METRON (2017) 75:109–124
DOI 10.1007/s40300-016-0104-9

Trivariate Burr-III copula with applications to income data

Lucio De Capitani[1] · **Federica Nicolussi**[2] · **Alessandro Zini**[1]

Received: 20 November 2015 / Accepted: 10 November 2016 / Published online: 28 November 2016

Abstract In this work, Bivariate Burr-III copula is extended to the trivariate case. This copula seems to be very general and analytically manageable and it provides an alternative to the commonly employed elliptical copulas (such as the Gaussian or the Stutent's *t* ones) since they have, roughly, the same number of parameters. Several applications to income and wine data are described in the paper. They show that the Trivariate Burr-III copula is, in general, able to capture the dependence structure implicit in observed trivariate data. Moreover, they show that the third-order interaction parameter results, in some cases, significant at 1% significance level while, in other cases, it can be removed from the fitted model. The ability of the Trivariate Burr-III copula in representing the dependence structure implicit in the considered data is compared with the ones of other well known copulas: the Clayton copula, the *t* copula, and the Skew-*t* copula. It results that the Trivariate Burr-III copula provides a good fitting and turns out to be the best performer in fitting the considered wine data but, on income data, the best performers are the *t* and Skew-*t* copulas. The over-performance of the last two copulas on income data is probably due to their ability in representing right-tail dependence (a kind of dependence that is not taken into account by the Trivariate Burr-III copula).

Keywords Income distribution · Trivariate Burr-III · Multivariate Copula

✉ Lucio De Capitani
lucio.decapitani1@unimib.it

Federica Nicolussi
federica.nicolussi@unicatt.it

Alessandro Zini
alessandro.zini@unimib.it

1 Department of Statistics and Quantitative Methods, Università degli Studi di Milano-Bicocca, via Bicocca degli Arcimboldi n. 8, 20126 Milan, Italy

2 Department of Mathematical Sciences, Mathematical Finance and Econometrics, Università Cattolica del Sacro Cuore, Milan, Italy

1 Introduction

The Univariate BUR-III distribution was introduced by Burr in 1942 [4]. In economics, this distribution is more widely known, after the introduction of an additional scale parameter, as the Dagum distribution [6]. Specifically, the Dagum cumulative distribution function with parameter (λ, δ, β) is given by

$$F(x) = \frac{1}{(1 + \lambda x^{-\delta})^{\beta}} \qquad x > 0, \quad \lambda > 0, \quad \delta > 0 \; \beta > 0. \tag{1}$$

Model (1) is well known for its successful applications on income, wage and wealth data (see, for example, [14], and the references therein). The first extension of this model to the multivariate case is due to Rodriguez [15] who introduced the so-called "Bivariate Burr-III distribution". The cumulative distribution of such model is given by

$$F_{X,Y}(x, y) = \begin{cases} \dfrac{1}{(1 + \alpha\lambda\gamma x^{-\theta} y^{-\delta} + \lambda x^{-\theta} + \gamma y^{-\delta})^{\epsilon}} & x > 0, \; y > 0 \\ 0 & otherwise \end{cases}, \tag{2}$$

with $\lambda > 0$, $\gamma > 0$, $\theta > 0$, $\delta > 0$, $\epsilon > 0$, and $0 \le \alpha \le (\epsilon + 1)$. Indeed, (2) is a bivariate Dagum distribution since it has Dagum univariate margins with parameters $(\lambda, \theta, \epsilon)$ and $(\gamma, \delta, \epsilon)$. This distribution has been deeply studied in [9] which provides its main dependence properties. For our purposes it is interesting to consider the copula implicit in distribution (2):

$$C_{X,Y}(u, v) = \left\{1 + \alpha(u^{-\frac{1}{\epsilon}} - 1)(v^{-\frac{1}{\epsilon}} - 1) + (u^{-\frac{1}{\epsilon}} - 1) + (v^{-\frac{1}{\epsilon}} - 1)\right\}^{-\epsilon}, \tag{3}$$

with $0 \le \alpha \le (\epsilon + 1)$.

In this paper the Bivariate Burr-III distribution and its copula are extended to the trivariate case. Some preliminary results on these distributions are provided along with some applications to income and wine data.

2 The Trivariate Burr-III distribution and its copula

In order to extend (3) to the trivariate context, the Trivariate Burr-III distribution is first built. Let $\mathbf{X} = (X_1, X_2, X_3)$ be a random vector where X_i, $i = 1, 2, 3$, follows the Dagum distribution with parameters $(\lambda_i, \theta_i, \epsilon)$. In the independence case the joint distribution of $\mathbf{X}$ is:

$$\begin{aligned} F_{\mathbf{X}}(x_1, x_2, x_3) &= \prod_{i=1}^{3} \frac{1}{\left(1 + \lambda_i x_i^{-\theta_i}\right)^{\epsilon}} \\ &= \Big[1 + \lambda_1 x_1^{-\theta_1} + \lambda_2 x_2^{-\theta_2} + \lambda_3 x_3^{-\theta_3} + \lambda_1\lambda_2 x_1 x_2^{\theta_1\theta_2} + \lambda_1\lambda_3 x_1 x_3^{\theta_1\theta_3} \\ &\quad + \lambda_2\lambda_3 x_2 x_3^{\theta_2\theta_3} + \lambda_1\lambda_2\lambda_3 x_1 x_2 x_3^{-\theta_1\theta_2\theta_3}\Big]^{-\epsilon}. \end{aligned} \tag{4}$$

In order to introduce some kind of dependence among the random variables (X_1, X_2, X_3), expression (4) can be modified by introducing the second order interaction parameters $(\alpha_{12}, \alpha_{13}, \alpha_{23})$ and the third order interaction parameter α_{123}:

$$F_{\mathbf{X}}(x_1, x_2, x_3) = \Big[1 + \lambda_1 x_1^{-\theta_1} + \lambda_2 x_2^{-\theta_2} + \lambda_3 x_3^{-\theta_3} + \alpha_{12}\lambda_1\lambda_2 x_1 x_2^{\theta_1\theta_2} + \alpha_{13}\lambda_1\lambda_3 x_1 x_3^{\theta_1\theta_3} + \alpha_{23}\lambda_2\lambda_3 x_2 x_3^{\theta_2\theta_3} + \alpha_{123}\lambda_1\lambda_2\lambda_3 x_1 x_2 x_3^{-\theta_1\theta_2\theta_3}\Big]^{-\epsilon} \quad (5)$$

with restrictions on the interaction parameters. It is worth noting that (5) has univariate Dagum margins and bivariate Burr-III margins. Then (5) is an extension to the trivariate case of the bivariate distribution (2).

