Universiti Malaysia Terengganu Journal of Undergraduate Research
Volume 7 Number 1, April 2025 27-38

eISSN: 2637-1138

ANALYSIS OF LIVING COST BY COUNTRIES USING VINE COPULA APPROACH

REDZA-SAFWAN KHAIRUDDIN AND WENDY LING SHINYIE*

Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.

**Corresponding author: sy_ling@upm.edu.my*
https://doi.org/10.46754/umtjur.v7i1.439
Received: 3 August 2023 Revised: 21 November 2024 Accepted: 21 February 2025 Published: 15 April 2025

Abstract: Living cost is defined as the amount of money needed to cover expenses for essential needs. At the onset of the pandemic, the cost of living rose while wages stagnated. As a result, people lost jobs. Therefore, this research aims to perform copula modelling for five variables; living cost, monthly wages, cost index, inflation rate, and purchasing power. A better understanding of the cost of living can help reduce future losses and enhance sustainable economic growth and employment. In this study, we conducted a living cost analysis using marginal distribution to generate a pseudo-observation. Then, we build the vine copula models, which are Regular vine (R-vine), Drawable vine (D-vine), and Canonical vine (C-vine). Then, we obtain and compare the Akaike Information Criterion (AIC) for the three models to determine the best-fitted model. The results of AIC reveal that both C-vine and D-vine are suitable for modelling living cost analysis.

Keywords: Copula modelling, living cost, sustainability, vine copula.

Introduction

Living cost is the amount needed to cover the expenses for essential needs such as food and other day-to-day goods, accommodation, utilities, transportation, healthcare, and taxes (Banton, 2021). In the same article, Banton highlighted that the cost of living can be an important factor in one's financial condition. For example, a person receiving a salary double the average cost of living in a particular city enjoys a better standard of living than their neighbours with the same commitments but receiving a lower salary. In another instance, in New York, which is listed as the city with the highest cost of living in the United States, we expect that the wages of ordinary citizens will be higher than those of residents in other cities to live comfortably. However, the number of homeless people in New York is increasing. The Bowery Mission Organisation reported that one in five New Yorkers lives in poverty. Presently, as a consequence of the coronavirus pandemic, an estimated one in seven New Yorkers have lost their jobs, and more than 50,000 people are at risk of eviction due to the creeping rise in living costs there (Versey & Russell, 2022).

Notably, monthly salaries should equal or exceed the cost of living in expensive cities. Average monthly salaries are recorded to determine information about the wages of the general population. We can learn from recent news that it was reported that the cost of living in Kuala Lumpur has risen to RM3,300 per month, while the average monthly net salary is RM4,259. Therefore, it could be inferred that compared to other Southeast Asian countries except Singapore, people living in Kuala Lumpur have a better quality of life (Foyerglobalhealth, n.d.).

The living cost index measures and compares the cost of living between a city and a metropolitan area. The index is used as a benchmark as it consists of various living expenses, making it a feasible standard to measure new workers. When new students graduate from college or university, they can utilise this index to clearly understand the cost of rent, transportation, and groceries.

Inflation occurs as the cost of necessities and services increases and the value of money decreases. According to personal views and currency exchange rates, inflation can be positive or negative (Fernando, 2021). The inflation rate measures the rate at which prices change monthly or annually. There are three types of inflation, as stated by the World Population Review: Demand-pull inflation, cost-push inflation, and built-in inflation. In particular, demand-pull inflation occurs when the demand for particular goods or services is high, but the production is low. Meanwhile, cost-push inflation occurs when the production cost increases. Moreover, built-in inflation occurs when prices and wages rise to maintain a viable living cost for the people.

Inflation will reduce the value of money's purchasing power and increase prices. Purchasing power is the value of a currency, expressed by the goods or services that can be bought by one unit of currency (Hayes, 2021). If there is inflation, the purchasing power will decrease, and the price will rise. In addition, purchasing power is measured by comparing the price of goods or services against a price index, such as the consumer price index. According to the Cost of Living Report 2020 released by ECA International, Hong Kong has again become the city with the highest cost of living for foreigners worldwide due to the unusually high housing cost. From those reports, we can observe that due to the pandemic, some European cities such as London and France, climbed up the ladder of the list. In contrast, Asian cities such as Mumbai which is considered the most expensive city in India, fell from the ranking in Asia due to currency fluctuations, which impacted the purchasing power of the people.

