendix.

4.4 | Assessing forecast performance: In- and out-of-sample forecasts

We also evaluated the performance of our mixed copula model by considering the forecasting performance. Therefore, we partitioned the data sample into in-sample and out-of-sample periods as 1990Q1-2014Q4 and 2015Q1-2020Q4, respectively. In addition, we also investigate whether WUI enhances the forecasting power in ADV, EM, and US. We then compare a VAR model that includes WUI with the VAR model without WUI.

To evaluate our in-sample and out-of-sample forecasting performance, we consider two loss functions, namely, root mean squares error (*RMSE*), $\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\mathbf{y}_t-\hat{\mathbf{y}}_t\right)^2}$, and mean absolute error (MAE), $\mathrm{MAE}=\frac{1}{n}\sum_{t=1}^{n}\left|\mathbf{y}_t-\hat{\mathbf{y}}_t\right|$, where n is the number of forecasting data (in and out of sample forecasts) and $\mathbf{y}_t=\{\mathrm{US}_t,\mathrm{ADV}_t,\mathrm{EM}_t\}$ and $\hat{\mathbf{y}}_t=\left\{\widehat{\mathrm{US}}_t,\widehat{\mathrm{ADV}}_t,\widehat{\mathrm{EM}}_t\right\}$ are the true observations and predicted values, respectively. However, there is sometimes no consensus for selection based on the loss; thus, it is not easy to find the best forecasting model that minimizes all loss functions. Ma et al. (2019) suggested that the comparison of the performance forecasting models in terms of a loss function cannot distinguish the predictive power among competing models in a statistically significant manner. Thus, we also conduct the model confidence set (MCS) test of Hansen et al. (2011) to analyze the robustness in and out-of-forecasting result.

4.4.1 | In-sample statistical performance

In this first experiment, we do in-sample forecasts to compare the overall performance of the 22 forecasting VAR(3) models with WUI and another 22 forecasting VAR(3) models without WUI. The in-sample statistical performance results of RMSE and MAE and their corresponding MCS *p*-values are provided in Table 7.

Table 7 shows the RMSE and MAE as well *p*-value of MCS test results by using bootstrap simulation at 1000 times. We note that models with lower RMSE and MAE and greater *p*-value imply that they are more accurate in prediction performance in all three forecasting horizons. We can see that Clayton–Joe copula-based VAR with WUI data is statistically superior to all the other models. The MAE and RMSE of the models are the lowest, whereas their MCS's *p*-values are equal to one. This analysis demonstrates that the Clayton–Joe copula-based VAR model has outstanding performance compared to other models for this in-sample forecast. We also find that most VAR models combined with WUI have better performance than the models without WUI, confirming that the EPU index has a good forecasting power in ADV, EM, and US. Hence, it can be concluded that if the WUI data are considered as the predictor variable in the VAR model, better predictions of the growths of the advanced stock market, emerging stock market, and the US GDP are obtained.

4.4.2 | Out-of-sample forecast

In this subsection, we consider only 3 forecasting VAR (3) models with WUI (Gaussian, Clayton, and Clayton–Jo) and these three models without WUI, because out-of-sample forecasting is considered to be the most relevant test for an econometric framework and predictors, rather than in-sample analyses (Campbell, 2008). We choose Gaussian as this model is equivalent to the classical VAR model; Clayton is the best fit single copula for joining the errors in VAR(3), and Clayton–Jo performs the best in the model in the in-sample forecasts.

Then, the parameters of these six models are optimized on a training set; the testing set is used to compare the quality of the models. The former is used to build the forecasting models, and the latter is used to evaluate the accuracy of the various models. Again, the forecasting quality of the forecasting models is measured using the MAE, RMSE, and MCS tests based on MAE and RMSE. The data set is divided into two parts for the use in training (1990Q1-2014Q4) and forecasting (2015Q1-2020Q4). The first part is used to build the forecasting models, whereas the second is intended to evaluate the various models' accuracy. The models are estimated recursively to include observations from the out-of-sample period. In the predictive comparison, we consider forecast horizons of horizons of 1–6 quarter-ahead forecasts. Particularly, the forecast up to 6 quarters ahead is computed, one more observation is added to the sample, and forecasts up to 6 quarters ahead are again generated, and so on.

