Vine copulas and applications to the European Union sovereign debt analysis

Dalu Zhang[a,*]

[a]*School of Economics, University of East Anglia, Norwich, NR4 7TJ, UK*

Abstract

European sovereign debt crisis has become a very popular topic since late 2009. In this paper, sovereign debt crisis is investigated by calculating the probabilities of the potential future crisis of 11 countries in the European Union. We use sovereign spreads of the European countries against Germany as targets and apply the GARCH based vine copula simulation technique. The methodology solves the difficulties of calculating the probabilities of rarely happening events and takes sovereign debt movement dependence, especially tail dependence, into consideration. Results indicate that Italy and Spain are the most likely next victims of the sovereign debt crisis, followed by Ireland, France and Belgium. The UK, Sweden and Denmark, which are outside the euro area, are the most financially stable countries in the sample.

1. Introduction

The ongoing European sovereign debt crisis originated in Greece, but the impact has spread all over the European Union especially in the euro area. On 8th Dec, 2009, rating agency Fitch cut Greece's long-term debt from A- to BBB+. Because of the lack of confidence in investing in Greek government bonds, the yield of 10-year government bonds jumped up significantly. In the mean time, the bond yield of peripheral European countries Spain and Portugal also increased along with Greece. In Ireland and Italy, however, the yields decreased. This phenomenon shows that yield differentials across European bond markets have not been wiped out completely,

*Corresponding author

Email address: Dalu.Zhang@uea.ac.uk (Dalu Zhang)

although accelerated financial integration among euro bond markets has been widely expected, since the macroeconomic and fiscal indicators have shown significant improvement for the higher risk euro markets, creating a potential for those members to converge with lower risk members in terms of bond returns. Finding the relationship between the yields of these countries' sovereign bonds might be a useful way to understand how they will influence each other, especially in extreme events. This information could then be used to assess the risk level of a sovereign bond. In order to achieve this, a GARCH based vine copula simulation method to analyse the sovereign debts in the European Union is proposed in this paper.

As a popular multivariate modeling tool, copula is widely used in many fields where the multivariate dependence matters, such as actuarial science (Frees et al., 1996), biomedical studies (Wang and Wells, 2000), engineering (Genest and Favre, 2007) and finance (Embrechts et al., 2003). In finance, the misuse of the copula method in the pricing of collateralized debt obligations (CDO) is considered by journalists to be one of the reasons that led to the global financial crisis of 2008 - 2009 (Salmon, 2009; Jones, 2009). The copula approach provides a method of isolating the description of the dependence structure and understanding the dependence at a deeper level. It expresses dependence on a quantile scale, which is useful for describing the dependence of extreme outcomes and is natural in a risk-management context. Due to the advantages of the copula method, it is an ideal tool for analysing the relationship of sovereign debts between countries in the European Union.

The main difficulty about sovereign debt crisis analysis is that the crisis rarely happens. It is extremely hard for statisticians to analyse an event which has never happened before. In order to solve this issue, this paper uses simulation methods to create unknown situations. This paper replicates 10000 iterations of a 365 future day simulation of sovereign spreads against Germany of 11 countries in the European Union. In the mean time, the relationships between the countries are considered. Then, the percentage chance of the crisis events is calculated, which is the probability of future crisis. In terms of defining crisis events, Sy (2004)'s definition of sovereign debt crisis is adopted, which is that sovereign spread against the US is more than 1000 basis points. In the same manner, a country experiencing a sovereign debt crisis is defined as being when its sovereign spread against Germany is greater than 1000 basis points in this research.

The contribution of this research is fourfold: firstly, this is the first analysis of extreme value and tail dependence of sovereign debt spread movement in the European Union; secondly, this study conducts the comparison between 11 countries in the European Union at the same time; thirdly, this paper uses vine copulas to deal with large numbers of dimensions and satisfies the wide range of dependence, flexible range of upper and lower tail dependence, computationally feasible density for estimation, and closure property under marginalization simultaneously; fourthly, which is also the key feature of this paper, the research identifies the risk level of sovereign debt in different countries in the European Union.

Daily 10-year government bond yields from 18/06/1997 to 12/03/2012 in Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Spain, Sweden and the UK are used in this research.

The results show that the estimated crisis probabilities of Greece and Portugal in the next 365 days are 100% and 99.77%, which is consistent with the situation that they are already in crisis. Spain and Italy show great potential to be the next victims in one year's time. France and Belgium show some instability in the results and the probability of crisis is fairly high: 63.13% and 60.14% respectively. Netherlands is next with an almost 1 in 4 chance of crisis and it is the most stable country in the euro area. In the mean time, countries outside the euro area in the sample which are the UK, Sweden and Denmark show the greatest stability in their sovereign bonds.

The remainder of the paper is as follows. Section 2 is a literature review in sovereign debt analysis and copula methods. Section 3 is the data description. Section 4 discusses the bivariate relationships of these pairs of countries. Section 5 explains the vine copula approach. Section 6 shows the results of simulation and calculation of the risk levels of the countries. And Section 7 concludes.

2. Literature review

The literature on sovereign debt analysis generally uses sovereign bond spread between a target country and a benchmark country to assess the default risk level of the target country. Structural approaches developed from the Merton model 1974 and reduced form models such as the Jarrow and Turnbull (1995) approach are the two main streams.

The structural approaches explain the sovereign spread endogenously using both enterprise value volatility and firm default definition. The pitfalls of these approaches are not only their difficulty and lack of accuracy to define appropriate country-specific proxy variables for the level of indebtedness, but also they disregard the fact that default incentives of a country are more complicated than those of enterprises. The reduced form approaches use different macro variables as the determinants of the sovereign default risk. Literature such as Reinhart et al. (2003), Eichengreen et al. (2003) and Goldstein and Turner (2004), analyse the sovereign debt risk of emerging market economies. Their focus is on the sustainability of the sovereign debt and the currency mismatches. They measure default risk by using country credit ratings. The disadvantage of these approaches is that these credit ratings are inefficient and cannot be adjusted in a timely manner to adapt to the market data when a big crisis is ongoing. Most recently, Dötz and Fischer (2010) use a GARCH-in-mean based reduced form model to analyse the factors triggering the sovereign spread movement in the European Union and the result shows that the expectation of loss is the main reason sovereign spreads widened during the recent global financial crisis. Nonetheless, both structural and reduced form approaches face a problem: they ignore the yields movement dependence with other countries, which is especially important inside the European Union.

Both multivariate extreme value theory (EVT) and copula method can solve these problems in order to capture the probabilities of rare events. Multivariate EVT is developed by de Haan and Resnick (1977) for limiting distribution of the componentwise maximum of independent and identically distributed (i.i.d.) random vectors. The technique has since been comprehensively developed. Although in the literature multivariate EVT are available in a d-dimensional context, the computational complexity increases significantly with the increase of d (Fougères, 2003). For instance, applications which are done by de Haan and de Ronde (1998), Bruun and Tawn (1998), de Haan and Ferreira (2006) as with most work done in the multivariate EVT context, are limited to 2 and 3 dimensions. With reference to copula method, there is a large body of literature using copulas in a financial context (Bouyé et al., 2000; Embrechts et al., 2003; Cherubini et al., 2004). Most of them are used to compute Value at Risk (VaR) and expected shortfall (ES) of the stock or bond portfolio by applying single copula families such as elliptical copulas and Archimedean copulas. There are many limitations on those copula families applied in the above literature. Elliptical copulas are widely used, but they cannot model the financial tail dependence very well (Patton, 2008). Archimedean copulas are not satisfactory for modeling with dimensions higher than two (Joe, 1997). Multivariate

Archimedean copulas only allow exchangeable structure with a narrower range of negative dependence in a higher dimension (McNeil and Neslehova, 2009). Partially symmetric copulas extend Archimedean to a class with a non-exchangeable structure, but the dependence they provide are not particularly flexible (Joe, 1993). Mix-id copulas in Joe and Hu (1996) provide flexible positive dependence by construction, but only upper tail dependence is flexible not lower tail. Demarta and McNeil (2005) provides multivariate skewed-t copulas, which model well, but are computationally more involved. Similarly to multivariate EVT, these copula methods experience limitation about dimension. Vine copulas were proposed by Joe (1996) and explained in detail by Bedford and Cooke (2002). At that time, vine copulas models were a graphical model using bivariate copulas to construct multivariate copulas. Aas et al. (2009) run statistical inference on two types of vines: canonical vine (C-vine) and drawable vine (D-vine). These models have been improved by Nikoloulopoulos et al. (2012) which can satisfy most of the features that should be included in a copula model: firstly, a wide range of dependence including both positive and negative dependence; secondly, a flexible range of upper and lower tail dependence; thirdly, and most importantly, computationally feasible density for estimation, even for high dimension estimation. According to the aim of this analysis, which is focusing on the interactions of the 11 countries and assessing the crisis probabilities of countries simultaneously, vine-copula method is preferred to other models above.

In this paper, a GARCH based Vine copula method is used to analyse the tail dependence and calculate probabilities of sovereign debt crisis of these countries in certain periods of time in the European Union.