Now, by substituting in (5) the inverse functions of the margins, the following trivariate copula is obtained:

$$C(u_1, u_2, u_3) = \Big[1 + \sum_{i=1}^{3}(u_i^{-\frac{1}{\epsilon}} - 1) + \alpha_{12}(u_1^{-\frac{1}{\epsilon}} - 1)(u_2^{-\frac{1}{\epsilon}} - 1) + \alpha_{13}(u_1^{-\frac{1}{\epsilon}} - 1)(u_3^{-\frac{1}{\epsilon}} - 1) + \alpha_{23}(u_2^{-\frac{1}{\epsilon}} - 1)(u_3^{-\frac{1}{\epsilon}} - 1) + \alpha_{123}(u_1^{-\frac{1}{\epsilon}} - 1)(u_2^{-\frac{1}{\epsilon}} - 1)(u_3^{-\frac{1}{\epsilon}} - 1)\Big]^{-\epsilon}. \quad (6)$$

It is worth noting that (6) is an extension of the bivariate copula (3) according to the following definition ([12], page 155).

Definition 1 A m-variate parametric family of copulas is an extension of a bivariate family if:

(i) all bivariate marginal copulas of the multivariate copula are in the given family;
(ii) all multivariate marginal copulas of order 3 to $m - 1$ have the same multivariate form.

As mentioned above, appropriate restrictions to the values of the parameters $(\alpha_{12}, \alpha_{13}, \alpha_{23})$ and α_{123} had to be imposed in order to assure that (6) is a well-defined trivariate copula. As sometimes happens (see the copula model BB4 in [12], pp. 152, point c), to find these restrictions is quite hard. The lemma below describes a necessary condition that must be satisfied in order to assure that (6) is a well-defined copula.

Lemma 1 *If $C(u_1, u_2, u_3)$ in (6) is a copula then*

(a) $\alpha_{12} \le (\epsilon + 1)$, $\alpha_{13} \le (\epsilon + 1)$*, and* $\alpha_{23} \le (\epsilon + 1)$;
(b) $\alpha_{123} \le (\epsilon + 1) min(\alpha_{12}\alpha_{13};\ \alpha_{12}\alpha_{23};\ \alpha_{13}\alpha_{23})$.

Proof Let U_1, U_2, and U_3 be three uniform random variables with joint distribution function $C(u_1, u_2, u_3)$. Condition a) follows from the restrictions on the parameters of the bivariate marginal distributions of (U_1, U_2), (U_1, U_3), and (U_2, U_3). In order to prove condition b), consider the bivariate conditional cdf

$$C_{12|3}(u_1, u_2) = P(U_1 \le u_1, U_2 \le u_2 | U_3 \le u_3) \quad (7)$$

$$= \frac{1}{u_3} C(u_1, u_2, u_3)$$

$$= \Big[1 + \sum_{i=1}^{3} v_i + \alpha_{12} v_1 v_2 + \alpha_{13} v_1 v_3 + \alpha_{32} v_2 v_3 + \alpha_{123} v_1 v_2 v_3\Big]^{-\epsilon} \quad (8)$$

with $v_i = (u_i^{-\frac{1}{\epsilon}} - 1)$. The bivariate density implicit in $C_{12|3}$ is

$$\frac{\partial}{\partial u_1 \partial u_2} \frac{1}{u_3} C(u_1, u_2, u_3) = \frac{u_1^{-\epsilon-1} u_2^{-\epsilon-1}}{\epsilon u_3} H^{-\epsilon-2}$$
$$\cdot \Bigg\{ \underbrace{(\epsilon + 1)\,[(1 + \alpha 12 v_1 + \alpha 23 v_3 + \alpha 123 v_1 v_3)\,(1 + \alpha 12 v_2 + \alpha 13 v_3 + \alpha 123 v_2 v_3)]}_{M_1}$$
$$+ \underbrace{H(\alpha_{12} + \alpha_{123} v_3)}_{M_2} \Bigg\} = \frac{u_1^{-\epsilon-1} u_2^{-\epsilon-1}}{\epsilon u_3} H^{-\epsilon-2} \{M_1 - M_2\} \tag{9}$$

where

$$H = 1 + \sum_{i=1}^{3} v_i + \alpha_{12} v_1 v_2 + \alpha_{13} v_1 v_3 + \alpha_{32} v_2 v_3 + \alpha_{123} v_1 v_2 v_3.$$

The density (9) must be positive and this fact is assured if

$$\frac{M_1}{M_2} \geq 1 \quad \text{for all} \quad (u_1, u_2, u_3) \in (0, 1)^3. \tag{10}$$

Now, observe that

$$\lim_{u_1 \to 1} \left(\lim_{u_2 \to 1} \left(\lim_{u_3 \to 0} \frac{M_1}{M_2} \right) \right) = \frac{(\epsilon + 1)\alpha_{23}\alpha_{13}}{\alpha_{123}}. \tag{11}$$

From (10) and (9) it results that if the bivariate density implicit in $C_{12|3}$ is positive on $(0, 1)^2$ then

$$\frac{(\epsilon + 1)\alpha_{23}\alpha_{13}}{\alpha_{123}} \geq 1$$

and, consequently

$$\alpha_{123} \leq (\epsilon + 1)\alpha_{13}\alpha_{23}.$$

Repeating the same procedure by conditioning with respect to U_1 and U_2 the additional constraints

$$\alpha_{123} \leq (\epsilon + 1)\alpha_{12}\alpha_{23} \quad \text{and} \quad \alpha_{123} \leq (\epsilon + 1)\alpha_{12}\alpha_{13}$$

are obtained and their intersections coincides with condition *b*).