According to the results of the out-of-sample forecasts in Table 8, we find that both RMSE and MAE as well as

TABLE 7 In sample forecasting performance evaluation based on MAE and RMSE

Model VAR lag 3	Model with WUI MAE	Model with WUI RMSE	Model without WUI MAE	Model without WUI RMSE
Independent	0.5862 (0.2786)	0.8231 (0.2821)	0.5903 (0.0000)	0.8537 (0.0000)
Gaussian	0.5708 (0.4561)	0.8034 (0.4392)	0.5951 (0.0000)	0.8631 (0.0000)
Student's *t*	0.5708 (0.4561)	0.8034 (0.4392)	0.5936 (0.0000)	0.8560 (0.0000)
Gumbel	0.6336 (0.0000)	0.8771 (0.0000)	0.6021 (0.0000)	0.8636 (0.0000)
Clayton	0.5619 (0.4734)	0.7947 (0.4880)	0.6013 (0.0000)	0.8621 (0.0000)
Frank	0.6365 (0.0000)	0.8601 (0.0000)	0.6027 (0.0000)	0.8637 (0.0000)
Joe	0.6297 (0.0000)	0.8707 (0.0000)	0.6288 (0.0000)	0.8687 (0.0000)
Gaussian–Student's *t*	0.5957 (0.0000)	0.8367 (0.0000)	0.5951 (0.0000)	0.8564 (0.0000)
Gaussian–Clayton	0.5957 (0.0000)	0.8367 (0.0000)	0.5982 (0.0000)	0.8606 (0.0000)
Gaussian–Frank	0.6006 (0.0000)	0.8138 (0.0000)	0.6011 (0.0000)	0.8629 (0.0000)
Gaussian–Gumbel	0.5919 (0.0000)	0.8377 (0.0000)	0.6006 (0.0000)	0.8607 (0.0000)
Gaussian–Joe	0.5857 (0.0000)	0.8367 (0.0000)	0.5980 (0.0000)	0.8578 (0.0000)
Student's *t*–Clayton	0.6322 (0.0000)	0.8583 (0.0000)	0.6027 (0.0000)	0.8621 (0.0000)
Student's *t*–Frank	0.6003 (0.0000)	0.8395 (0.0000)	0.6074 (0.0000)	0.8676 (0.0000)
Student's *t*–Gumbel	0.6028 (0.0000)	0.8357 (0.0000)	0.5910 (0.0000)	0.8543 (0.0000)
Student's *t*–Joe	0.6554 (0.0000)	0.8791 (0.0000)	0.6355 (0.0000)	0.9021 (0.0000)
Clayton–Frank	0.8722 (0.0000)	1.0829 (0.0000)	0.5977 (0.0000)	0.8592 (0.0000)
Clayton–Gumbel	0.5907 (0.0000)	0.8370 (0.0000)	0.5961 (0.0000)	0.8381 (0.0000)
Clayton–Joe	**0.5388 (1.0000)**	**0.7598 (1.0000)**	0.5951 (0.0000)	0.8370 (0.0000)
Frank–Gumbel	0.8751 (0.0000)	1.0853 (0.0000)	0.8971 (0.0000)	1.1583 (0.0000)
Frank–Joe	0.6280 (0.0000)	0.8601 (0.0000)	0.6014 (0.0000)	0.8590 (0.0000)
Gumbel–Joe	0.6086 (0.0000)	0.8459 (0.0000)	0.5974 (0.0000)	0.8593 (0.0000)

Note: The bold numbers indicate the lowest error rate (MAE and RMSE). The parentheses () denote *p*-value of the MCS test, and if the *p*-value is larger than 0.1, it indicates that these models can survive in the MCS test, and the higher the *p*-value, the better forecasting accuracy of the model.
Abbreviations: MAE, mean absolute error; RMSE, root mean squares error; VAR, vector autoregression.