3. Data

Daily 10-year government bond yields from 18/06/1997 to 12/03/2012 in Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Spain, Sweden and the UK are used in this research. All data are collected from Thomson Reuters ECOWIN. The target variable, sovereign spread against Germany, is calculated as

$$\Delta(i_j - i^*),$$

where $j = 1, \ldots, d$, i_j is 10-year government bond yield of a target country, and i^* is 10-year government bond yield of Germany.

4. Bivariate copula analysis

4.1. GARCH filter

Vine copula modeling proceeds in three stages. In the first stage, the model for the individual variables (i.i.d) is selected, which is the marginal distribution. For financial time series data, a GARCH filter with innovation being student-t distribution is applied for the purpose of making the data independent and identically distributed (Aas and Berg, 2009). Using Box-Jenkins analysis method (Box and Jenkins, 1970), all $\Delta(i_j - i^*)$ are determined to be MA(1) process. In order to find the best model to fit the series, MA(1)-GARCH(1,1), MA(1)-EGARCH(1,0)[1] and MA(1)-TGARCH(1,1) are proposed in this stage. Q-statistic (Ljung and Box, 1978) and ARCH LM test (Engle, 1982) are conducted at the same time for testing autocorrelation of residuals and squared residuals respectively.

The MA(1)-GARCH(1,1) model can be expressed as follows:

$$\Delta(i - i^*)_{t,j} = \mu_j + \epsilon_{t,j} + \theta\epsilon_{t-1,j}, \tag{1}$$

$$\epsilon_{t,j} = z_{t,j}\sigma_{t,j}, \tag{2}$$

$$\sigma^2_{t,j} = \alpha_{0,j} + \alpha_{1,j}\epsilon^2_{t-1,j} + \beta_{1,j}\sigma^2_{t-1,j}, \tag{3}$$

where $j = 1, \ldots, d$, $t = 1, \ldots, T$, $\Delta(i - i^*)$ is sovereign spread against Germany (i^*) of a target country (i), $z_t \sim T(0, 1, \nu)$, The conditions of coefficients that ensure positive volatility and existence of second moment are $\alpha_1 > 0$, $\beta_1 > 0$ and $\alpha_1 + \beta_1 < 1$.

The MA(1)-EGARCH(1,0) model may generally be specified as follows:

$$\Delta(i - i^*)_{t,j} = \mu + \epsilon_{t,j} + \theta\epsilon_{t-1,j}, \tag{4}$$

$$\epsilon_{t,j} = z_{t,j}\sigma_{t,j}, \tag{5}$$

$$\ln\sigma^2_{t,j} = \alpha_{0,j} + \gamma_{1,j}(|\frac{\epsilon_{t-1,j}}{\sigma_{t-1,j}}| - E|\frac{\epsilon_{t-1,j}}{\sigma_{t-1,j}}|) + \beta_{1,j}\ln\sigma^2_{t-1,j}, \tag{6}$$

where $j = 1, \ldots, d$, $t = 1, \ldots, T$, $\Delta(i - i^*)$ is sovereign spread against Germany (i^*) of a target country (i), $z_t \sim T(0, 1, \nu)$.

The MA(1)-TGARCH(1,1) model is represented by the expression:

$$\Delta(i - i^*)_{t,j} = \mu + \epsilon_{t,j} + \theta\epsilon_{t-1,j}, \tag{7}$$

$$\epsilon_{t,j} = z_{t,j}\sigma_{t,j}, \tag{8}$$

$$\sigma_{t,j} = \alpha_{0,j} + \alpha_{1,j}|z_{t-1,j}| + \beta_{1,j}\sigma_{t-1,j} + \delta_{1,j}z_{t-1,j}, \tag{9}$$

[1] MA(1)-EGARCH(1,1) was also considered, and all the coefficients α_1 are insignificant.

where $j = 1, \ldots, d$, $t = 1, \ldots, T$, $\Delta(i - i^*)$ is sovereign spread against Germany (i^*) of a target country (i), $z_t \sim T(0, 1, \nu)$. The conditions of coefficients which guarantee positive conditional volatility are $\alpha_0 > 0$, $\alpha_1 > 0$, $\beta_1 > 0$, $|\delta_1| < \alpha_1$ and ${\alpha_1}^2 + {\beta_1}^2 + {\delta_1}^2 + 2\alpha_1\beta_1\nu_1 < 1$, where $\nu_1 = \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}$ for z_t is student-t distributed (Rodriguez and Ruiz, 2012).

Table 1,2 and 3 present the results of MA(1)-GARCH(1,1), MA(1)-EGARCH(1,0), MA(1)-TGARCH(1,1), respectively. In Table 1, all the coefficients satisfy the condition $\alpha_1 > 0$, $\beta_1 > 0$ and $\alpha_1 + \beta_1 < 1$, which ensure the positive conditional volatility and confirm the existence of second moment of a standard GARCH model. In Table 3, all the coefficients meet the requirements $\alpha_0 > 0$, $\alpha_1 > 0$, $\beta_1 > 0$,$|\delta_1| < \alpha 1$ and ${\alpha_1}^2 + {\beta_1}^2 + {\delta_1}^2 + 2\alpha_1\beta_1\nu_1 < 1$, which guarantees positive conditional volatility as well as the existence of the second moment of a TGARCH model. According to Akaike information criterion (Akaike, 1974), MA(1)-TGARCH(1,1) model fits the data the best, and then MA(1)-EGARCH(1,0), and last place is MA(1)-GARCH(1,1). However, in MA(1)-TGARCH(1,1) model, coefficients δ of DEN, FRA, and POR are insignificant in 95% confidence interval, which means there is no threshold effect in these models. In the mean time, ARCH LM tests of MA(1)-TGARCH(1,1) in FRA, POR and UK indicate autocorrelation of squared standardized residuals. The above results suggest that MA(1)-TGARCH(1,1) fit for BEL, GRE, IRE, ITA, NET, SPA, SWE the best. The next best model MA(1)-EGARCH(1,0) is considered for DEN, FRA, POR and UK. ARCH LM tests of MA(1)-EGARCH(1,0) imply that there are autocorrelations in squared standardized residuals for FRA and UK. With the insignificant coefficients of threshold parameter in MA(1)-TGARCH(1,1), this suggests the series of FRA and UK could be symmetric. Q-Statistics are mostly insignificant in 95% significance level, which represents no autocorrelation in the residuals.

In summary, the best model fit for BEL, GRE, IRE, ITA, NET, SPA and SWE is MA(1)-TGARCH(1,1); the best model fit for DEN and POR is MA(1)-EGARCH(1,0); and the best model fit for FRA and UK is MA(1)-GARCH(1,1).

Table 1: Results of MA(1)-GARCH(1,1)

	BEL	DEN	FRA	GRE	IRE	ITA	NET	POR	SPA	SWE	UK
μ	-0.0002	-0.00031	2.07E-05	-0.00031	-0.00024	-0.00027	-0.00014	-0.00016	-0.0003	-0.00046	2.32E-05
θ	-0.35547*	-0.46821*	-0.47978*	-0.27817*	-0.33931*	-0.28703*	-0.49628*	-0.35214*	-0.31153*	-0.23441*	-0.19524*
α_0	3.49E-06*	3.12E-05*	2.82E-06*	3.31E-05*	8.15E-06*	1.91E-06*	2.46E-06	2.15E-05*	2.83E-06*	5.02E-05*	2.13E-05*
α_1	0.142989*	0.169135*	0.113557*	0.192742*	0.154521*	0.098451*	0.135867*	0.206958*	0.117413*	0.125711*	0.055155*
β_1	0.856011*	0.806349*	0.885443*	0.806258*	0.844479*	0.900549*	0.863133*	0.792042*	0.881587*	0.835442*	0.930337*
ν	4.74224*	5.109619*	4.925748*	4.378324*	4.900171*	5.2668*	4.430659*	4.015165*	4.774563*	5.781193*	5.646772*
$\alpha_1 + \beta_1$	0.999	0.975484	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.961153	0.985493
AIC	-4.87767	-4.57883	-4.99516	-2.75439	-4.16988	-4.34567	-5.3233	-3.96323	-4.56947	-4.1455	-3.9174
Q-stat for standardized residuals											
lag1	2.653	0.406	0.307	2.998	6.952	0.008	0.005	1.299	1.367	1.368	0.645
lag3	2.999	0.828	0.772	4.339	7.494	0.586	0.793	2.442	3.087	3.603	2.808
lag7	10.323	7.547	6.528	6.456	10.531	4.343	3.125	4.442	5.409	9.304	6.06
ARCH LM test											
lag2	0.5451	0.059	4.324	0.003	1.692	0.826	0.047	1.497	1.244	5.369	3.268
lag5	1.3756	0.223	9.373	0.241	3.104	2.16	0.348	4.204	4.48	6.654	4.043
lag10	3.862	0.562	11.387	0.582	4.894	3.05	1.111	9.164	6.899	7.933	7.527

Note:* is significant in the 95% confidence interval.