The necessary condition provided by Lemma 1 have been numerically tested in order to understand if they can be considered also as sufficient condition. Specifically, the non-negativity of the copula density function implicit in $C(u_1, u_2, u_3)$ have been tested on a very fine grid of $(0, 1)^3$ for a wide range of parameters values. This numerical experiment suggests that the necessary condition is reasonably sufficient too. However, we underline that this fact has not been confirmed analytically.

3 Applications to income data: methodological aspects

Here we try to model the joint distribution of different income sources. We consider the survey on household income and wealth provided in [5] which covers 8151 households whose total disposable income is split in four main income sources: income from salaries (S); income from profession (P); income from capital gains (C); income from net-transfers and pensions (T).

We study the dependence of the income deriving from work (i.e. S and/or P) and the other two income sources (i.e. C and T). In particular, the trivariate dataset (S, C, T), (P, C, T) and $(S + P, C, T)$ will be analyzed.

The study of a model with the joint distribution of S and P is discarded since S and P are almost incompatible.

A simple inspection of these trivariate dataset highlights that there is a high number of observed income sources that are equal to zero or in its proximity.

In detail, in the (S, C, T) dataset only about the 17% of observations have the three sources jointly different from zero, in the (P, C, S) this percentage reduces about to 6% while in the $(S + P, C, T)$ dataset it is equal to 21% (it is worth noting that the percentage of observations with the income sources $(S + P, C, T)$ jointly different from 0 is lower than $17\% + 6\% = 23\%$ since S and P are jointly equal to 0 in about the 2% of observations).

This fact emphasizes that an absolutely continuous model is not adequate to represent the joint distribution of (S, C, T), (P, C, T) and $(S+P, C, T)$. For instance, the joint distribution of (S, C, T) can be split as follow:

$$\begin{aligned} F_{SCT}(s, c, t) = \; & F_{SCT}(s, c, t|S \neq 0, C \neq 0, T \neq 0) P(S \neq 0, C \neq 0, T \neq 0) \\ & + F_{SCT}(s, c, 0|S \neq 0, C \neq 0) P(S \neq 0, C \neq 0, T = 0) \\ & + F_{SCT}(s, 0, t|S \neq 0, T \neq 0) P(S \neq 0, C = 0, T \neq 0) \\ & + F_{SCT}(0, c, t|C \neq 0, T \neq 0) P(S = 0, C \neq 0, T \neq 0) \\ & + F_{SCT}(s, 0, 0|S \neq 0) P(S \neq 0, C = 0, T = 0) + \\ & + F_{SCT}(0, c, 0|C \neq 0) P(S = 0, C \neq 0, T = 0) \\ & + F_{SCT}(0, 0, t|T \neq 0) P(S = 0, C = 0, T \neq 0) \\ & + F_{SCT}(0, 0, 0) P(S = 0, C = 0, T = 0) \end{aligned} \tag{12}$$

A similar decomposition can be done concerning (P, C, T) and $(S + P, C, T)$. Each part of the model can now be fitted separately.

At first we consider the estimation problem of the trivariate distributions $F_{SCT}(s, c, t|S \neq 0, C \neq 0, T \neq 0)$, $F_{PCT}(p, c, t|P \neq 0, C \neq 0, T \neq 0)$, and $F_{(S+P)CT}(s + p, c, t|S + P \neq 0, C \neq 0, T \neq 0)$ which can be reasonably assumed absolutely continuous. In particular, considering, for example, (S, C, T), the estimate of $F_{SCT}(s, c, t|S \neq 0, C \neq 0, T \neq 0)$ is given by:

$$\hat{F}_{SCT}(s, c, t|S \neq 0, C \neq 0, T \neq 0) = \hat{C}(\hat{G}_S(s), \hat{G}_C(c), \hat{G}_T(t)) \tag{13}$$

where $\hat{C}$ denotes estimated trivariate Burr-III copula in formula (6) while $\hat{G}_\bullet$ is the fitted model for the single income component "•". In the following, the estimates are obtained by considering a "pseudo"-CML method. In particular, the Canonical Maximum Likelihood method (CML method, see [10]) is related to the IFM method proposed in [13] and consists in a two step approach in which the ML estimates of the marginal distributions are estimated first and, at second, the estimates of the remaining dependence parameters are obtained by maximizing the likelihood related to the empirical copula function. Here we adopt a slight modification of this method. In detail: (a) the parameters of the marginal distribution are estimated by following the minimum χ^2-method in place of the ML method (see [1], [2], and [8], for application of this estimation method in incomes modeling); (b) the parameters of the copula function are separately obtained by using two different estimation methods: the minimum χ^2 method and the ML method. Concerning the estimation of the marginal distributions, three different models for single income components are here considered. The three-parameters Zenga Distribution (ZD) [16], with cumulative distribution function

$$G(x) = \begin{cases} \dfrac{1}{B(\alpha;\theta)} \displaystyle\sum_{i=0}^{\infty} \left(IB\left(\frac{x}{\mu} : \alpha+i;\theta\right) - \left(\frac{\mu}{x}\right)^{\frac{1}{2}} IB\left(\frac{x}{\mu} : \alpha+i+\frac{1}{2};\theta\right)\right) & 0 < x \le \mu \\ 1 - \dfrac{1}{B(\alpha;\theta)} \displaystyle\sum_{i=0}^{\infty} \left(IB\left(\frac{x}{\mu} : \alpha+i;\theta\right) - \left(\frac{\mu}{x}\right)^{\frac{1}{2}} IB\left(\frac{x}{\mu} : \alpha+i+\frac{1}{2};\theta\right)\right) & \mu < x. \end{cases} \tag{14}$$

The four-parameters generalization of ZD (GZD) [7], with cumulative distribution function

$$\tilde{G}(x) = \begin{cases} C \displaystyle\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} (-1)^j \frac{\gamma^j}{j!} \left(IB\left(\frac{x}{\mu} : \alpha+i+j;\theta\right) - \left(\frac{\mu}{x}\right)^{\frac{1}{2}} IB\left(\frac{x}{\mu} : \alpha+i+j+\frac{1}{2};\theta\right)\right) & 0 < x \le \mu \\ 1 - C \displaystyle\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} (-1)^j \frac{\gamma^j}{j!} \left(IB\left(\frac{x}{\mu} : \alpha+i+j;\theta\right) - \left(\frac{\mu}{x}\right)^{\frac{1}{2}} IB\left(\frac{x}{\mu} : \alpha+i+j+\frac{1}{2};\theta\right)\right) & \mu < x, \end{cases} \tag{15}$$

where $C = (B(\alpha;\theta)\,{}_1F_1(\alpha;\theta+\alpha;-\gamma))^{-1}$. Finally, even the Dagum distribution (1) is considered.