Over time, the cost of living continues to rise while wages remain stagnant, creating a gap between the minimum wage and the expenses required to maintain a reasonable cost of living. Due to the current epidemic situation, this problem has been brought to an extreme level, which aggravated global extreme poverty and has reached 150 million in 2021. According to the biennial Poverty and Shared Prosperity Report, the extremely poor people living on less than US$1.90 a day are most likely to affect 9.1% to 9.4% of the world's population. It is also reported that some middle-income countries will fall into the extreme poverty lines, a trend already observable in Malaysia. Additionally, the substantial difference in daily commodity prices is increasing. However, the minimum wage has decreased from RM1,200 to RM1,100, making it more challenging for new graduates to find a job with a higher income. Moreover, countries with low inflation rates will have a higher cost of living than countries with high inflation rates. Therefore, this study investigates the relationship between the cost of living, wages, cost of living index, inflation rate, and purchasing power in selected countries including Malaysia.

The main objective of this research is to perform copula modelling for the five selected variables; living cost, monthly wages, cost index, inflation rate, and purchasing power. The 30 selected countries are grouped according to their similar cultural, climatic, and geographical conditions: South America (Brazil and Argentina), North America (United States of America, Canada, and Mexico), East Asia (China, Japan, and South Korea), Southeast Asia (Malaysia, Singapore, Indonesia, Vietnam, and Thailand), South Asia (Bangladesh, India, and Pakistan), West Asia (United Arab Emirates, Turkey, and Qatar), East Europe (Hungary and Russia), North Europe (United Kingdom, Denmark, Sweden, and Norway), South Europe (Italy), and West Europe (Switzerland, Germany, France, and Netherlands). The data used in this study are secondary data obtained from WorldData.info in the year 2020, and the vine copula is the copula model used in this study.

Living cost analysis has been widely researched in the economic sector. Many methods use multiple linear regression to study the variables (Halim *et al*., 2023). However, the Gaussian copula model is preferred for this paper due to its common use nowadays. Based on the data from 30 countries, this paper analyses the cost of living using the copula function. In the next section, we will explain the literature review and methods used, followed by the results and discussions, as well as the conclusion.

Literature Review

Living cost analysis is critical for understanding the expenses of residing in a specific urban environment. Research in this area predominantly employs regression methods, such as multiple linear regression and logarithmic regression. Cebula and Toma (2007) analysed interstate living-cost differentials. They discovered that states with right-to-work laws, colder climates, and higher levels of toxic chemical releases generally have lower overall living costs. However, earlier work by Cebula (1989) identified issues of heteroskedasticity and multicollinearity in geographic living cost analyses using multiple linear regression.

Alternatively, Timmins (2006) explored Brazilian living costs and highlighted the roles of urban market disintegration and congestion in driving costs higher. Beyond regression techniques, Abakah *et al*. (2021) applied pair copula models to examine international bond markets. They demonstrated that regular vine (R-vine) copulas provide superior modelling of interdependencies between markets and improved goodness of fit and mean squared error.

Similarly, Zhou and Ji (2021) concluded that R-vine copulas incorporating dependence dynamics outperformed static and regime-switching vine copula models, offering better accuracy in capturing complex dependencies. This suggests that while traditional regression methods are foundational, advanced techniques like vine copulas can enhance the analysis of living cost dependencies, especially in multifactorial scenarios.

The application of vine copulas in economic and social studies has gained prominence due to their flexibility in modelling complex, high-dimensional dependencies. In analysing the cost of living across countries, vine copulas offer a unique ability to capture non-linear relationships and tail dependencies between multiple economic factors. Notably, vine copulas have been employed to study multivariate dependencies in economic indicators such as housing prices, income distribution, and market behaviours.

On the other hand, Zimmer (2015) highlighted the utility of vine copulas in modelling housing price co-movements, highlighting their effectiveness in capturing dynamic dependencies across regions. Vine copulas are instrumental in integrating data from diverse sources. For instance, social media and historical flood data were combined using this approach, exhibiting its adaptability to incorporate unconventional datasets (Ansell & Valle, 2021).

Studies have extended the use of vine copulas to model multi-country dependencies in mortality and financial stability, emphasising their role in cross-border analyses. The methodology's flexibility allows researchers to address asymmetrical dependencies and extreme scenarios. While not directly applied to living cost studies in existing literature, similar methodologies in housing and agricultural markets illustrate the potential for analysing cost-of-living indices. Applying vine copulas to these domains could uncover latent relationships among factors such as currency exchange rates, inflation, and purchasing power (Sukcharoen & Leatham, 2018). Moreover, incorporating vine copula models into cost-of-living analyses can provide robust insights, particularly in understanding tail dependencies during economic crises or inflationary periods. Therefore, future research could explore coupling vine copulas with machine learning for enhanced predictive capabilities in economic modelling.