TABLE 8 Out-of-sample forecasting performance evaluation based on MAE and RMSE

	1 step	2 steps	3 steps	4 steps	5 steps	6 steps	Average
MAE							
Model with WUI							
Gaussian	0.6233 (0.0000)	0.7032 (0.0000)	0.7832 (0.0000)	0.8211 (0.0000)	0.8991 (0.0000)	0.9821 (0.0000)	4.8120
Clayton	0.6021 (0.0000)	0.6722 (0.0000)	0.7692 (0.0000)	0.8133 (0.0000)	0.8676 (0.0000)	0.9722 (0.0000)	4.6966
Clayton–Joe	**0.5619 (1.0000)**	**0.6023 (1.0000)**	**0.6792 (1.0000)**	**0.7522 (1.0000)**	**0.8322 (1.0000)**	**0.9091 (1.0000)**	4.3369
Model without WUI							
Gaussian	0.6499 (0.0000)	0.7322 (0.0000)	0.8122 (0.0000)	0.8821 (0.0000)	0.9732 (0.0000)	1.3923 (0.0000)	5.4419
Clayton	0.6389 (0.0000)	0.7301 (0.0000)	0.8232 (0.0000)	0.8702 (0.0000)	1.0013 (0.0000)	1.4902 (0.0000)	5.5539
Clayton–Joe	0.6032 (0.0000)	0.7032 (0.0000)	0.8232 (0.0000)	0.8652 (0.0000)	0.9622 (0.0000)	1.2399 (0.0000)	5.1969
	1 step	2 steps	3 steps	4 steps	5 steps	6 steps	Average
RMSE							
Model with WUI							
Classical VAR	0.8711 (0.0000)	0.9221 (0.0000)	0.9928 (0.0000)	1.1021 (0.0000)	1.2091 (0.0000)	1.2839 (0.0000)	6.3811
Clayton	0.8537 (0.0000)	0.8932 (0.0000)	0.9632 (0.0000)	1.0800 (0.0000)	1.1228 (0.0000)	1.2773 (0.0000)	6.1902
Clayton–Joe	**0.7801 (1.0000)**	**0.8445 (1.0000)**	**0.8922 (1.0000)**	**0.9763 (1.0000)**	**1.0933 (1.0000)**	**1.2003 (1.0000)**	5.7867
Model without WUI							
Classical VAR	0.8941 (0.0000)	0.9334 (0.0000)	1.0012 (0.0000)	1.1720 (0.0000)	1.2681 (0.0000)	1.7211 (0.0000)	6.9899
Clayton	0.8753 (0.0000)	0.9291 (0.0000)	0.9872 (0.0000)	1.1562 (0.0000)	1.3721 (0.0000)	1.9822 (0.0000)	7.3021
Clayton–Joe	0.8561 (0.0000)	0.8921 (0.0000)	0.9872 (0.0000)	1.1449 (0.0000)	1.2432 (0.0000)	1.5928 (0.0000)	6.7163

Note: The bold numbers indicate the lowest error rate (MAE and RMSE). The *p*-values of the MCS test are larger than the 0.1, indicating that these models can survive in the MCS test, and the higher the *p*-value means the better forecasting accuracy by the model.
Abbreviations: MAE, mean absolute error; RMSE, root mean squares error; VAR, vector autoregression; WUI, world uncertainty index.

their corresponding MCS tests provide a consensus result. Based on the RMSE and MSE criterion, Clayton–Joe copula-based VAR with WUI still performs the best in all forecasting horizons, 1 up to 6 steps ahead. The model's score is the smallest in term of RMSE and MAE when compared with other models, including those models without WUI, indicating the deviation between the predicted values obtained from Clayton–Joe copula-based VAR with WUI and their actual observations are smaller with a better forecasting accuracy than other competing models. Furthermore, based on the MCS test results, we observe that all *p*-values of our Clayton–Joe copula-based VAR with WUI are equal to 1.000. These results are consistent with that in Table 7. We also find that the model combining with WUI has a better performance than the model without WUI, confirming that the WUI has a good forecasting power in the growths of the advanced stock market, emerging stock market, and the US GDP.