A location parameter h has been added to these models in order to accounts for the possible negative values of some income components. Negative values occur especially for the C component (probably due to losses derived from financial investments) and for the T component (due to the payment of alimony and gifts).

We observed a high concentration of values in the interval $(-100; 100)$ in the distributions of the component C and T (which correspond to interests or coupon of small investments or negligible transfers). For this reason we take into account a further model with the restrictions $C \notin (-100; 100)$ and $T \notin (-100; 100)$ instead of $C \neq 0$ and $T \neq 0$. In detail, in the following applications the distribution functions

$$F_{SCT}(s, c, t | S \neq 0, C \notin (-100; 100), T \notin (-100; 100)),$$

$$F_{PCT}(p, c, t | P \neq 0, C \notin (-100; 100), T \notin (-100; 100))$$

and

$$F_{(S+P)CT}(s+p, c, t | S+P \neq 0, C \notin (-100; 100), T \notin (-100; 100))$$

will be estimated too.

The distributions $F_{SCT}(s, c, 0|S \neq 0, C \neq 0)$, $F_{SCT}(s, 0, t|S \neq 0, T \neq 0)$ and $F_{SCT}(0, c, t|C \neq 0, T \neq 0)$ (2nd–4th rows of formula 13) are estimated by considering the bivariate BUR III copula in formula (3).

The distributions $F_{SCT}(s, 0, 0|S \neq 0)$, $F_{SCT}(0, c, 0|C \neq 0)$ and $F_{SCT}(0, 0, t|T \neq 0)$ (5th-7th rows of formula 13) are estimated by considering the univariate distributions ZD, GZD and Dagum in formula (14), (15) and (1), respectively. The same is done for the dataset (P,C,T) and (S + P,C,T).

The estimation is repeated for the dataset with restrictions on C and T.

4 Applications to income data: results

The parameter estimates of the Trivariate Burr-III copula in the three dataset (S, C, T), (P, C, T) and $(S + P, C, T)$ are reported in Table 1. These estimates are obtained, as explained in the previous Section, by using the minimum χ^2 method (the number of classes k used to implement this method is reported in the third column of that Table 1) and the ML method.

Table 1 Parameter estimates of the Burr-III copula saturated model and of the reduced model with $\alpha_{123} = 0$

Dataset	n	k	mi	α_{12}	α_{13}	α_{23}	α_{123}	ϵ	χ^2	
Estimates obtained via the minimum χ^2 estimation method										
(S, C, T)	1419	4^3	0	0.4894	1.1235	0.3929	0.1823	0.4607	0.3084	
(S, C, T)	1419	4^3	0	0.4279	1.1574	0.3160	–	0.8775	0.3382	
(S, C, T)	1278	4^3	100	0.5243	1.1203	0.4768	0.207	0.4056	0.3166	
(S, C, T)	1278	4^3	100	0.4727	1.1473	0.4150	–	0.8579	0.3495	
(P, C, T)	475	4^3	0	0.4998	1.3425	0.2157	≈0	1.1175	0.3604	
(P, C, T)	475	4^3	0	0.4997	1.3423	0.2158	–	1.1174	0.3605	
(P, C, T)	448	4^3	100	0.5940	1.26792	0.2847	0.1685	0.7193	0.3653	
(P, C, T)	448	4^3	100	0.5361	1.3373	0.1932	–	1.0535	0.3735	
(S + P, C, T)	1706	4^3	0	0.4994	1.22	0.3645	0.1690	0.4438	0.2901	
(S + P, C, T)	1706	4^3	0	0.4454	1.2907	0.2981	–	0.8135	0.3272	
(S + P, C, T)	1548	4^3	100	0.5363	1.2019	0.3995	0.1625	0.4293	0.2723	
(S + P, C, T)	1548	4^3	100	0.5107	1.2563	0.3466	–	0.7373	0.3010	
Dataset	n	–	mi	α_{12}	α_{13}	α_{23}	α_{123}	ϵ	l	LR
Estimates obtained via the Maximum Likelihood estimation method										
(S, C, T)	1419		0	0.4816	1.2248	0.3666	0.1902	0.3337	147.15	
(S, C, T)	1419		0	0.4502	1.1808	0.3152	–	0.9819	105.72	82.86
(S, C, T)	1278		100	0.5359	1.1068	0.4162	0.1928	0.3528	113.54	
(S, C, T)	1278		100	0.4979	1.2223	0.3500	–	1.0154	79.07	68.94
(P, C, T)	475		0	0.5214	1.0502	0.3378	0.1586	0.3026	55.96	
(P, C, T)	475		0	0.4887	1.3311	0.2387	–	0.9844	42.29	27.34
(P, C, T)	448		100	0.5084	1.06188	0.3363	0.1468	0.3209	53.43	
(P, C, T)	448		100	0.5107	1.2647	0.2394	–	0.8620	45.04	16.78
(S + P, C, T)	1706		0	0.5065	1.2009	0.3468	0.1813	0.3801	212.87	
(S + P, C, T)	1706		0	0.4531	1.2766	0.2882	–	0.8538	166.97	91.8
(S + P, C, T)	1548		100	0.5114	1.1088	0.3668	0.1417	0.3560	177.83	
(S + P, C, T)	1548		100	0.4890	1.3191	0.3105	–	0.8690	140.63	74.4

k The number of classes used to implement the minimum χ^2 method, n The number of observations in the corresponding dataset; mi is the threshold of the sources C and T, χ^2 the value of the χ^2-statistic and l the value of the log-likelihood function, LR the likelihood ratio statistic to test the full models in the odd rows against the reduced models in the pair rows

Concerning the minimum χ^2 estimates, it can be observed that, for each dataset, the multivariate Burr-III copula has a goodness of fit less or equal to 0, 35 except for (P, C, T), which has the smallest number of observations. The parameters estimates obtained with the minimum χ^2 or with the ML methods are quite similar. Moreover, as expected, the value of the χ^2 statistics and of the log-likelihood function in the last column of Table 1 worsen when passing from the saturated copula model to the constrained one (with $\alpha_{123} = 0$) and, in some cases, a sensible difference is observed. In order to evaluate if the parameter α_{123} is significantly different from zero, a likelihood ratio test is performed. In all cases the null hypothesis $H_0 : \alpha_{123} = 0$ is rejected at a significance level $\alpha = 0.01$.