Methodology

Material Used

In this study, the data used are national living cost, monthly income, cost index, inflation rate, and purchasing power obtained from the WorldData.info website. The selected countries are Brazil, Argentina, the United States of America, Canada, Mexico, China, Japan, South Korea, Malaysia, Singapore, Indonesia, Vietnam, Thailand, Bangladesh, India, Pakistan, the United Arab Emirates, Turkey, Qatar, Hungary, Russia, the United Kingdom, Denmark, Sweden, Norway, Italy, Switzerland, Germany, France, and the Netherlands.

Copula Theory

A copula is a multivariate probability distribution in which the marginal probability distribution of each variable is uniform (Saarela, 2015). It is a popular form of dependency modelling between disparate processes with different underlying distributions. In this study, copula modelling is selected to analyse living costs, as it is easier to model and estimate the distribution of random vectors by estimating marginals and copulas separately (Voraprateep, 2017).

Marginal Distribution

The marginal distribution is the probability distribution of variables contained in the subset, which offers the variables in the subset various values of probabilities without reference to the values of other variables. The empirical distribution function, commonly known as the Empirical Cumulative Distribution Function (ECDF), is used to estimate the cumulative distribution function and is associated with empirical measures of a sample. It is a step function that increases by 1/*n* when *n* increases. It can be defined as:

$$\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{x_i \le t},$$

where *n* is the sample size, x_i is individual data points in the sample for *i* =1,..., *n,* and *t* is a value for which the ECDF is evaluated.

Other than that, this paper used pseudo-observations to apply a general linear model to estimate a measure of association to quantify the relationship between two or more variables, which is defined as:

$$u_{ij} = \frac{r_{ij}}{n+1} \quad \text{for } i \in \{1,\dots,n\} \text{ and } j \in \{1,\dots,d\},$$

where r_{ij} is the rank of the element *j* in relation to *i*.

The pseudo-observations can be computed by component-wise applying the empirical distribution functions to the data and scaling the result by:

$$\frac{n}{n+1}.$$

This asymptotically negligible scaling factor is used to force the variates to fall inside the open unit hypercube, for example, to avoid problems with density evaluation at the boundaries.

Copula

The Archimedean copula is an associative class of copulas, and it is popular as it allows modelling dependence in arbitrarily high dimensions with only one parameter governing the strength of dependence. Embrechts and Hofert (2011) state that the Archimedean copula has the form of:

$$H(x) = C\big(F_1(x_1),\dots,F_d(x_d)\big) \qquad x \in R^d.$$

Some examples of Archimedean copulas are the Clayton and Gumbel copulas. The Clayton copula is a copula that allows any specific non-zero level of (lower) tail dependency between individual variables. It can be defined as:

$$C_\theta(u_1, \ldots, u_d) = \left(\sum_{i}^{d}\left(u_i^{-\theta}\right) - d + 1\right)^{-1/\theta}.$$

Boateng *et al*. (2022) indicated that the Gumbel copula is a copula that allows any specific level of (upper) tail dependency between individual variables and is defined as:

$$C_\theta(u_1, \ldots, u_d) = \exp\left(-\left(\sum_{u}^{t}(-\log u_i)^\theta\right)^{\frac{1}{\theta}}\right).$$

Other than that, there are elliptical copulas where an elliptical distribution joins univariate margins. The most popular example of an elliptical copula is the Gaussian copula. All the univariate marginal distributions are normal and the multivariate joint distribution is a multivariate normal distribution. It is elliptical and symmetric, which offers nice analytical properties. It can be defined as:

$$C_\Sigma(u_1, \ldots, u_d) = \Phi_\Sigma\left(N^{-1}(u_1), \ldots, N^{-1}(u_d)\right),$$

where Σ is the covariance matrix, which is the parameter of the Gaussian copula, and *N*-1 is the quantile function.

Vine Copula

Vine is a tool for labelling constraints in high-dimensional probability distributions. There are three kinds of vine copula; R-vine, Drawable vine (D-vine), and Canonical vine (C-vine). The degree of a node is the number of edges attached to it. Some advantages of the vine copula strength include its high flexibility and easily constructed model.

R-vine Copula

The R-vine has the simplest degree structure. It is considered as the most flexible than D-vine and C-vine distribution. Figure 1, displayed below, is the R-vine for five variables.