Table 2: Results of MA(1)-EGARCH(1,0)

	BEL	DEN	FRA	GRE	IRE	ITA	NET	POR	SPA	SWE	UK
μ	-0.00019	-0.00029	1.92E-05	0.000187	-0.00014	-0.00017	-0.00018	-0.00014	-0.00026	-0.00039	-0.00011
θ	-0.34243*	-0.45268*	-0.48048*	-0.24924*	-0.33173*	-0.27014*	-0.48491*	-0.33622*	-0.30217*	-0.23033*	-0.18705*
$\alpha 0$	-0.05887*	-0.15162*	-0.06527*	-0.06603*	-0.03301*	-0.01592*	-0.03989	-0.05047*	-0.02427*	-0.24019*	-0.05975*
$\beta 1$	0.992167*	0.979542*	0.991508*	0.987961*	0.995293*	0.997969*	0.995033*	0.992279*	0.99662*	0.965449*	0.991166*
$\gamma 1$	0.276111*	0.173309*	0.257267*	0.265489*	0.198386*	0.162121*	0.189471*	0.273383*	0.208379*	0.200036*	0.092029*
ν	3.996966*	5.207032*	4.26362*	3.466955*	4.458685*	4.523871*	3.726167*	3.519717*	4.033725*	5.839347*	5.807454*
AIC	-4.88445	-4.58675	-5.00522	-2.78533	-4.16871	-4.36446	-5.33835	-3.97388	-4.58738	-4.14811	-3.92529
Q-stat for standardized residuals											
lag1	2.29	1.158	0.12	2.065	2.995	0.222	0.029	0.001	2.058	2.011	0.001
lag3	2.676	1.911	0.409	3.76	3.824	1.289	0.307	0.098	6.505	4.747*	1.731
lag7	10.442	8.312	6.89	7.761	8.618	5.067	2.124	2.291	7.725	10.026	5.45
ARCH LM test											
lag2	0.581	0.592	7.594*	0.173	0.611	2.05	0.0097	5.552	1.698	5.609	42.111*
lag5	0.914	0.72	12.381*	0.327	0.801	2.788	0.202	6.975	3.048	6.075	42.579*
lag10	2.151	1.147	18.847*	0.584	1.195	3.239	0.49	12.208	4.343	6.561	45.534*

Note:* is significant in the 95% confidence interval.

Table 3: Results of MA(1)-TGARCH(1,1)

	BEL	DEN	FRA	GRE	IRE	ITA	NET	POR	SPA	SWE	UK
μ	-0.00013	-0.00034	1.68E-05	4.19E-05	-0.0001	-0.00015	-0.00014	-0.00016	-0.00024	-0.00044	-0.00012
θ	-0.34019*	-0.44969*	-0.48155*	-0.24573*	-0.33402*	-0.26643*	-0.48293*	-0.33827*	-0.30044*	-0.22685*	-0.18653*
$\alpha 0$	0.000142*	0.000588*	0.000174*	0.000535*	0.000158*	6.12E-05*	0.000101	0.000345*	0.00011*	0.000913*	0.000357
$\alpha 1$	0.156665*	0.113818*	0.14433*	0.19591*	0.125661*	0.108066*	0.144631*	0.174769*	0.139214*	0.105767*	0.053928*
$\beta 1$	0.876154*	0.890765*	0.89129*	0.851958*	0.894857*	0.910654*	0.891647*	0.870463*	0.8955*	0.892559*	0.949712*
δ_1	-0.02407*	-0.0016	-0.01639	-0.03508*	-0.02487*	-0.02553*	-0.01252*	-0.01672	-0.02581*	0.00323*	0.00139*
ν	3.937861*	5.176186*	4.265713*	3.5826*	4.623747*	4.55136*	3.770753*	3.598538*	4.161302*	5.829874*	5.724568*
condition	0.986182	0.956137	0.999887	0.994651	0.980641	0.984313	0.995846	0.997733	0.999783	0.949083	0.981343
AIC	-4.89272	-4.5909	-5.00669	-2.7918	-4.17984	-4.37049	-5.35006	-3.98503	-4.59574	-4.15006	-3.92382
Q-stat for standardized residuals											
lag1	1.133	1.59	0.025	2.935	2.621	0.272	0.085	0.142	3.204	1.923	0.022
lag3	1.702	2.183	0.353	3.066	3.093	0.923	0.591	0.231	5.182	4.702	1.722
lag7	9.507	9.011	6.984	10.861	7.329	4.623	3.03	2.062	7.576	9.93	5.643
ARCH LM test											
lag2	1.818	1.086	10.45*	0.697	1.277	2.241	0.432	9.936*	1.269	5.387	42.88*
lag5	2.603	1.189	15.02*	0.907	1.601	3.219	0.555	11.554*	2.601	5.752	43.48*
lag10	4.145	1.479	21.17*	1.201	2.162	3.669	0.891	16.062*	3.963	6.123	46.99*

Note:* is significant in the 95% confidence interval. The "condition" is the calculated condition ${\alpha_1}^2+{\beta_1}^2+{\delta_1}^2+2\alpha_1\beta_1\nu_1$, where $\nu_1 = \sqrt{\frac{\nu-2}{\pi}}\frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}$ and if it is smaller than 1, there will be guaranteed positive conditional volatility and second moment for TGARCH model.

4.2. Bivariate copula analysis

In the second stage, pairs of data using various families are modeled in order to select the proper copula family by goodness-of-fit tests. Different copula families have different characteristics of tail dependence that allow us to identify the tail-dependence between different pairs. In the third stage, we construct a vine using copula families which are estimated in the second step. The vine type selection and copula indexing are involved in this stage as well.

In the first stage, the different GARCH filters are applied in this research. In the second stage, the Vuong (1989) test and the Clarke (2007) test are used to select the best copulas that fit the pairs as goodness-of-fit tests. These two tests compare two models against each other. Based on their null hypothesis, the tests will identify the better model by a statistically significant decision. Belgorodski (2010) proposes a method using these two tests for copula selection.

Using this method, a bivariate copula model A is compared with all other possible bivariate copula models. If copula model A outperforms another copula model, a score of "+1" is assigned to model A, and at the same time a score of "−1" will be added to the other copula model. No score will be added when the test cannot identify which model is better. There is a total score which sums up the scores we get from all these pairwise comparisons. Both the Vuong test and the Clarke test are likelihood ratio based and use the common Kullback-Leibler information criterion, which measures the distance between two statistical models. For instance, c_1 and c_2 are two bivariate copula with estimated parameters $\hat{\theta}_1$ and $\hat{\theta}_2$ respectively. The Vuong test requires a sum, ν, of the log differences of their point-wise likelihoods m_i. For observations $u_{i,j}$, $i = 1, \ldots, N$, $j = 1, 2$,

$$m_i = log \left[\frac{c_1(u_{i,1}, u_{i,2}|\hat{\theta}_1)}{c_2(u_{i,1}, u_{i,2}|\hat{\theta}_2)} \right], \tag{10}$$

and then

$$\nu = \frac{\frac{1}{n}\sum_{i=1}^{N} m_i}{\sqrt{\sum_{i=1}^{N}(m_i - \bar{m})^2}}. \tag{11}$$

The null hypothesis of the Vuong test is

$$H_0 : E(m_i) = 0, \forall i = 1, \ldots, N.$$

Vuong (1989) shows that ν is asymptotically standard normal distributed. Therefore, model A is preferred against model B at level α if

$$\nu > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right). \tag{12}$$

In the same manner, if $\nu < -\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$,then model B is chosen. Nonetheless, if $|\nu| \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$, then the test cannot identify if there is a better one which will not reject the null hypothesis of the test as well.

On the other hand, the null hypothesis of the Clarke test is

$$H_0 : P(m_i > 0) = 0.5, \forall i = 1, \ldots, N,$$

and the test statistic is specified as

$$B = \sum_{i=1}^{N} \mathbf{1}_{(0,\infty)}(m_i), \tag{13}$$

where $\mathbf{1}$ is proposed by Clarke (2007) as the indicator of the function. It is binomial distributed with parameters N and $p = 0.5$. Based on this, the critical values can be obtained. Model A is considered statistically equivalent with model B if B is not significantly different from the expected value $Np = \frac{N}{2}$. Both test statistics from equations (12) and (13) can be corrected for the number of parameters used in the models by using AIC.

Table 4 and Table 5 show the goodness-of-fit test results of bivariate copula modelling. 11 copula families are chosen which include Gaussian, Student-t, Clayton, Gumbel, Frank, BB1, BB7, and the survival copulas of the Clayton (s.Clayton), Gumbel (s.Gumbel), BB1 (s.BB1) and BB7 (s.BB7)[2] in both tests. In these candidates, families represent various strengths of tail behaviour. For instance, Frank copulas show tail independence which is also considered as a benchmark for tail dependence, Gumbel copulas show only upper tail dependence while Clayton copulas show only lower tail dependence. Student-t copulas show reflection symmetric upper and lower tail dependence and BB families show different upper and lower tail dependence. From the results of the Vuong test, student-t copula family fits 53 out of 55 pairs best in all 11 copula families, although t copula of three pairs (NET.SWE, SPA.SWE, NET.DEN) share the highest score

[2]In terms of bivariate copula families and their functions and properties, please see Appendix I.

with both survival form of BB1 and survival form of BB7 families. Additionally, both survival form of BB1 and survival form of BB7 families indicate asymmetric upper and lower tail dependence. GRE.SWE and POR.SWE are modeled best by Frank copula, which shows no tail dependence of the pairs, according to Vuong test. On the other hand, the Clarke test shows that student-t copula family fits all 55 pairs better than the others, which means these pairs tend to have symmetric upper and lower tail dependence.