In order to evaluate the goodness of fit of the estimated copula models, a comparison with the fitting obtained by using the trivariate Clayton copula (which has only one parameter), the trivariate Student's *t* copula (which has 4 parameters), and the trivariate Skew-*t* copula (which has 7 parameters, see [3]) is performed. For these copulas, the parameters estimates are obtained by using the minimum χ^2 method. The obtained results are reported in Tables 2, 3, and 4, for the Clayton, *t*, and Skew-*t* copulas, respectively. In order to make a fair comparison among the various models, in Table 5 the adjusted χ^2 statistics are reported. From that table it emerges that the Burr-III copula provides a good fitting. However, the best performers turned out to be the Student's *t* copula and the Skew-*t* copula. It is worth noting that for these two copula models there are not explicit parameters for the interaction of the third order. Nonetheless, these models seems to be the most promising among the considered ones. One possible motivation of this empirical evidence can be the fact that the bivariate

Table 2 Parameter estimates of the Clayton copula

Dataset	k	n	mi	θ	χ^2
(S, C, T)	4^3	1419	0	0.1744	0.4520
(S, C, T)	4^3	1278	100	0.1379	0.4470
(P, C, T)	4^3	475	0	0.1264	0.4959
(P, C, T)	4^3	448	100	0.1356	0.5157
(S + P, C, T)	4^3	1706	0	0.1381	0.4849
(S + P, C, T)	4^3	1548	100	0.1161	0.4600

k The number of classes used to implement the minimum χ^2 method, n the number of observations in the corresponding dataset, mi the threshold of the sources C and T, χ^2 the value of the χ^2-statistic

Table 3 Parameter estimates of the *t* copula

Dataset	k	n	mi	v	ρ_{12}	ρ_{13}	ρ_{23}	χ^2
(S, C, T)	4^3	1419	0	10	0.2837	−0.0552	0.3155	0.3004
(S, C, T)	4^3	1278	100	12	0.2729	−0.0620	0.2702	0.3263
(P, C, T)	4^3	475	0	5	0.2295	−0.1072	0.3746	0.2875
(P, C, T)	4^3	448	100	6	0.2225	−0.1078	0.3870	0.3166
(S + P, C, T)	4^3	1706	0	8	0.2797	−0.1066	0.3703	0.2793
(S + P, C, T)	4^3	1548	100	12	0.2581	−0.1221	0.3245	0.2794

In the table k indicates the number of classes used to implement the minimum χ^2 method; n is the number of observations in the corresponding dataset; mi is the threshold of the sources C and T and χ^2 denotes the value of the χ^2-statistic

Table 4 Parameter estimates of the Skew-t copula

Dataset	k	n	mi	v	ρ_{12}	ρ_{13}	ρ_{23}	α_{12}	α_{13}	α_{23}	χ^2
(S, C, T)	4^3	1419	0	6	0.3039	−0.3390	0.2181	1.3641	0.2082	−0.8461	0.2384
(S, C, T)	4^3	1278	100	9	0.3458	−0.3682	0.1039	1.3092	0.3450	−0.9022	0.2756
(P, C, T)	4^3	475	0	5	0.2185	−0.1414	0.3641	−0.2772	0.0002	0.2393	0.2871
(P, C, T)	4^3	448	100	5	0.2163	−0.1064	0.3848	−0.0108	0.0712	0.0372	0.3189
(S + P, C, T)	4^3	1706	0	8	0.3942	−0.3367	0.2164	1.8326	0.4102	−0.5464	0.2486
(S + P, C, T)	4^3	1548	100	10	0.4269	−0.3357	0.1582	1.8414	0.5545	−0.5293	0.2400

k The number of classes used to implement the minimum χ^2 method, n the number of observations in the corresponding dataset, mi the threshold of the sources C and T, χ^2 the value of the χ^2-statistic

Table 5 Adjusted χ^2 statistic for the four copulas: Clayton, t, Skew-t and Burr-III

Dataset	mi	Adjusted χ^2-statistic			
		Clayton	t	Skew-t	Burr-III
(S, C, T)	0	0.0174	0.0131	0.0119	0.0140
(S, C, T)	100	0.0172	0.0142	0.0138	0.0144
(P, C, T)	0	0.0191	0.0125	0.0144	0.0164
(P, C, T)	100	0.0198	0.0138	0.0160	0.0166
(S + P, C, T)	0	0.0187	0.0122	0.0124	0.0132
(S + P, C, T)	100	0.0177	0.0122	0.0120	0.0124

Table 6 Estimated parameters of the ZD and of the GZD obtained on the dataset (S, C, T)

Comp.	Dist.	k	mi	h	μ	α	θ	γ	χ^2	Adjusted-χ^2
S	ZD	18	0	−3894	23,122	2.8268	2.8554		0.1570	0.0121
S	ZD	18	100	−1.1598	20,195	1.8519	2.4381		0.2082	0.0160
S	GZD	18	0	5890	15,988	0.1972	14.286	−24.279	0.1263	0.0105
S	GZD	18	100	−2883	22,708	2.7587	2.7586	1.4881	0.2047	0.0171
C	ZD	11	0	2.2401	7756	1.3767	2.4386		0.4492	0.0749
C	ZD	11	100	0.2081	6754	1.3426	3.4070		0.4483	0.0747
C	GZD	11	0	2366	6892	0.1259	23.142	−35.83	0.4139	0.0828
C	GZD	11	100	3053	6573	0.2250	25.12	−36.460	0.4328	0.0866
T	ZD	12	0	4.0311	17,567	0.50296	1.3228		0.1363	0.0195
T	ZD	12	100	3.0897	19,287	0.5019	1.2827		0.3005	0.0429
T	GZD	12	0	−881.2	17,928	0.6126	1.6670	−0.3703	0.1100	0.0183
T	GZD	12	100	−632.3	18,057	0.4799	3.3033	−4.2529	0.1258	0.0210

The notation for the parameters is the same adopted in [16] and in [7]

Burr-III copulas implicit in the proposed trivariate extension do not exhibit tail dependence while the Student's t and Skew-t copulas incorporates also this kind of dependence.