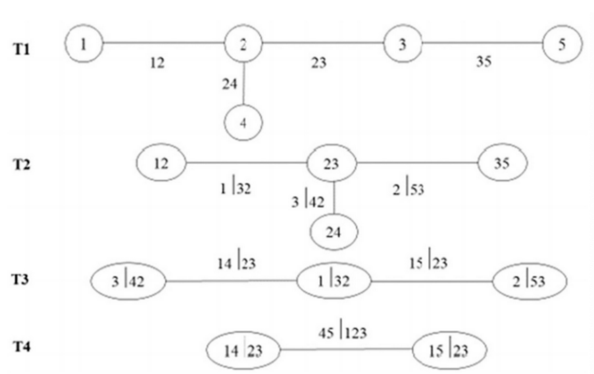


Figure 1: R-vine for five variables (reproduced from "Re-examination of international bond market dependence: Evidence from a pair copula approach," by Abakah *et al*., 2021, International Review of Financial Analysis, 74, 101678)

$$
\begin{aligned}
&f_{12345} = f_1(z_1)f_2(z_2)f_3(z_3)f_4(z_4)f_5(z_5) \cdot c_{12}\big(F_1(z_1), F_2(z_2)\big) \cdot c_{23}\big(F_2(z_2), F_3(z_3)\big),\\
&\cdot c_{24}\big(F_2(z_2), F_4(z_4)\big) \cdot c_{35}\big(F_3(z_3), F_5(z_5)\big) \cdot c_{32|1}\left(F_{2|1}(z_2|z_1), F_{3|1}(z_3|z_1)\right),\\
&\cdot c_{24|3}(F_{2|3}(z_2|z_3), F_{4|3}(z_4|z_3)) \cdot c_{53|2}(F_{3|2}(z_3|z_2), F_{5|2}(z_5|z_2),\\
&\cdot c_{23|14}(F_{2|14}(z_2|z_{14}), F_3(z_3|z_{14})) \cdot c_{15|23}(F_{1|23}(z_1|z_{23}), F_{5|23}(z_5|z_{23}),\\
&\cdot c_{45|123}(F_{4|123}(z_1|z_{123}), F_{5|123}(z_5|z_{123})).
\end{aligned}
$$

D-vine Copula

For the D-vine, the decomposition of the joint densities consists of the pair-copula densities evaluated at the conditional distribution functions and for specified indices and marginal densities (Bedford & Cooke, 2002; Czado, 2019). The D-vine assigns every node of degree 1 or 2. Figure 2 illustrates the D-vine in five variables.

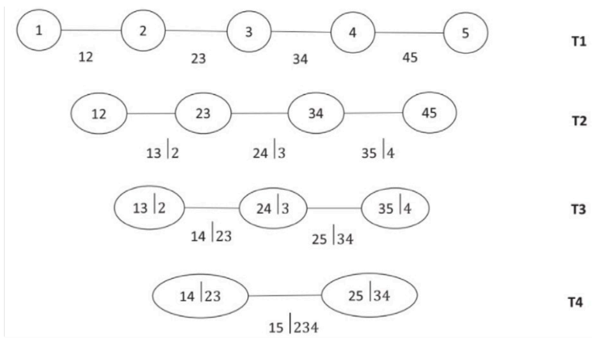


Figure 2: D-vine for five variables (reproduced from "Re-examination of international bond market dependence: Evidence from a pair copula approach," by Abakah *et al*., 2021, International Review of Financial Analysis, 74, 101678)

$$
\begin{aligned}
&f_{12345} = f_1(z_1)f_2(z_2)f_3(z_3)f_4(z_4)f_5(z_5) \cdot c_{12}\big(F_1(z_1), F_2(z_2)\big) \cdot c_{23}\big(F_2(z_2), F_3(z_3)\big),\\
&\cdot c_{34}\big(F_3(z_3), F_4(z_4)\big) \cdot c_{45}\big(F_4(z_4), F_5(z_5)\big) \cdot c_{13|2}\left(F_{1|2}(z_1|z_2), F_{3|2}(z_3|z_2)\right),\\
&\cdot c_{24|3}\left(F_{2|3}(z_2|z_3), F_{4|3}(z_4|z_3)\right) \cdot c_{35|4}\left(F_{3|4}(z_3|z_4), F_{5|4}(z_5|z_4)\right),\\
&\cdot c_{14|23}\left(F_{1|23}(z_1|z_{23}), F_4(z_4|z_{23})\right) \cdot c_{25|34}\left(F_{2|34}(z_2|z_{34}), F_{5|34}(z_5|z_{34})\right),\\
&\cdot c_{15|234}(F_{1|234}(z_1|z_{234}), F_{5|234}(z_5|z_{234})).
\end{aligned}
$$

C-vine Copula

The C-vine assigns one node in each tree to the highest degree. It is generally used when a key variable that controls the relationships of the datasets can be highlighted. Figure 3 presents the C-vine structure for five variables.