Table 4: Bivariate goodness-of-fit Vuong test

Pairs	Gaussian	t	Clayton	Gumbel	Frank	BB1	BB7	s.Clayton	s.Gumbel	s.BB1	s.BB7
BEL.DEN	-6	**10**	-3	-5	-7	4	4	-9	4	4	4
BEL.FRA	-7	**10**	-8	-2	-3	6	2	-9	2	7	2
BEL.GRE	-7	**10**	-7	0	-7	4	6	-6	-2	4	5
BEL.IRE	-8	**10**	-7	-1	-5	4	4	-8	1	5	5
BEL.ITA	-6	**10**	-9	-1	-6	5	5	-7	-1	5	5
BEL.NET	-7	**10**	-8	0	2	5	0	-9	1	6	0
BEL.POR	-8	**10**	-8	-3	-3	4	2	-8	3	7	4
BEL.SPA	-8	**10**	-8	0	-4	5	4	-8	-1	5	5
BEL.SWE	-2	**7**	-2	-1	0	-1	-1	-1	1	1	-1
BEL.UK	-5	**8**	-5	-4	0	4	2	-10	3	5	2
DEN.FRA	-6	**10**	-6	-3	-7	4	4	-7	3	4	4
DEN.GRE	-6	**10**	-7	-5	3	3	-1	-9	4	6	2
DEN.IRE	-6	**10**	-6	-1	-4	4	2	-8	3	3	3
DEN.ITA	-7	**10**	-6	-3	-6	4	4	-7	3	4	4
DEN.NET	-4	**6**	-6	-3	-8	5	5	-8	1	**6**	**6**
DEN.POR	-7	**10**	-7	-3	3	4	0	-8	2	5	1
DEN.SPA	-6	**10**	-6	-4	-6	4	4	-7	3	4	4
DEN.SWE	-4	**10**	-9	-3	-4	6	2	-9	1	7	3
DEN.UK	-5	**10**	-8	0	-2	4	2	-6	-1	4	2
FRA.GRE	-8	**10**	-8	3	-4	3	3	-6	-1	4	4
FRA.IRE	-8	**10**	-8	0	2	4	1	-8	1	5	1
FRA.ITA	-6	**10**	-9	-1	-5	5	5	-8	-1	5	5
FRA.NET	-7	**10**	-7	-2	-1	6	1	-10	3	6	1
FRA.POR	-7	**10**	-7	-2	2	5	1	-10	2	5	1
FRA.SPA	-8	**10**	-8	0	0	7	0	-8	0	7	0
FRA.SWE	-6	**9**	-3	-1	-1	0	-1	-3	2	3	1
FRA.UK	-5	**8**	-6	-4	-5	4	4	-9	4	5	4
GRE.IRE	-7	**10**	-7	0	-7	4	6	-7	-1	3	6

Continued on next page

Table 4 –continued from previous page

Pairs	Gaussian	t	Clayton	Gumbel	Frank	BB1	BB7	s.Clayton	s.Gumbel	s.BB1	s.BB7
GRE.ITA	-8	**10**	-8	0	-5	5	3	-7	0	5	5
GRE.NET	-6	**10**	-6	-1	-4	3	2	-6	0	5	3
GRE.POR	-7	**10**	-10	4	-5	4	4	-5	-3	4	4
GRE.SPA	-7	**10**	-7	0	-7	4	6	-6	-2	4	5
GRE.SWE	-1	7	-6	-6	**8**	1	-4	-9	3	7	0
GRE.UK	-3	**9**	-5	0	0	3	-1	-6	0	3	0
IRE.ITA	-7	**10**	-7	-1	-7	4	4	-7	1	5	5
IRE.NET	-8	**10**	-8	1	2	5	1	-8	1	4	0
IRE.POR	-7	**10**	-7	-1	-7	5	5	-7	-1	5	5
IRE.SPA	-8	**10**	-7	-1	-5	4	4	-8	1	5	5
IRE.SWE	0	**7**	-1	-3	2	-1	-1	-3	-1	3	-2
IRE.UK	-6	**10**	-5	0	2	3	0	-8	2	3	-1
ITA.NET	-8	**10**	-8	1	0	6	1	-8	1	5	0
ITA.POR	-8	**10**	-8	-2	-3	5	2	-8	2	7	3
ITA.SPA	-6	**10**	-9	-1	-6	5	6	-7	-1	5	4
ITA.SWE	-3	**10**	-7	0	0	2	-1	-6	0	3	2
ITA.UK	-6	**10**	-6	0	-6	3	3	-7	3	3	3
NET.POR	-7	**10**	-8	-1	3	5	0	-9	1	6	0
NET.SPA	-7	**10**	-8	-1	3	5	0	-9	2	5	0
NET.SWE	-6	**7**	-4	-5	-6	1	1	-6	4	**7**	**7**
NET.UK	-5	**10**	-6	-4	-6	4	4	-8	3	4	4
POR.SPA	-9	**10**	-7	-2	-3	5	2	-8	2	7	3
POR.SWE	2	6	-4	-4	**10**	0	-6	-5	1	3	-3
POR.UK	-4	**8**	-4	-4	7	3	-4	-10	6	5	-3
SPA.SWE	1	**2**	0	-3	1	0	0	-5	0	**2**	**2**
SPA.UK	-5	**8**	-5	-5	-1	4	2	-10	4	6	2
SWE.UK	-4	**7**	-6	0	-3	4	3	-6	0	4	1

Note: There is no order in the pair names. Bold format indicates the best candidate.

Table 5: Bivariate goodness-of-fit Clarke test

Pairs	Gaussian	t	Clayton	Gumbel	Frank	BB1	BB7	s.Clayton	s.Gumbel	s.BB1	s.BB7
BEL.DEN	-9	**10**	-3	-3	-6	3	6	-9	2	1	8
BEL.FRA	-6	**10**	-8	-2	8	6	-3	-10	2	4	-1
BEL.GRE	-9	**10**	-8	0	-4	3	7	-7	-2	5	5
BEL.IRE	-9	**10**	-6	-4	6	5	2	-9	0	3	2
BEL.ITA	-6	**10**	-10	-1	8	5	-1	-8	-1	5	-1
BEL.NET	-6	**10**	-8	0	8	6	-3	-10	2	4	-3
BEL.POR	-9	**10**	-6	-4	8	6	0	-9	0	4	0
BEL.SPA	-6	**10**	-9	0	4	7	-1	-9	-1	7	-2
BEL.SWE	-10	**10**	-5	-1	-6	3	1	-6	6	2	6
BEL.UK	-8	**10**	-5	-4	-3	5	4	-10	3	3	5
DEN.FRA	-9	**10**	-5	-2	-5	3	6	-9	0	4	7
DEN.GRE	-8	**10**	-6	-3	4	3	-2	-10	5	4	3
DEN.IRE	-9	**10**	-6	-2	-5	4	5	-8	2	3	6
DEN.ITA	-9	**10**	-6	-2	-5	5	5	-8	0	3	7
DEN.NET	-8	**10**	-5	-2	-5	3	7	-10	0	4	6
DEN.POR	-9	**10**	-6	-3	4	2	-2	-9	5	5	3
DEN.SPA	-9	**10**	-4	-2	-7	3	6	-8	2	1	8
DEN.SWE	-6	**10**	-8	-3	8	6	1	-10	0	3	-1
DEN.UK	-8	**10**	-8	0	-3	4	4	-8	-1	6	4
FRA.GRE	-10	**10**	-7	-1	-4	4	6	-7	-1	6	4
FRA.IRE	-9	**10**	-6	-3	8	4	0	-9	1	2	2
FRA.ITA	-6	**10**	-9	-1	8	3	1	-9	-1	4	0
FRA.NET	-6	**10**	-8	-1	8	5	-2	-10	2	5	-3
FRA.POR	-8	**10**	-6	-3	8	4	-3	-10	2	4	2
FRA.SPA	-6	**10**	-8	-1	8	5	-3	-10	2	5	-2
FRA.SWE	-10	**10**	-6	-1	-6	2	3	-6	5	2	7
FRA.UK	-8	**10**	-6	-3	-3	6	4	-10	1	4	5
GRE.IRE	-10	**10**	-8	0	-4	2	7	-6	-2	5	6