5 | CONCLUSION

The study attempts to extend the VAR model using the mixed copula approach to join the VAR model's multivariate marginal distribution. Thus, the model becomes more flexible in dealing with various complicated dependence structures of the joint distribution. Several copula classes are considered and compared in this study: Elliptical, Archimedean, and mixed Copula, which is a combination of two or more copula families from Elliptical and Archimedean classes. The estimation technique used in this study is the maximum likelihood method. We employ the full maximum likelihood (FML) (one-step estimation) rather than the inference for margins (IFM) (two-step estimation) as we aim to improve the likelihood of the VAR model in order to gain more accurate parameter estimates. Two experimental studies are suggested to validate the accuracy of the one-step maximum likelihood estimation as well as the performance of our model.

We further investigate the performance of the mixed copula-based VAR model using real data analysis. Our overall result shows that the Clayton–Joe copula-based VAR model provides the best fit for our data. This model's BIC is lower than other copula-based VAR specifications as well as the traditional VAR model. In addition, we further conduct the in- and out-of-sample forecasts. The results suggest that Clayton–Joe copula-based VAR model with WUI performs superiority in this application study. Our findings have important economic implications for stock investors and the US policymakers to pay more attention to the change of worldwide EPU.

We leave to study further the issue of deriving the mathematical proof of asymptotic properties and consistency. In addition, our proposed model and estimation can be straightforwardly applied to other copula families. Future research may consider extending our copula-based VAR to compute the nonlinear impulse response functions. Finally, the Bayesian approach is another estimation that could be used to estimate the large parameter estimates in our proposed model.

ACKNOWLEDGMENT

We are grateful to Siem Jan Koopman and two anonymous referees for useful comments and suggestions on earlier versions of this manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data used to support this analysis are available from https://www.dallasfed.org/institute/dgei/gdp.aspx and https://worlduncertaintyindex.com/data/, and code of the proposed model generated during the study is available from the corresponding author upon reasonable request.

ORCID

Woraphon Yamaka https://orcid.org/0000-0002-0787-1437

REFERENCES

Ahir, H., Bloom, N. A., & Furceri, D. (2018). World uncertainty index. Stanford University Mimeo. http://policyuncertainty.com/wui_quarterly.html

Al-Thaqeb, S. A., & Algharabali, B. G. (2019). Economic policy uncertainty: A literature review. *The Journal of Economic Asymmetries*, *20*(C), e00133. https://doi.org/10.1016/j.jeca.2019.e00133

Brechmann, E. C., & Czado, C. (2015). COPAR—Multivariate time series modeling using the copula autoregressive model. *Applied Stochastic Models in Business and Industry*, *31*(4), 495–514. https://doi.org/10.1002/asmb.2043

Campbell, J. Y. (2008). Viewpoint: Estimating the equity premium. *Canadian Journal of Economics*, *41*, 1–21. https://doi.org/10.1111/j.1365-2966.2008.00453.x

Castelnuovo, E., Lim, G. C., & Pellegrino, G. (2017). A short review of the recent literature on uncertainty. *The Australian Economic Review*, *50*(1), 68–78. https://doi.org/10.1111/1467-8462.12210

Christopoulos, D., McAdam, P., & Tzavalis, E. (2021). Dealing with endogeneity in threshold models using copulas. *Journal of Business & Economic Statistics*, *39*(1), 166–178. https://doi.org/10.1080/07350015.2019.1647213

Fan, Y., & Patton, A. J. (2014). Copulas in econometrics. *Annual Review of Economics*, *6*(1), 179–200. https://doi.org/10.1146/annurev-economics-080213-041221

Fermanian, J.-D. (2017). Recent Developments in Copula Models. *Econometrics*, *5*, 34. https://doi.org/10.3390/econometrics5030034

Grossman, V., Mack, A., & Martínez-García, E. (2014). Database of Global Economic Indicators (DGEI): A methodological note. *Journal of Economic and Social Measurement*, *39*(3), 163–197. https://doi.org/10.3233/JEM-140391