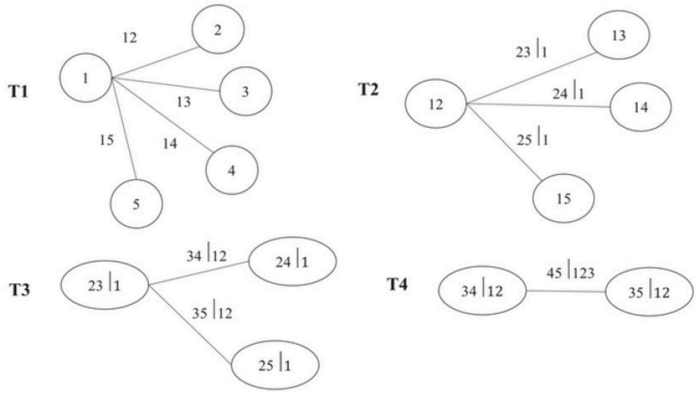


Figure 3: C-vine for five variables (reproduced from "Re-examination of international bond market dependence: Evidence from a pair copula approach," by Abakah *et al*., 2021, International Review of Financial Analysis, 74, 101678)

$$
\begin{aligned}
f_{12345} &= f_1(z_1)f_2(z_2)f_3(z_3)f_4(z_4)f_5(z_5) \cdot c_{12}\big(F_1(z_1),F_2(z_2)\big) \cdot c_{13}\big(F_1(z_1),F_3(z_3)\big),\\
&\cdot c_{14}\big(F_1(z_1),F_4(z_4)\big) \cdot c_{15}\big(F_1(z_5),F_1(z_5)\big) \cdot c_{23|1}\left(F_{2|1}(z_2|z_1),F_{3|1}(z_3|z_1)\right),\\
&\cdot c_{24|1}\left(F_{2|1}(z_2|z_1),F_{4|1}(z_4|z_1)\right) \cdot c_{25|1}\left(F_{2|1}(z_2|z_1),F_{5|1}(z_5|z_1)\right),\\
&\cdot c_{34|12}\left(F_{3|12}(z_3|z_{12}),F_4(z_4|z_{12})\right) \cdot c_{35|12}\left(F_{3|12}(z_3|z_{12}),F_{5|12}(z_5|z_{12})\right),\\
&\cdot c_{45|123}(F_{4|123}(z_1|z_{123}),F_{5|123}(z_5|z_{123})).
\end{aligned}
$$

Goodness of Fit

A goodness of fit test is used to observe how well the data fit a distribution from a population with a normal distribution (Wielaard, 2021). The Akaike Information Criterion (AIC) is used to estimate the quality of each model relative to other models. The AIC formula is used as presented below:

$$\mathrm{AIC} = 2k - 2\ \mathrm{In}\ (\hat{L}).$$

Results and Discussion

The data used in this study is the living cost data collected from WorldData.info for 33 countries across the globe from the year 2020, which includes monthly income, cost index, purchasing power, and inflation rate. Accordingly, pseudo-observation is used to generate a marginal distribution of our data to apply a general linear model to estimate the measure of association between the variables. Consequently, it will be used to generate a vine copula using:

$$u_{ij} = \frac{r_{ij}}{n+1} \quad \text{for } i \in \{1, \dots, n\} \text{ and } j \in \{1, \dots, d\}.$$

Then, data is scaled using $\frac{n}{n+1}$.

As stated in the previous section, we utilise the Gaussian copula family to generate our Vine copula as it is easy to work with, and its strength is its high edibility and easily constructed model. This section will examine three vine copula models: R-vine copula structure, D-vine, and C-vine copula structure.

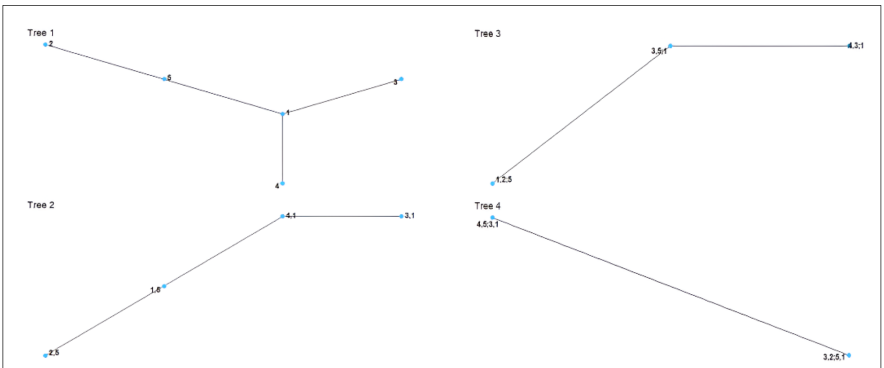
R-vine

From our pseudo-observation data, we can generate our first vine copula which is R-vine, using R programming. The R-vine copula analysis uses c,c var type, Gaussian family, zero rotation, and 1 degree of freedom.