Continued on next page

Table 5 –continued from previous page

Pairs	Gaussian	t	Clayton	Gumbel	Frank	BB1	BB7	s.Clayton	s.Gumbel	s.BB1	s.BB7
GRE.ITA	-10	**10**	-8	-2	7	3	2	-6	-1	4	1
GRE.NET	-10	**10**	-7	-1	-4	6	5	-7	-1	4	5
GRE.POR	-9	**10**	-9	0	7	4	2	-6	-4	3	2
GRE.SPA	-9	**10**	-9	0	-4	4	6	-6	-2	6	4
GRE.SWE	-7	**10**	-7	-3	7	2	-3	-10	4	5	2
GRE.UK	-9	**10**	-7	-1	-4	4	2	-8	6	3	4
IRE.ITA	-9	**10**	-6	-4	1	6	3	-9	0	4	4
IRE.NET	-9	**10**	-7	-1	8	6	0	-8	0	2	-1
IRE.POR	-10	**10**	-7	-2	1	5	3	-7	-1	5	3
IRE.SPA	-10	**10**	-6	-4	6	5	2	-8	0	3	2
IRE.SWE	-9	**10**	-6	-1	-4	2	0	-7	4	6	5
IRE.UK	-9	**10**	-6	-3	-1	3	1	-9	7	3	4
ITA.NET	-7	**10**	-7	-2	8	5	-1	-10	0	4	0
ITA.POR	-9	**10**	-6	-2	8	6	-2	-9	0	4	0
ITA.SPA	-6	**10**	-10	-1	5	7	-1	-8	-2	6	0
ITA.SWE	-8	**10**	-8	-2	-4	6	4	-8	2	4	4
ITA.UK	-9	**10**	-5	-2	-5	3	6	-9	3	2	6
NET.POR	-8	**10**	-6	-2	8	4	-4	-10	2	5	1
NET.SPA	-8	**10**	-6	-1	8	5	-4	-10	2	5	-1
NET.SWE	-10	**10**	-4	-3	-7	1	5	-6	2	4	8
NET.UK	-8	**10**	-6	-2	-4	5	5	-10	0	3	7
POR.SPA	-10	**10**	-6	-3	8	6	-1	-8	0	4	0
POR.SWE	-7	**9**	-8	-2	7	2	-3	-9	6	4	1
POR.UK	-8	**10**	-6	-3	6	1	-3	-10	6	5	2
SPA.SWE	-9	**10**	-5	-2	-6	4	0	-7	4	5	6
SPA.UK	-8	**10**	-6	-3	-3	5	4	-10	2	4	5
SWE.UK	-8	**10**	-8	-1	-4	4	6	-8	0	5	4

Note: There is no order in the pair names. Bold format indicates the best candidate.

5. Vine Copula approach

5.1. Introduction of Vine Copulas

In order to improve the copula method with regard to a wider range of dependence, a more flexible range of upper and lower tail dependence, a larger dimension and a computationally feasible density for estimation, vine copulas became a handy copula technique.

A d-variate copula $C(u_1, \ldots, u_d)$ is a cumulative distribution function (cdf) with uniform marginals on the unit interval. According to Sklar (1959), if $F_j(x_j)$ is the cdf of a univariate continuous random variable X_j, then $C(F_1(x_1), \ldots, F_d(x_d))$ is a d-variate distribution for $X = (X_1, \ldots, X_d)$ with marginal distributions $F_j, j = 1, \ldots, d$. Conversely, if $F_j, j = 1, \ldots, d$ is continuous, then there exists a unique copula C as

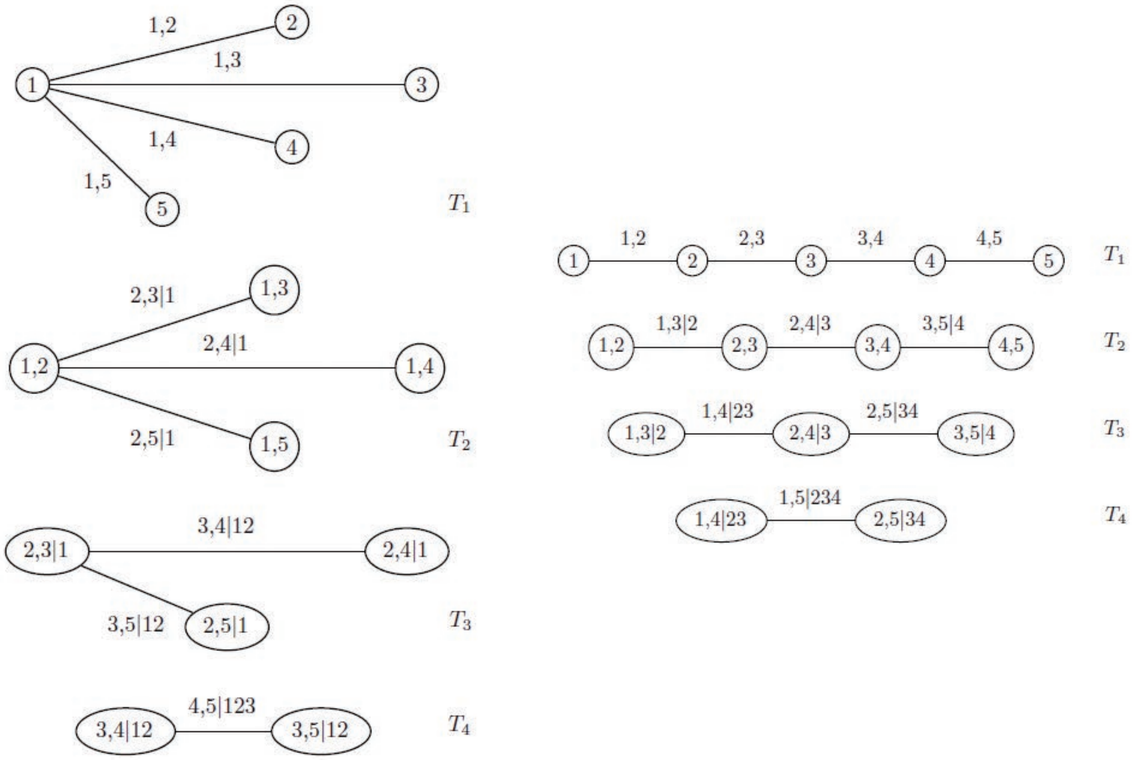
$$F(x) = C(F_1(x_1), \ldots, F_d(x_d)), \forall x = (x_1, \ldots, x_d), \tag{14}$$

which is called the theorem of Sklar (1959). While a d-dimensional vine copula are built by $d(d-1)$ bivariate copulas in a $(d-1)$-level tree form. There are different ways to construct a copula tree. C-vines and D-vines are the selected tree types in this paper. In a C-vine tree, the dependence with respect to one particular variable, called first root node, is modeled by bivariate copulas for each pair. Conditioned on this variable, pair wise dependencies with respect to a second variable are modeled, which is called the second root node. In general, a root node is chosen in each tree and all pairwise dependencies with respect to this node are modeled conditioned on all previous root nodes (see Figure 1 left panel). According to Aas et al. (2009), this gives the following decomposition of a multivariate density,

$$f(x) = \prod_{k=1}^{d} f_k(x_k) \times \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{i,i+j|1:(i-1)}(F(x_i|x_1, \ldots, x_{i-1}), (F(x_{i+j}|x_1, \ldots, x_{i-1})|\boldsymbol{\theta}_{i,i+j|1:(i-1)}), \tag{15}$$

where $f_k, k = 1, \ldots, d$, denote the marginal densities and $c_{i,i+j|1:(i-1)}$ bivariate copula densities with parameter(s) $\boldsymbol{\theta}_{i,i+j|1:(i-1)}$ (here $i_k : i_m$ means $i_k, \ldots, i_m$). And the outer product runs over the $d-1$ trees and root nodes i, while the inner product refers to the $d-i$ pair-copulas in each tree $i = 1, \ldots, d-1$.

A D-vine chooses the order of these pairs in a different way (see Figure 1 right panel). In the first level of the tree, the dependence of the first and second variable, the second and the third, the third and the fourth, and so



Source: Brechmann and Schepsmeier (2012)

Figure 1: Examples of 5-dimensional C- (left) and D-vine (right)

on, are used. That means in a 5-dimensional vine copula, in the first level of the tree, pairs $(1,2),(2,3),(3,4),(4,5)$ have been modeled. While in the second level of the tree, conditional dependence of the first and third given the second variable (pair $(1,3|2)$), the second and fourth given the third (pair $(2,4|3)$), and so on. In this way it continues to construct the third level up to the $d-1$ level. According to Aas et al. (2009) the density of a D-vine is,

$$f(x) = \prod_{k=1}^{d} f_k(x_k) \times$$
$$\prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{j,i+i|(j+1):(j+i-1)}(F(x_j|x_{j+1},\ldots,x_{j+i-1}),(F(x_{j+i}|x_{j+1},\ldots,x_{j+i-1})|\theta_{j,j+i|x_{j+1},\ldots,x_{j+i-1}}), \tag{16}$$

where the outer product runs over the $d-1$ trees, while the pairs in each tree are designated by the inner product. In order to get the conditional distribution functions $F(x|\mathbf{v})$ for an m-dimensional vector $\mathbf{v}$, one can sequentially apply the following relationship,

$$h(x|\mathbf{v},\boldsymbol{\theta}) := F(x|\mathbf{v}) = \frac{\partial C_{xv_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})|\boldsymbol{\theta})}{\partial F(v_j|\mathbf{v}_{-j}}, \tag{17}$$

where v_j is an arbitrary component of $\mathbf{v}$ and $\mathbf{v}_{-j}$ denotes the $(m-1)$-dimensional vector $\mathbf{v}$ excluding v_j. Further $C_{xv_j|\mathbf{v}_{-j}}$ is a bivariate copula distribution function with parameter(s) θ specified in tree m.