Finally, the estimation of the marginal distribution had to be considered. The results obtained by applying the minimum χ^2 method on the ZD and on the GZD are reported in Tables 6, 7 and 8 for the dataset (S, C, T), (P, C, T), and $(S + P, C, T)$, respectively. The ones related to the Dagum distribution are given in Table 9. From these tables it emerges that

Table 7 Parameter estimates of the ZD and of the GZD obtained on the dataset (P, C, T)

Comp.	Dist.	k	mi	h	μ	α	θ	γ	χ^2	Adjusted-χ^2
P	ZD	18	0	2.4419	18,140.54	1.0500	2.41916		0.1739	0.0134
P	ZD	18	100	1.5475	19,018	0.994	0.9948		0.2108	0.0162
P	GZD	18	0	414.89	17897	0.8097	3.6190	−3.0831	0.1713	0.0143
P	GZD	18	100	−1839	20,420	2.4237	0.2516	9.1023	0.1825	0.0152
C	ZD	11	0	−315.21	13,554	2.58816	6.8519		0.2450	0.0408
C	ZD	11	100	1.3314	13,604	4.7552	12.1484		0.1133	0.0189
C	GZD	11	0	1361.75	12,577	0.0571	81.72	−113.36	0.1123	0.0225
C	GZD	11	100	1305	12,811	0.0843	81.5993	−113.36	0.0514	0.0103
T	ZD	12	0	0.3056	16,955	3.4733	123.77		1.7049	0.2436
T	ZD	12	100	−901	15,884	0	20.238		1.5201	0.2172
T	GZD	12	0	−461.2	20,658	0.2963	6.0022	−10.908	0.1927	0.0321
T	GZD	12	100	−65.57	19,720	0.2100	8.9470	−16.519	0.2067	0.0345

The notation for the parameters is the same adopted in [16] and in [7]

Table 8 Parameter estimates of the ZD and of the GZD obtained on the dataset $(S + P, C, T)$

Comp.	Dist.	k	mi	h	μ	α	θ	γ	χ^2	Adjusted-χ^2
S + P	ZD	18	0	0.27	21,639	1.991	3.225		0.2294	0.0176
S + P	ZD	18	100	0.26	21,851	1.8414	2.9153		0.2332	0.0179
S + P	GZD	18	0	5737	18,175	0.1373	24.0782	−38.2197	0.1553	0.0129
S + P	GZD	18	100	4933	19,271	0.2264	16.8936	−27.42	0.1946	0.0162
C	ZD	11	0	−4020	11,973	2.0355	2.5092		0.2748	0.0458
C	ZD	11	100	1.47	9404	2.2456	4.1942		0.2437	0.0406
C	GZD	11	0	164	9758	0.0588	28.3357	−46.77	0.1810	0.0362
C	GZD	11	100	2287	8247	0.1538	43.5449	−62.31	0.1938	0.0388
T	ZD	12	0	0.79	18,627	0.5113	1.2589		0.1933	0.0276
T	ZD	12	100	0.042	18,522	0.5566	1.2803		0.1472	0.0210
T	GZD	12	0	−321	17,938	0.3685	4.0662	−6.437	0.1492	0.0249
T	GZD	12	100	−390	18,583	0.4567	2.8070	−3.780	0.1116	0.0187

The notation for the parameters is the same adopted in [16] and in [7]
h The location parameter added to the original proposal of the ZD and GZD, *k* the number of classes used to implement the minimum χ^2 estimation method

the GZD is outperformed (in terms of Adjusted χ^2) the ZD and the Dagum distribution in 12 cases out of 18. As observed in [7] and [8] and emphasized by the graphical representation in Fig. 1, the over-performance of the GZD is particularly evident in the cases in which the left tail of the distribution has an irregular behavior. Finally, in order to get a look to the shape of the estimated joint distributions, in Figs. 2, 3 and 4 the plots of the estimated bivariate densities implicit in the dataset $(S + P, C, T)$ are provided.

Table 9 Parameter estimates of the Dagum distribution obtained on the dataset (S, C, T),(P, C, T) and $(S + P, C, T)$ respectively

Comp.	k	mi	h	λ	θ	ϵ	χ^2	Adjusted-χ^2
S	18	0	−2888	640	1.3840	1078	0.5563	0.0428
S	18	100	−2939	881	1.4257	1186	0.5443	0.0419
C	11	0	−5393	3737	1.863	6539	0.4840	0.0807
C	11	100	−490	754	1.5785	1045	0.4002	0.0667
T	12	0	−10408	12,536	1.9439	12,598	0.3617	0.0517
T	12	100	−27446	67,460	1.9553	11,676	0.6829	0.0976
P	18	0	−11813	500,838	2.4561	77,794	0.2352	0.0181
P	18	100	−21737	178,982,734	3.4612	19,787,024	0.2411	0.0185
C	11	0	−667	745	1.3333	168.25	0.3143	0.0524
C	11	100	−1820	586	1.7689	16614	0.1301	0.0217
T	12	0	−39525	$10e^{10}$	4.3011	$3.5e^{10}$	0.3131	0.0447
T	12	100	−18926	16,039	1.7805	5703	0.6013	0.0859
S + P	18	0	−17069	$3181\ e^{16}$	4.360	1.708	0.2160	0.0166
S + P	18	100	−16265	$1883e^{16}$	4.305	1.557	0.2245	0.0173
C	11	0	−3428	$6987e^{7}$	2.702	0.894	0.3347	0.0558
C	11	100	−2436	$2696e^{7}$	2.705	1.862	0.2202	0.0367
T	12	0	−25675	$1208e^{25}$	6.105	0.744	0.1481	0.0212
T	12	100	−26680	$1168e^{26}$	6.293	0.736	0.1271	0.0182

h The location parameter added to the original proposal of the ZD and GZD, k the number of classes used to implement the minimum χ^2 estimation method