Table 1: Summary for R-vine copula [Variables: Cost of living (1), monthly income (2), cost index (3), purchasing power (4), inflation rate (5)]

Tree	Edge	Conditioned	Conditioning	Variables Types	Family	Rotation	Parameters	df	tau	loglik
1	1	4, 1		c, c	Gaussian	0	0.78	1	0.571	13.390
1	2	3, 1		c, c	Gaussian	0	0.91	1	0.724	25.986
1	3	1, 5		c, c	Gaussian	0	- 0.74	1	- 0.534	11.250
1	4	2, 5		c, c	Gaussian	0	0.55	1	- 0.371	4.627
2	1	4, 3	1	c, c	Gaussian	0	0.086	1	0.055	0.117
2	2	3, 5	1	c, c	Gaussian	0	- 0.061	1	- 0.039	0.058
2	3	1, 2	5	c, c	Gaussian	0	- 0.039	1	- 0.025	0.021
3	1	4, 5	3, 1	c, c	Gaussian	0	- 0.18	1	- 0.117	0.514
3	2	3, 2	5, 1	c, c	Gaussian	0	- 0.31	1	- 0.204	1.557
4	1	4, 2	5, 3, 1	c, c	Gaussian	0	- 0.065	1	- 0.041	0.061

To form the tree, the variable needs to have the lowest log-likelihood value. From Table 1, the first tree consists of edge 4, yielding the lowest value at 4.627. To form the second tree, edge 3 is selected as the log-likelihood value, which is the lowest at 0.021. Next, edge 1 yields a 0.514 log likelihood value to form the third tree. Finally, the fourth tree is formed with edge 1, resulting in a 0.061 log-likelihood value. Correspondingly, we plot the R-vine tree structure:



(a) R-vine tree 1 and 2 pairs (b) R-vine tree 3 and 4 pairs

Figure 4: R-vine tree structure: *C*4,5:3,1-*C*3,2:5,1

D-vine

Next, using the same pseudo-observation data set, we can generate D-vine copula:

Table 2: Summary for D-vine copula [Variables: Cost of living (1), Monthly income (2), Cost index (3), Purchasing power (4), Inflation rate (5)]

Tree	Edge	Conditioned	Conditioning	Var Types	Family	Rotation	Parameters	df	tau	loglik
1	1	1, 2		c, c	Gaussian	0	- 0.43	1	- 0.280	12.16
1	2	2, 3		c, c	Gaussian	0	- 0.53	1	- 0.353	19.30
1	3	3, 4		c, c	Gaussian	0	0.73	1	0.521	41.81
1	4	4, 5		c, c	Gaussian	0	- 0.65	1	- 0.447	29.68
2	1	1, 3	2	c, c	Gaussian	0	0.89	1	0.694	77.45
2	2	2, 4	3	c, c	Gaussian	0	- 0.09	1	- 0.058	0.51
2	3	3, 5	4	c, c	Gaussian	0	- 0.51	1	- 0.339	14.55
3	1	1, 4	3, 2	c, c	Gaussian	0	0.44	1	0.287	11.21
3	2	2, 5	4, 3	c, c	Gaussian	0	0.28	1	0.180	4.56
4	1	1, 5	4, 3, 2	c, c	Gaussian	0	- 0.37	1	- 0.240	7.13

From Table 2, the first tree consists of edge 1 at a 12.16 log-likelihood value. Then, the second tree have edge 2 and produced a 0.51 log-likelihood value. Next, the third tree uses edge 2 to obtain the log-likelihood value of 4.56. Finally, the fourth tree consists of edge 1 and produces a result of a 7.13 log-likelihood value.

After that, we generate tree structure for D-vine copula:

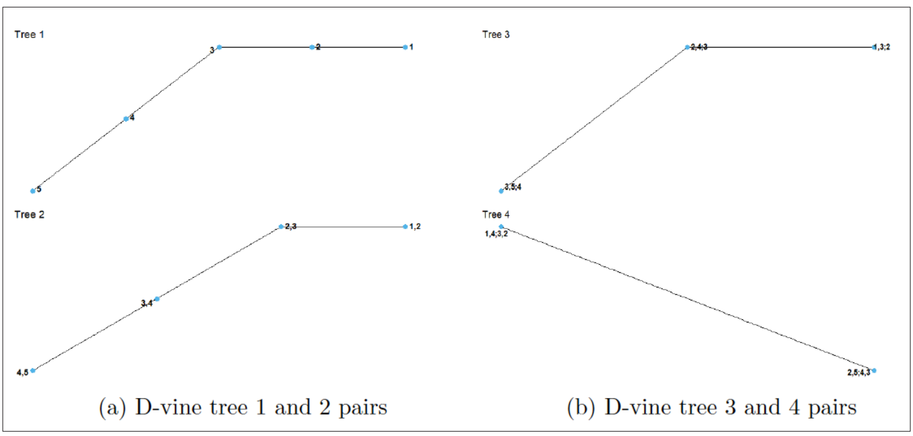


Figure 5: D-vine tree structure: $C_{1, 4, 3, 2}$ - $C_{2, 5, 4, 3}$

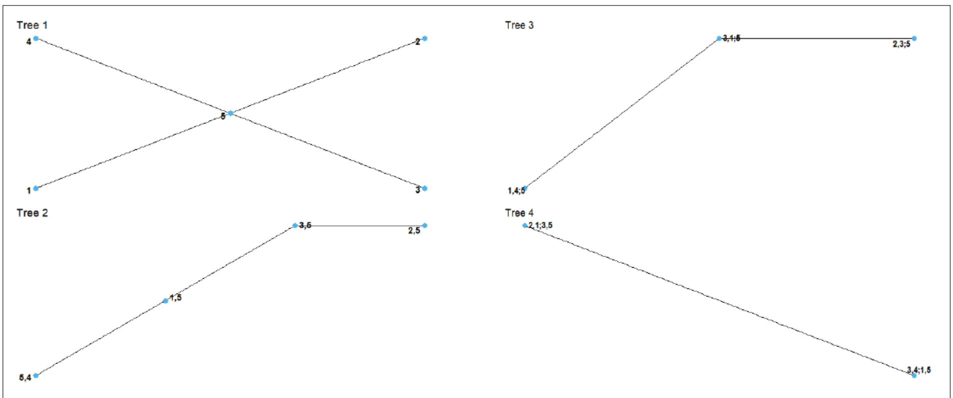
C-vine

As for C-vine copula, we need to specify a node in the tree to the highest degree. Hence, we choose the first node prior to generating the copula.

Table 3: Summary for C-vine copula [Variables: Cost of living (1), Monthly income (2), Cost index (3), Purchasing power (4), Inflation rate (5)]

Tree	Edge	Conditioned	Conditioning	Var Types	Family	Rotation	Parameters	df	tau	loglik
1	1	2, 5		c, c	Gaussian	0	0.56	1	0.376	21.64
1	2	3, 3		c, c	Gaussian	0	- 0.74	1	- 0.530	41.25
1	3	1, 5		c, c	Gaussian	0	- 0.77	1	- 0.558	46.35
1	4	5, 4		c, c	Gaussian	0	- 0.65	1	- 0.447	29.68
2	1	2, 3	5	c, c	Gaussian	0	- 0.21	1	- 0.135	2.67
2	2	3, 1	5	c, c	Gaussian	0	0.78	1	0.573	48.54
2	3	1, 4	5	c, c	Gaussian	0	0.58	1	0.397	22.48
3	1	2, 1	3, 5	c, c	Gaussian	0	0.27	1	0.173	4.26
3	2	3, 4	1, 5	c, c	Gaussian	0	0.072	1	0.046	0.28
4	1	2, 4	1, 3, 5	c, c	Gaussian	0	- 0.14	1	- 0.089	1.17

From Table 3, the first tree consists of edge 1 as specified, resulting in a 21.64 log-likelihood value. Then, the second tree uses edge 1, producing a 2.67 log-likelihood value. For the third tree, edge 2 is selected as the log-likelihood value is 0.28. Finally, the fourth tree has edge 1 at a 1.17 log-likelihood value. Consequently, we draw the C-vine tree structure as specified:



(a) C-vine tree 1 and 2 pairs (b) C-vine tree 3 and 4 pairs

Figure 6: C-vine structure: $C_{2, 1, 3, 5}$ - $C_{3, 4, 1, 5}$

Goodness-of-fit Test

From the model, each model AIC can be calculated using the formula:

$$\text{AIC} = 2k - 2In(\hat{L})$$

We aim to find the most suitable fitted model for our dataset.

Table 4: AIC comparison for each model

Model	AIC
R-vine	- 95.1615
D-vine	- 416.7066
C-vine	- 416.659

Table 4 summarises the AIC values for all three models. When comparing the AIC of all three models, the D-vine copula has the lowest value, which is -416.7066, followed by the C-vine with an AIC of -416.659. Therefore, C-vine and D-vine copulas are suitable fitting models for our living cost data.

Conclusions

This study aimed to analyse the living cost among 33 selected countries using copula modelling and conduct a goodness of fit test with these variables to achieve the best model fit. Based on the analysis, copula models were chosen to fit the selected living cost data. In this study, we select the vine copula model to measure the data. The selected models are R-vine, D-vine, and C-vine copula using Gaussian copula.

Before fitting the data, the marginal distribution is generated from the data set using the pseudo-observations. Subsequently, to test the goodness of fit of the distribution to the data set, the AIC value of each method was obtained and compared.