5.2. Vine copula estimation

Vine copulas can be constructed by the bivariate copulas estimated in section 4.2. Two types of vine are chosen to be estimated, C-vine and D-vine. Then one will choose the better vine based on their value of log-likelihood.

First, a C-vine has been conducted. In order to achieve the best performance of the C-vine, $d-1$ pairs of countries should be carefully chosen. According to Aas and Berg (2009), empirical rules can be applied to select to vine order.

1. Select the first root node that has strong dependence with all other variables;
2. List the most dependent variables with the first root node as decreasing in dependence order;
3. List the least dependent variables with the first root node as increasing in dependence order;
4. Sequentially list the least dependent variable with the previous selected.

Table 6 shows the dependence of pairs according to Kendall's τ. Kendall's τ is a rank correlation coefficient which developed by Kendall (1938). It is calculated as follows. Let (x_1, y_1), (x_2, y_2), …, (x_n, y_n) be a set of observations of the joint random variables X and Y respectively, such that all the values of (x_i) and (y_i) are unique. Any pair of observations (x_i, y_i) and (x_j, y_j) are said to be concordant if the ranks for both elements agree: that is, if both $x_i > x_j$ and $y_i > y_j$ or if both $x_i < x_j$ and $y_i < y_j$. They are said to be discordant, if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, the pair is neither concordant nor discordant.

$$\tau = \frac{(\text{number of concordant pair}) - (\text{number of discordant pair})}{\frac{1}{2}n(n-1)}$$

The first root node should have strong dependence with all other variables.

Table 6: Country-pair dependence based on Kendall's τ

Pair	τ	Pair	τ	Pair	τ
BEL.DEN	0.113469	FRA.GRE	0.165981	IRE.SWE	0.08461
BEL.FRA	0.526231	FRA.IRE	0.250222	IRE.UK	0.122281
BEL.GRE	0.193722	FRA.ITA	0.337805	ITA.NET	0.313732
BEL.IRE	0.29413	FRA.NET	0.56691	ITA.POR	0.358453
BEL.ITA	0.412142	FRA.POR	0.310606	ITA.SPA	0.507106
BEL.NET	0.523141	FRA.SPA	0.461769	ITA.SWE	0.126307
BEL.POR	0.356663	FRA.SWE	0.066616	ITA.UK	0.107109
BEL.SPA	0.537487	FRA.UK	0.154579	NET.POR	0.299719
BEL.SWE	0.058031	GRE.IRE	0.202541	NET.SPA	0.443678
BEL.UK	0.151617	GRE.ITA	0.299755	NET.SWE	0.071075
DEN.FRA	0.13127	GRE.NET	0.13684	NET.UK	0.160818
DEN.GRE	0.201002	GRE.POR	0.297087	POR.SPA	0.384536
DEN.IRE	0.112917	GRE.SPA	0.228495	POR.SWE	0.12385
DEN.ITA	0.123432	GRE.SWE	0.198992	POR.UK	0.193094
DEN.NET	0.16681	GRE.UK	0.114702	SPA.SWE	0.082487
DEN.POR	0.179232	IRE.ITA	0.274777	SPA.UK	0.139775
DEN.SPA	0.103206	IRE.NET	0.255291	SWE.UK	0.127193
DEN.SWE	0.348671	IRE.POR	0.337959		
DEN.UK	0.180609	IRE.SPA	0.314839		

Note: There are no order in the pair names.

In this case, Ireland shows the strongest dependence with others. Applying the rest of the rules, the order of the C-vine is chosen as SPA, BEL, DEN, FRA, GRE, IRE, ITA, NET, POR, SWE, UK. The estimated dependence parameters are shown in Table 7. The log-likelihood function of the C-vine copula with parameter $\boldsymbol{\theta}_{CV}$ is as follows:

$$\ell_{CV}(\boldsymbol{\theta}_{CV}|\boldsymbol{u}) = \sum_{k=1}^{N}\sum_{i=1}^{d-1}\sum_{j=1}^{d-i}\log[c_{i,i+j|1:(i-1)}(F_{i|1:(i-1)}, F_{i+j|1:(i-1)}|\boldsymbol{\theta}_{i,i+j|1:(i-1)})], \tag{18}$$

where $F_{j|i_1:i_m} := F(u_{k,j}|u_{k,i_1}, \cdots, u_{k,i_m})$ and the marginal distribution are uniform.

Table 7: Estimated C-vine copula parameters (Log-likelihood = 12934.83)

	Level-1									
margin*:	91	92	93	94	95	96	97	98	9a	9b
family:	t	t	t	t	t	t	t	t	t	t
$\hat{\theta}_1$	0.785636	0.162843	0.692354	0.343322	0.508532	0.757037	0.65892	0.616534	0.124844	0.215362
$\hat{\theta}_2$	2.0001	4.599867	2.0001	2.126921	2.0001	2.0001	2.03829	2.0001	15.77393	7.204589
	Level-2									
margin:	12|9	13|9	14|9	15|9	16|9	17|9	18|9	1a|9	1b|9	
family:	t	t	t	t	t	t	t	t	t	
$\hat{\theta}_1$	0.078301	0.494936	0.045666	0.168873	0.181871	0.525246	0.206182	0.00662	0.109588	
$\hat{\theta}_2$	8.578611	3.110824	4.048904	3.658444	3.96442	3.452198	3.831343	16.10222	9.224822	
	Level-3									
margin:	23|19	24|19	25|19	26|19	27|19	28|19	2a|19	2b|19		
family:	t	t	t	t	t	t	t	t		
$\hat{\theta}_1$	0.079817	0.28738	0.112875	0.088872	0.129475	0.208854	0.525698	0.242231		
$\hat{\theta}_2$	11.87698	8.23072	11.3741	15.98879	11.707	9.281302	3.867485	7.188617		
	Level-4									
margin:	34|129	35|129	36|129	37|129	38|129	3a|129	3b|129			
family:	t	t	t	t	t	90.Clayton	Frank			
$\hat{\theta}_1$	-0.01066	0.021689	-0.00146	0.507135	0.025434	-0.01267	0.49029			
$\hat{\theta}_2$	8.751509	11.44785	8.198117	4.60616	8.407895	0	0			
	Level-5									
margin:	45|1239	46|1239	47|1239	48|1239	4a|1239	4b|1239				
family:	t	t	Frank	t	Frank	t				
$\hat{\theta}_1$	0.173071	0.234101	-0.41015	0.27752	1.047847	0.047483				
$\hat{\theta}_2$	6.216486	8.23026	0	4.326902	0	16.16467				
	Level-6					Level-7				
margin:	56|12349	57|12349	58|12349	5a|12349	5b|12349	67|123459	68|123459	6a|123459	6b|123459	
family:	Frank	t	t	270.Joe	t	t	t	t	t	
$\hat{\theta}_1$	0.553639	0.023637	0.283618	-1.01931	0.04726	-0.02059	0.117399	0.094077	-0.02442	
$\hat{\theta}_2$	0	23.86815	7.225896	0	20.07778	16.18606	12.6435	16.46889	16.44031	
	Level-8 (|9123456)			Level-9 (|91234567)		Level-10 (|912345678)				
margin:	78	7a	7b	8a	8b	ab				
family:	Joe	t	t	270.Clayton	Frank	Gaussian				
$\hat{\theta}_1$	1.003137	-0.00739	0.046393	-0.0234	0.807638	0.072738				
$\hat{\theta}_2$	0	25.25985	13.24346	0	0	0				

Note(*): It shows the bivariate margin under condition, and 1=Belgium, 2=Denmark, 3=France, 4=Greece, 5=Ireland, 6=Italy, 7=Netherlands, 8=Portugal, 9=Spain, a=Sweden and b=UK.

In the case of D-vine, the empirical rule for first tree selection chooses an order of the variables that intends to capture as much dependence as possible. According to Belgorodski (2010), it is equivalent to solving the Traveling Salesman Problem (TSP). The TSP can be solved using the Cheapest Insertion Algorithm (Rosenkrantz et al., 1977). The log-likelihood function of a D-vine copula with parameter $\boldsymbol{\theta}_{DV}$ is as follows:

$$\ell_{DV}(\boldsymbol{\theta}_{DV}|\boldsymbol{u}) = \sum_{k=1}^{N}\sum_{i=1}^{d-1}\sum_{j=1}^{d-i} \log[c_{j,j+i|(j+1):(j+i-1)}(F_{j|(j+1):(j+i-1)}, F_{j+i|(j+1):(j+i-1)}|\boldsymbol{\theta}_{j,j+i|(j+1):(j+i-1)})]. \tag{19}$$

Using information from Table 6 with the algorithm, the order of D-vine is chosen as IRE, POR, GRE, ITA, SPA, BEL, FRA, NET, UK, DEN, SWE. Table 8 shows the estimated dependence parameters.