5 Applications to wine data

The results obtained on income data could discourage the application of the Trivariate Burr-III copula. However it is not difficult to find trivariate real data in which Trivariate Burr-III copula performs substantially better than the t and Skew-t copulas. In this section we present such an application. The data here considered concern the concentrations of 13 different chemicals in wines grown in the same region in Italy that are derived from three different cultivars [11]. The chemicals are: Alcohol (Al), Malic acid (M), Ash (As), Alcalinity of ash (AA), Magnesium (Mag), Total phenols (TP), Flavanoids (F), Nonflavanoid phenols (NP), Proanthocyanins (Pa), Color intensity (CI), Hue (H), OD280/OD315 of diluted wines (OD) and Proline (Pl) . The sample is composed by $n = 178$ units. Similarly to the previous section, t, Skew-t and Trivariate Burr-III copulas are fitted to all the possible $\binom{13}{3}$ trivariate data implicit in the original 13-variate data-set. Similarly to the previous section the parameters of these copulas are estimated by using the minimum χ^2 method with 3 classes for each marginal distribution. The estimation of the marginal distributions is not performed here since the aim of the present section is to show the potentiality of the Trivariate Burr-III copula.

The results obtained over the $\binom{13}{3}$ trivariate data-set clearly exhibit the potential usefulness of the Trivariate Burr-III copula: in almost all cases it provides the better goodness-of-fit. Moreover, the parameter estimates obtained on these data are characterized by the fact that, usually, estimates of the third-order interaction parameter are substantially null. For this reason the reduced Trivariate Burr-III copula (in which the parameter α_{123} is set equal to 0)

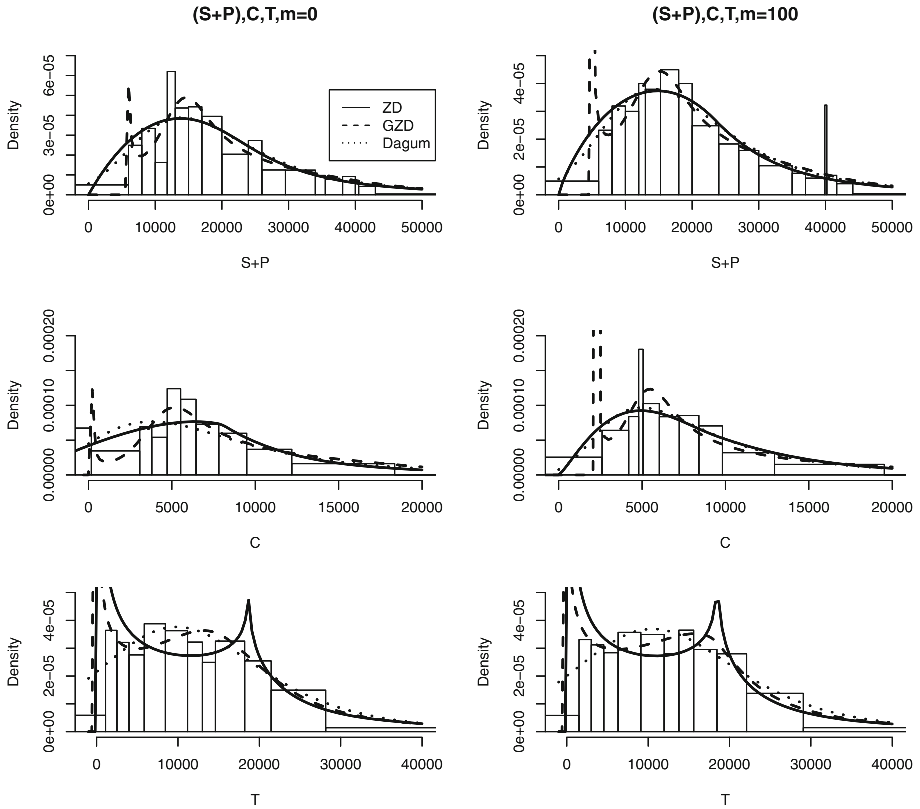

Fig. 1 Histogram of the marginal distributions of the dataset $(S + P, C, T)$ and graphical representation of the estimated ZD, GZD and Dagum distributions

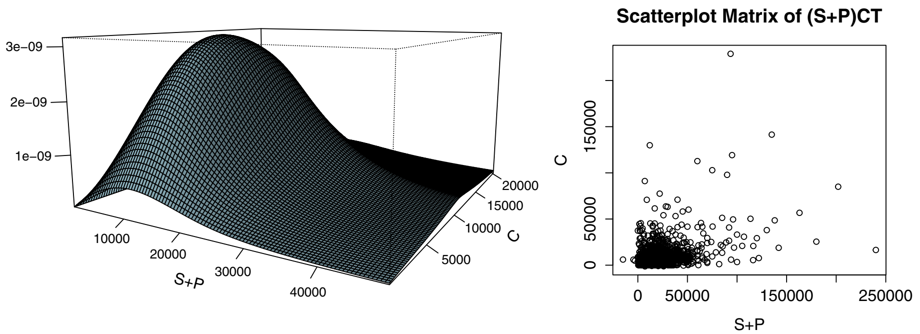

Fig. 2 Bivariate density estimate and scatterplot of observed data of component $(S + P, C)$ in the data-set $(S + P, C, T)$

is estimated too and the obtained results confirm the very good performance obtained with the full model.

Below, we do not report the result obtained over all the $\binom{13}{3}$ possible trivariate data-set but in Table 10 we provide only 9 of the most representative results.