Among these models, the D-vine model produces the lowest AIC value, closely followed by the C-vine copula with the second-lowest AIC value. This indicates that C- and D-vine copulas are well-suited for modelling the living cost data set. Notably, the cost of living fluctuates greatly as it is affected by many factors. In this study, only five variables were selected as research objects. Thus, it is suggested that different types of models be included in this research to analyse and compare the better-fitted models and improve the analysis's accuracy.

The results of this study aim to provide policymakers and decision-makers with research evidence of variables to ensure that the cost of living crisis will not become a crisis for sustainable living.

Acknowledgements

We acknowledge the financial support from the Ministry of Education Malaysia for the fundamental research grant scheme (FRGS/1/2019/STG06/UPM/02/7).

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

Abakah, E. J. A., Addo, E., Gil-Alana, L. A., & Tiwari, A. K. (2021). Re-examination of international bond market dependence: Evidence from a pair copula approach. *International Review of Financial Analysis*, *74*, Article 101678.

Ansell, L., & Valle, L. D. (2022). Social Media Integration of flood data: A Vine Copula-Based approach. *Journal of Environmental Informatics*, *39*(2), 97-110.

Banton, C. (2025, January 5). *Cost of living: Definition, how to calculate, index, and example*. Investopedia. Retrieved March 26, 2023 from https://www.investopedia.com/terms/c/cost-of-living.asp

Bedford, T., & Cooke, R. M. (2002). Vines: A new graphical model for dependent random variables. *Annals of Statistics*, *30*(4), 1031-1068.

Boateng, M. A., Omari-Sasu, A. Y., Avuglah, R. K., & Frempong, N. K. (2022). A mixture of Clayton, Gumbel and Frank Copulas: A complete dependence model. *Journal of Probability and Statistics*, *2022*, 1-7.

Cebula, R. J. (1989). The analysis of geographic living-cost differentials: A brief empirical note. *Land Economics*, *65*(1), 64-67.

Cebula, R. J., & Toma, M. (2007). An empirical analysis of determinants of interstate living-cost differentials, 2005. *The Journal of Regional Analysis and Policy*, *38*(3), 222-228. https://ageconsearch.umn.edu/record/133000/files/08-3-2.pdf

Czado, C. (2019). *Analyzing dependent data with vine copulas: A practical guide with R*. New York: Springer. https://link.springer.com/content/pdf/10.1007/978-3-030-13785-4.pdf

Embrechts, P., & Hofert, M. (2011). Comments on: Inference in multivariate Archimedean copula models. *Test*, *20*(2), 263-270.

Fernando, J. (2025, January 25). *Inflation: What it is and how to control inflation rates*. Investopedia. Retrieved March 26, 2023 from https://www.investopedia.com/terms/i/inflation.asp

Foyerglobalhealth. (n.d.) *The cost of living in Malaysia*. Global Health Foyer Group. Retrieved November 20, 2023, from https://www.foyerglobalhealth.com/destination/the-cost-of-living-in-malaysia/

Halim, G. A., Agustin, P., Adiwijayanto, E. & Ohyver, M. (2023). Estimation of cost of living in a particular city using multiple regression analysis and correlation of residual assumptions through appropriate methods. *Procedia Computer Science*, *216*, 613-619.

Saarela, S. (2015). Use of remotely sensed auxiliary data for improving sample-based forest inventories. Dissertationes Forestales, *2015*(201). https://doi.org/10.14214/df.201

Sukcharoen, K., & Leatham, D. (2018). Analyzing extreme comovements in agricultural and energy commodity markets using a regular vine copula method. *International Journal of Energy Economics and Policy*, *8*(5), 193-201.

Timmins, C. (2006). Estimating spatial differences in the Brazilian cost of living with household location choices. *Journal of Development Economics*, *80*(1), 59-83.

Versey, H. S., & Russell, C. N. (2022). The impact of COVID-19 and housing insecurity on lower-income Black women. *Journal of Social Issues*, *79*(2), 773-793. https://doi.org/10.1111/josi.12555

Voraprateep, J. (2017). *Statistical properties of liquid protein-water molecular system dynamics* (Doctoral dissertation, Aston University). https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.774113

Wielaard, J. (2021). *Model selection and goodness of fit tests for conditional copula models* (Master's thesis, University of Twente).

Zhou, R., & Ji, M. (2021). Modelling mortality dependence: An application of dynamic vine copula. *Insurance Mathematics and Economics*, *99*, 241-255. https://doi.org/10.1016/j.insmatheco.2021.03.022

Zimmer, D. M. (2015). Analyzing comovements in housing prices using vine copulas. *Economic Inquiry*, *53*(2), 1156-1169.