Table 8: Estimated D-vine copula parameters (Log-likelihood = 12805.13)

	Level-1									
margin*:	58	84	46	69	91	13	37	7b	b2	2a
family:	t	t	t	t	t	t	t	t	t	t
$\hat{\theta}_1$	0.554977	0.474093	0.458832	0.757037	0.785636	0.75319	0.793209	0.252631	0.279172	0.519913
$\hat{\theta}_2$	2.0001	2.0001	2.139375	2.0001	2.0001	2.0001	2.0001	5.253718	5.590975	3.629961
	Level-2									
margin:	54|8	86|4	49|6	61|9	93|1	17|3	3b|7	72|b	ba|2	
family:	t	t	t	t	t	t	t	t	t	
$\hat{\theta}_1$	0.13404	0.442381	0.09262	0.181871	0.266287	0.411756	0.074764	0.208051	0.06887	
$\hat{\theta}_2$	5.023617	2.842461	6.450827	3.96442	3.266378	3.317322	22.79576	7.173716	14.1263	
	Level-3									
margin:	56|48	89|46	41|69	63|19	97|13	1b|37	32|b7	7a|2b		
family:	t	t	t	t	t	BB8	t	t		
$\hat{\theta}_1$	0.17359	0.347971	0.014044	0.007594	0.083558	1.139907	0.002258	-0.04332		
$\hat{\theta}_2$	5.093834	4.668656	5.872747	8.001854	6.340184	0.859126	14.36096	10.73387		
	Level-4									
margin:	59|648	81|469	43|169	67|913	9b|137	12|37b	3a|2b7			
family:	t	t	t	t	t	t	Frank			
$\hat{\theta}_1$	0.184206	0.177932	0.007428	-0.02296	0.040587	0.000576	0.116236			
$\hat{\theta}_2$	6.146175	7.053155	10.4331	15.10297	17.18287	11.79698	0			
	Level-5									
margin:	51|8469	83|1469	47|1369	6b|7913	92|b137	1a|237b				
family:	t	t	Frank	t	t	270.Clayton				
$\hat{\theta}_1$	0.08633	0.056068	-0.24771	0.007181	0.021326	-0.01147				
$\hat{\theta}_2$	7.543744	12.72372	0	13.75464	28.68429	0				
	Level-6					Level-7				
margin:	53|18469	87|13469	4b|71369	62|b7913	9a|2b137	57|318469	8b|713469	42|b71369	6a|2b7913	
family:	Frank	s.BB8	t	t	BB8	t	Frank	s.Gumbel	t	
$\hat{\theta}_1$	0.169382	1.038112	0.104391	0.09152	1.098464	0.020014	0.974141	1.181762	0.1298	
$\hat{\theta}_2$	0	0.985067	14.97414	24.76819	0.929845	14.60283	0	0	20.13141	
	Level-8			Level-9		Level-10				
margin:	5b	82	4a	52	8a	5a				
cond.	7318469	b713469	2b71369	b7318469	2b713469	2b7318469				
family:	t	BB8	Frank	t	Gaussian	270.Joe				
$\hat{\theta}_1$	0.009865	1.263496	0.899777	-0.00524	-0.03135	-1.02386				
$\hat{\theta}_2$	22.60227	0.845965	0	22.41893	0	0				

Note(*): It shows the bivariate margin under condition, and 1=Belgium, 2=Denmark, 3=France, 4=Greece, 5=Ireland, 6=Italy, 7=Netherlands, 8=Portugal, 9=Spain, a=Sweden and b=UK.

According to the results from Table 7 and 8, the log-likelihood of C-vine is 12934.83, while the log-likelihood of D-vine is 12805.13. Therefore, C-vine is superior to D-vine.

6. Simulation

In this paper, we intend to forecast the probabilities of sovereign crisis in these 11 countries in the future year. Therefore, the sovereign spreads of each country for the next 365 days need to be generated. 365 groups of error terms based on the C-vine copula parameters are simulated. 365 is the forecast horizon of this research and it can be changed depending on the purpose of forecast. We apply these group of error terms back into the GARCH filters estimated in Section 4.1 to get the next 365 days' sovereign spreads movements of each country. Future sovereign spreads can be calculated by adding spreads movement to the spreads of previous day from 12/03/2012 which is the last day in the sample. We apply the definition of sovereign crisis is stated in Section 1, which is that sovereign spread against Germany is more than 1000 basis points. Therefore, if one or more of these simulated spreads are greater than 10%, the sovereign crisis in the following year will be counted. This process is repeated 10000 times, and the times with sovereign crisis divided by 10000 will be the probabilities of sovereign debt crisis. The relationship can be represented by the expression as follows:

$$k_i = \begin{cases} 1 & \text{if there is at least one crisis event in future } h\text{-day simulation} \\ 0 & \text{if there is no event in future } h\text{-day simulation} \end{cases}$$

The probability of the soveriegn debt crisis is expressed as

$$Pr = \frac{\sum_{i=1}^{N} k_i}{N},$$

where k is a dummy in order to identify whether there will be one or more crisis in the forecasting horizon, h is the forecast horizon (365 days) , i is the ith simulation, Pr is the probability of sovereign crisis in target country and N is the total number of simulations (10000).

Table 9: Probability of sovereign debt crisis in next 365 days

Countries	BEL	DEN	FRA	GRE	IRE	ITA	NET	POR	SPA	SWE	UK
Probability	60.14%	8.74%	62.13%	100%	71.60%	81.08%	25.86%	99.77%	87.17%	5.45%	12.87%

Table 9 presents the results of the estimated probabilities of sovereign crisis in the next 365 days. According to the results, Greece has the highest probability which is 100% and is followed by Portugal (99.77%), which are consistent with the fact that they are already in crisis. Spain (87.17%) and Italy (81.08%) have extremely high probabilities of entering crisis. Ireland has a 71.06% chance of entering crisis. The probability of crisis in France and Belgium are 62.13% and 60.14% which are fairly high, and for France it is higher than expected. Netherlands (25.86%) shows a fairly low probability of crisis, the most stable in the euro area. The probability of countries outside the euro area such as the UK (12.87%), Denmark(8.74%) and Sweden(5.45%) are very low which reveals the stability of sovereign debt in these countries.

7. Conclusion

This paper provides a method to calculate the probability of sovereign debt crisis which is an infrequent event. The sovereign spreads against Germany are simulated and the dependence of those time series is considered by applying vine copula models in the mean time. It is extremely useful in assessing the risk level of sovereign debt crisis in the European Union. We examined 11 countries in the European Union. Results show that Greece and Portugal have an extremely high probability of sovereign debt crisis. Spain and Italy are potentially the next victims of sovereign debt crisis. Unexpectedly, France and Belgium show a fairly high risk level. Netherlands enjoys the lowest probability of crisis in the euro area in the sample. The UK, Denmark and Sweden show strong stability of their sovereign debt and being outside the euro area might be the reason for this. According to the results, the probability calculated in this paper appears to be a very good indicator of sovereign debt default risk level. In addition, it is a better indicator than sovereign credit default swap (CDS), because sovereign CDS is an over the counter (OTC) traded financial instrument, which makes tracking all the trades difficult to achieve. This indicator can make a contribution to alerting the European Central Bank (ECB) or governments of those countries in the European Union, as well as ranking the risk level of each government bond in the European Union for investors.

Aas, K. and Berg, D. (2009), 'Models for construction of multivariate dependence - a comparison study', *European Journal of Finance* **15**(7-8), 639–659.

Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009), 'Pair-copula constructions of multiple dependence', *Insurance: Mathematics and Economics* **44**(2), 182 – 198.

Akaike, H. (1974), 'A new look at the statistical model identification', *Automatic Control, IEEE Transactions on* **19**(6), 716–723.

Bedford, T. and Cooke, R. M. (2002), 'Vines: A new graphical model for dependent random variables', *The Annals of Statistics* **30**(4), 1031–1068.

Belgorodski, N. (2010), Selecting pair-copula families for regular vines with application to the multivariate analysis of european stock market indices., Master's thesis, Technische Universität München.

Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli, T. (2000), 'Copulas for finance - a reading guide and some applications', *SSRN* .

Box, G. E. P. and Jenkins, G. M. (1970), *Time Series Analysis: Forecasting and Control*, Holden-Day, Inc, San Francisco.

Brechmann, E. C. and Schepsmeier, U. (2012), Modeling dependence with c- and d-vine copulas: The r-package cdvine, Reference manual, R-CRAN.
URL: *http://cran.r-project.org/web/packages/CDVine/vignettes/ CDVine-package.pdf*

Bruun, J. T. and Tawn, J. A. (1998), 'Comparison of approaches for estimating the probability of coastal flooding', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **47**(3), 405–423.

Cherubini, U., Luciano, E. and Vecchiato, W. (2004), *Copula methods in Finance*, John Wiley & Son Ltd, Chichester.

Clarke, K. A. (2007), 'A simple distribution-free test for nonnested model selection', *Political Analysis* **15**(3), 347–363.

de Haan, L. and de Ronde, J. (1998), 'Sea and wind: Multivariate extremes at work', *Extremes* **1**(1), 7–45.

de Haan, L. and Ferreira, A. (2006), *Extreme Value Theory: An Introduction*, Springer-Verlag, New York.

de Haan, L. and Resnick, S. (1977), 'Limit theory for multivariate sample extremes', *Z.Wahrscheinlichkeitstheorie und Verw. Gebiete* **40**(4), 317–337.

Demarta, S. and McNeil, A. J. (2005), 'The t copula and related copulas', *International Statistical Review* **73**(1), 111–129.