Gupta, R., Lau, C. K. M., & Wohar, M. E. (2019). The impact of U.S. uncertainty on the euro area in good and bad times: Evidence from a quantile structural vector autoregressive model. *Empirica*, *46*, 353–368. https://doi.org/10.1007/s10663-018-9400-3

Gupta, R., Ma, J., Risse, M., & Wohar, M. E. (2018). Common business cycles and volatilities in U.S. states and MSAs: The role of economic uncertainty. *Journal of Macroeconomics*, *57*, 317–337. https://doi.org/10.1016/j.jmacro.2018.06.009

Gupta, R., Olasehinde-Williams, G. A., & Wohar, M. E. (2020). The impact of U.S. uncertainty shocks on a panel of advanced and emerging market economies: The role of exchange rate, trade and financial channels. *The Journal of International Trade & Economic Development*, *29*(6), 711–721. https://doi.org/10.1080/09638199.2020.1720785

Hamilton, J. D. (1994). *Time series analysis* (Vol. 2, pp. 690–696). Princeton University Press.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497. https://doi.org/10.1080/09638199.2020.1720785

Hofert, M., Mächler, M., & McNeil, A. J. (2012). Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis*, *110*, 133–150. https://doi.org/10.1016/j.jmva.2012.02.019

Hu, L. (2006). Dependence patterns across financial markets: A mixed copula approach. *Applied Financial Economics*, *16*(10), 717–729. https://doi.org/10.1080/09603100500426515

Joe, H., & Xu, J. J. (1996). *The Estimation Method of Inference Functions for Margins for Multivariate Models*. University of British Columbia.

Kaewsompong, N., Maneejuk, P., & Yamaka, W. (2020). Bayesian estimation of Archimedean copula-based SUR quantile models. *Complexity*, *2020*, 6746303. https://doi.org/10.1155/2020/6746303

Liu, J., Wang, M., Ma, J., Rahman, S., & Sriboonchitta, S. (2020). A simultaneous stochastic frontier model with dependent error components and dependent composite errors: An application to Chinese banking industry. *Mathematics*, *8*(2), 238. https://doi.org/10.3390/math8020238

Louzada, F., & Ferreira, P. H. (2016). Modified inference function for margins for the bivariate clayton copula-based SUN Tobit Model. *Journal of Applied Statistics*, *43*(16), 2956–2976. https://doi.org/10.1080/02664763.2016.1155204

Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer. https://doi.org/10.1007/978-3-540-27752-1

Ma, R., Zhou, C., Cai, H., & Deng, C. (2019). The forecasting power of EPU for crude oil return volatility. *Energy Reports*, *5*, 866–873. https://doi.org/10.1016/j.egyr.2019.07.002

Maneejuk, P., Pastpipatkul, P., & Sriboonchitta, S. (2016). Economic growth and income inequality: Evidence from Thailand. In Y. Tang, V.-N. Huynh, & J. Lawry (Eds.), *International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making* (pp. 649–663). Springer. https://doi.org/10.1007/978-3-319-49046-5_55

Maneejuk, P., Yamaka, W., & Sriboonchitta, S. (2018). Mixed-copulas approach in examining the relationship between oil prices and ASEAN's stock markets. In L. H. Anh, L. S. Dong, V. Kreinovich, & N. Ngoc Thach (Eds.), *Economics for Financial Applications* (pp. 531–541). Springer. https://doi.org/10.1007/978-3-319-73150-6_42

Mohti, W., Dionísio, A., Ferreira, P., & Vieira, I. (2019). Contagion of the subprime financial crisis on frontier stock markets: A copula analysis. *Economies*, *7*(1), 15. https://doi.org/10.3390/economies7010015

Nelsen, R. B. (2006). *An introduction to copulas*. Springer-Verlag.