Dötz, N. and Fischer, C. (2010), What can EMU countries' sovereign bond spreads tell us about market perceptions of default probabilities during the recent financial crisis?, Discussion paper series 1: Economic studies, Deutsche Bundesbank, Research Centre.

Eichengreen, B., Hausmann, R. and Panizza, U. (2003), Currency mismatches, debt intolerance and original sin: Why they are not the same and why it matters, Technical report, National Bureau of Economic Research, Inc.

Embrechts, P., Lindskog, F. and McNeil, A. (2003), 8 modelling dependence with copulas and applications to risk management, *in* S. T. Rachev, ed., 'Handbook of Heavy Tailed Distribution in Finance', Elsevier, pp. 329–384.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation', *Econometrica* **50**(4), 987–1007.

Fougères, A.-L. (2003), Multivariate extremes, *in* B. Finkenstädt and H. Rootzén, eds, 'Extreme Values in Finance, Telecommunications, and theEnvironment', Chapman and Hall/CRC.

Frees, E. W., Carriere, J. and Valdez, E. (1996), 'Annuity valuation with dependent mortality', *The Journal of Risk and Insurance* **63**(2), pp. 229–261.

Genest, C. and Favre, A. (2007), 'Everything you always wanted to know about copula modeling but were afraid to ask', *Journal of Hydrologic Engineering* **12**(4), 347–368.

Goldstein, M. and Turner, P. (2004), *Controlling Currency Mismatches in Emerging Markets*, Peterson Institute for International Economics, Washington, DC.

Jarrow, R. A. and Turnbull, S. M. (1995), 'Pricing derivatives on financial securities subject to credit risk', *Journal of Finance* **50**(1), 53–85.

Joe, H. (1993), 'Parametric families of multivariate distributions with given margins', *Journal of Multivariate Analysis* **46**(2), 262–282.

Joe, H. (1996), 'Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters', *Lecture Notes-Monograph Series* **28**, pp. 120–141.

Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.

Joe, H. and Hu, T. (1996), 'Multivariate distributions from mixtures of max-infinitely divisible distributions', *Journal of Multivariate Analysis* **57**(2), 240–265.

Jones, S. (2009), 'The Formula That Felled Wall St'.
URL: *http://www.ft.com/cms/s/0/912d85e8-2d75-11de-9eba- 00144feabdc0.html*

Kendall, M. G. (1938), 'A new measure of rank correlation', *Biometrika* **30**(1-2), 81–93.

Ljung, G. M. and Box, G. E. P. (1978), 'On a measure of lack of fit in time series models', *Biometrika* **65**(2), 297–303.

McNeil, A. J. and Neslehova, J. (2009), 'Multivariate archimedean copulas, d-monotone functions and l_1-norm symmetric distributions', *The Annals of Statistics* **37**, 3059–3097.

Merton, R. C. (1974), 'On the pricing of corporate debt: The risk structure of interest rates', *Journal of Finance* **29**(2), 449–70.

Nikoloulopoulos, A. K., Joe, H. and Li, H. (2012), 'Vine copulas with asymmetric tail dependence and applications to financial return data', *Computational Statistics & Data Analysis* .

Patton, A. J. (2008), Copula-based models for financial time series, Ofrc working papers series, Oxford Financial Research Centre.

Reinhart, C. M., Rogoff, K. S. and Savastano, M. A. (2003), 'Debt intolerance', *Brookings Papers on Economic Activity* **34**(1), 1–74.

Rodriguez, M. J. and Ruiz, E. (2012), 'Revisiting several popular garch models with leverage effect: Differences and similarities', *Journal of Financial Econometrics* **10**(4), 637–668.

Rosenkrantz, D., Stearns, R. and Lewis, II, P. (1977), 'An analysis of several heuristics for the traveling salesman problem', *SIAM Journal on Computing* **6**(3), 563–581.

Salmon, F. (2009), 'Recipe for Disaster: The Formula That Killed Wall Street'.
URL: *http://www.wired.com/ print/techbiz/it/magazine/17-03/wp_quant*

Sklar, A. (1959), 'Fonctions de répartition á n dimensions et leurs marges.', *Publications de l'Institut de Statistique de l'Université de Paris,* **8**, 229–231.

Sy, A. N. (2004), 'Rating the rating agencies: Anticipating currency crises or debt crises?', *Journal of Banking & Finance* **28**(11), 2845–2867.

Vuong, Q. H. (1989), 'Likelihood ratio tests for model selection and non-nested hypotheses', *Econometrica* **57**(2), pp. 307–333.

Wang, W. and Wells, M. T. (2000), 'Model selection and semiparametric inference for bivariate failure-time data', *Journal of the American Statistical Association* **95**(449), pp. 62–72.

Appendix I: Properties of the Bivariate Copula Families

I.1 Elliptical copulas

Gaussian copula function is as follows:

$$C(u_1, u_2) = \Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2))$$

Bivariate Student-t copula is as follows:

$$C(u_1, u_2) = t_{\rho,\nu}(t^{-1}(u_1), t^{-1}(u_2))$$

Table .10: Properties of the elliptical copula families

Name	Parameter range	Kendall's τ	Tail dep.(l, u)
Gaussian	$\rho \in (-1, 1)$	$\frac{2}{\pi}\arcsin(\rho)$	(0,0)
Student-t	$\rho \in (-1, 1), \nu > 2$	$\frac{2}{\pi}\arcsin(\rho)$	$\left(2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right), 2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)\right)$

I.2 Archimedean copulas

The bivariate acrchimedean copulas function is:

$$C(u_1, u_2) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2))$$

where $\varphi : [0, 1] \to [0, \infty]$ is a continuous strictly decreasing convex such that $\varphi(1) = 0$ and $\varphi^{[-1]}$ is the pseudo-inverse as follows:

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0), \\ 0, & \varphi(0) \leq t \leq \infty \end{cases}$$

I.3 Rotations of the copulas

In addition to the families presented in the last 2 sections, there are rotated versions of Clayton, Gumbel, Joe, BB1, BB6, BB7 and BB8 families in order to deal with more dependence structure. When the families are rotated by 180 degrees, they are also called the survival forms of the families. The copula function of these copulas will be calculated as follows:

$$\begin{aligned} C_{90}(u_1, u_2) &= u_2 - C(1 - u_1, u_2), \\ C_{180}(u_1, u_2) &= u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2), \\ C_{270}(u_1, u_2) &= u_1 - C(u_1, 1 - u_2), \end{aligned}$$

Where C_{90},C_{180} and C_{270} are the copula C rotated by 90,180 and 270 degree respectively.

Table .11: Properties of bivariate Archimedean copula families

Name	Function	Para. range	Kendall's τ	Tail dep.(l,u)
Clayton	$\frac{1}{\theta}(t^{-\theta}-1)$	$\theta > 0$	$\frac{\theta}{\theta+2}$	$(2^{-\frac{1}{\theta}})$
Gumbel	$(-\log t)^{\theta}$	$\theta \geq 1$	$1-\frac{1}{\theta}$	$(0, 2-2^{\frac{1}{\theta}})$
Frank	$-\log\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$	$\theta \in \Re$	$1-\frac{4}{\theta}+\frac{4D_1(\theta)}{\theta}^{*}$	$(0,0)$
Joe	$-\log(1-(1-t)^{t}heta)$	$\theta > 1$	$1+\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{\frac{2(1-\theta)}{\theta}}dt$	$(0, 2-2^{\frac{1}{\theta}})$
BB1	$(t^{-\theta}-1)^{-\delta}$	$\theta > 0, \delta \geq 1$	$1-\frac{2}{\delta(\theta+2)}$	$(2^{-\frac{1}{\theta\delta}}, 2-2^{\frac{1}{\theta}})$
BB6	$(-\log(1-(1-t)^{\theta}))^{\delta}$	$\theta \geq 1, \delta \geq 1$	$1+\frac{4}{\theta\delta}\int_0^1(-\log(1-(1-t)^{\theta})\times$ $(1-t)(1-(1-t^{-\theta})))dt$	$(0, 2-2^{\frac{1}{\theta\delta}})$
BB7	$(1-(1-t)^{\theta})^{-\delta}$	$\theta \geq 1, \delta > 0$	$1+\frac{4}{\theta\delta}\int_0^1(-(1-(1-t)^{\theta})^{\delta+1}\times$ $\frac{(1-(1-t)^{\theta})^{-\delta}-1}{(1-t)^{\theta}-1})dt$	$(2^{-\frac{1}{\theta}}, 2-2^{\frac{1}{\theta}})$
BB8	$-\log\left(\frac{1-(1-\delta t)^{\theta}}{1-(1-\delta)^{\theta}}\right)$	$\theta \geq 1, \delta \in (0,1]$	$1+\frac{4}{\theta\delta}\int_0^1(-\log\left(\frac{(1-t\delta)^{\theta}-1}{(1-\delta^{\theta}-1)}\right)\times$ $(1-t\delta)(1-(1-t\delta^{-\theta})))dt$	$(0,0)^{**}$

Note: * $D_1(\theta)=\int_0^{\theta}\frac{c/\theta}{\exp(x)-1}dx$ is the Debye function.

** For $\delta = 1$ the upper tail dependence coefficient is $2-2^{\frac{1}{\theta}}$.