Nguyen, C., Bhatti, M. I., Komorníková, M., & Komorník, J. (2016). Gold price and stock markets nexus under mixed-copulas. *Economic Modelling*, *58*, 283–292. https://doi.org/10.1016/j.econmod.2016.05.024

Pastpipatkul, P., Maneejuk, P., & Sriboonchitta, S. (2017). Testing the validity of economic growth theories using copula-based seemingly unrelated quantile kink regression. In V. Kreinovich, S. Sriboonchitta, & V. N. Huynh (Eds.), *Robustness in Econometrics* (pp. 523–541). Springer. https://doi.org/10.1007/978-3-319-50742-2_32

Pastpipatkul, P., Maneejuk, P., Wiboonpongse, A., & Sriboonchitta, S. (2016). Seemingly unrelated regression based copula: An application on Thai rice market. In V.-N. Huynh, V. Kreinovich, & S. Sriboonchitta (Eds.), *Causal Inference in Econometrics* (pp. 437–450). Springer. https://doi.org/10.1007/978-3-319-27284-9_28

Shumway, R. H., & Stoffer, D. S. (2011). Time series regression and exploratory data analysis. In R. H. Shumway & D. S. Stoffer (Eds.), *Time Series Analysis and Its Applications* (pp. 47–82). Springer. https://doi.org/10.1007/978-1-4419-7865-3_2

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, *48*(1), 1–48. https://doi.org/10.2307/1912017

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publications de l'Institut Statistique de l'Université de Paris*, *8*, 229–231.

Tansuchat, R., & Maneejuk, P. (2018). Modeling Dependence with Copulas: Are Real Estates and Tourism Associated? In V.-N. Huynh, M. Inuiguchi, D. H. Tran, & T. Denoeux (Eds.), *International symposium on integrated uncertainty in knowledge modelling and decision making* (pp. 302–311). Springer. https://doi.org/10.1007/978-3-319-75429-1_25

Tsay, R. S. (2005). *Analysis of Financial Time Series* (Vol. 543). John Wiley and Sons. https://doi.org/10.1002/0471746193

Xu, G., & Gao, W. (2019). Financial risk contagion in stock markets: Causality and measurement aspects. *Sustainability*, *11*(5), 1402. https://doi.org/10.3390/su11051402

Yamaka, W., & Thongkairat, S. (2020). A mixed copula-based vector autoregressive model for econometric analysis. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *28*(Supp01), 113–121. https://doi.org/10.1142/S0218488520400103

AUTHOR BIOGRAPHIES

Woraphon Yamaka received bachelor's, master's, and PhD degrees in economics from Chiang Mai, Thailand, in 2011, 2014, and 2017, respectively. He has been the assistant professor of the Faculty of Economics, Chiang Mai University, since 2018, where he is currently the Vice-Director of the Centre of Excellence in Econometrics. Since 2015, he has published more than 200 articles in journals that are indexed in Scopus and ISI. His current h-index in Scopus is 11. His research interests include economics, econometrics, and machine learning.

Rangan Gupta is a Professor at the Department of Economics, University of Pretoria. He holds a PhD in Economics from the University of Connecticut, USA. He conducts research in the fields of macroeconomics and financial economics. His research has appeared in high impact factor international journals such as *Energy Economics, International Journal of Forecasting, Journal of Banking and Finance, Journal of Financial Markets, Journal of Forecasting, Journal of International Money and Finance*, and *Macroeconomic Dynamics*.

Sukrit Thongkairat received the bachelor's degree and the master's degree in economics from Chiang Mai University, in 2014 and 2016, respectively. Since 2018, he has published 12 articles that are indexed in Scopus and ISI. His research interests include machine learning, finance, statistics, and econometrics. His current h-index in Scopus is 3.

Paravee Maneejuk has been an Associate Professor at the Faculty of Economics, Chiang Mai University,

Thailand, since 2018. Her research interests include information theory, economic development, growth, and applied econometrics. She has also worked with a research team from the Centre of Excellence in Econometrics and has published about 100 articles in SCOPUS/ISI with a Scopus h-index as 10. She is currently working with several research teams in applying econometric models to solve several economic, finance, and business issues.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Yamaka, W., Gupta, R., Thongkairat, S., & Maneejuk, P. (2023). Structural and predictive analyses with a mixed copula-based vector autoregression model. *Journal of Forecasting*, *42*(2), 223–239. https://doi.org/10.1002/for.2902