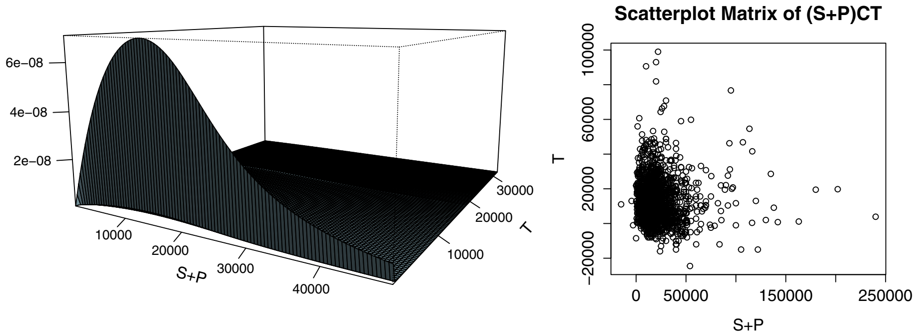

Fig. 3 Bivariate density estimate and scatterplot of observed data of component $(S + P, T)$ in the data-set $(S + P, C, T)$

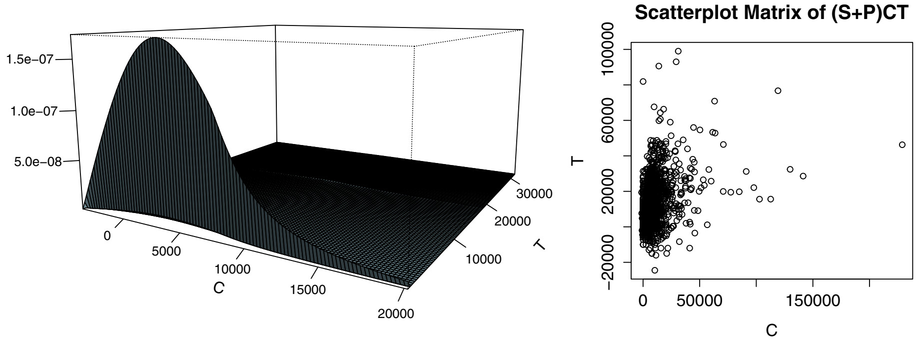

Fig. 4 Bivariate density estimate and scatterplot of observed data of component (C, T) in the data-set $(S + P, C, T)$

6 Conclusion and further research

An extension of the Burr-III copula to the trivariate case is proposed and applied to an income dataset by considering three income components and the ZD [16], the GZD [7], and the Dagum as marginal distributions. A comparison with the Clayton, t, and Skew-t copulas shows that the last two models perform better in all the cases, probably because they possess a certain degree of tail dependence, which Burr-III has not. However, it is not difficult to find trivariate real data in which Trivariate Burr-III copula performs substantially better than the t and Skew- t copulas, as shown by the applications on wine data presented in this paper.

There are several properties of the Burr-III copula that are to be deepened in future work, mainly concerning theoretical aspects. Specifically, it would be important to deepen the result of Lemma 1 by providing the full characterization of the parametric space. This result is in fact essential in order to implement an estimation algorithm providing parameters estimates that certainly correspond to well defined distributions. A further interesting topic that can be explored in future research is the extension to the multivariate context of the Trivariate Burr-III copula. Indeed, similarly to the trivariate case, starting from the independence case it is possible to introduce the following d-variate function:

Table 10 Parameter estimates and goodness-of-fit of the Burr-III, t, and Skew-t copula on wine data

Data	α_{12}	α_{13}	α_{23}	α_{123}	ϵ	χ^2	χ^2_{adj}
Burr-III copula							
Al As Pa	0.4481	≈0	0.4502	≈0	0.5573	0.2648	0.0883
AA Mag OD	5.8861	7.234	1.5182	≈0	10.4183	0.3112	0.1037
Al NP CI	1.4512	0.0251	1.0124	≈0	0.4512	0.3259	0.1086
M F OD	1.3989	1.3980	≈0	≈0	0.3980	0.5941	0.1980
M Pa CI	1.9035	0.2773	1.0066	≈0	0.9035	0.4424	0.1475
As F H	0.7782	0.8522	≈0	≈0	0.6154	0.3435	0.1145
As F OD	0.7826	1.0098	≈0	≈0	0.3540	0.3848	0.1283
As H OD	0.8902	1.1479	≈0	≈0	0.8281	0.4433	0.1478
AA CI OD	2.6859	5.0632	5.3434	≈0	8.005	0.4110	0.1367
Reduced Burr-III copula							
Al As Pa	0.4482	≈0	0.4503	–	0.5573	0.2648	0.0662
AA Mag OD	5.8947	7.2437	1.5209	–	10.4370	0.3112	0.0778
Al NP CI	1.5197	≈0	1.0906	–	0.5197	0.3332	0.0833
M F OD	1.3982	1.3982	≈0	–	0.3982	0.5941	0.1485
M Pa CI	2.5149	≈0	1.3660	–	1.5149	0.4607	0.1152
As F H	0.7783	0.8521	≈0	–	0.6158	0.3435	0.0859
As F OD	0.7826	1.0098	≈0	–	0.3541	0.3848	0.0962
As H OD	0.8899	1.1478	≈0	–	0.8284	0.4433	0.1108
AA CI OD	2.6858	5.0645	5.3456	–	8.0070	0.4110	0.1028

Data	v	ρ_{12}	ρ_{13}	ρ_{23}	χ^2	χ^2_{adj}
t copula						
Al As Pa	46	0.2873	0.7125	0.2687	0.3204	0.0801
AA Mag OD	52	−0.2726	−0.3481	−0.0337	0.3822	0.0956
Al NP CI	49	−0.1883	0.6962	0.0274	0.3893	0.0973
M F OD	52	−0.4222	−0.3765	0.7894	0.6086	0.1521
M Pa CI	48	−0.3199	0.3341	0.0382	0.4908	0.1227
As F H	53	0.0422	−0.0225	0.6276	0.4276	0.1069
As F OD	46	0.0.575	−0.0782	0.7896	0.4993	0.1248
As H OD	47	0.03571	−0.0873	0.4984	0.5435	0.1359
AA CI OD	7	−0.0959	−0.3049	−0.3582	0.5049	0.1262

Data	α_{12}	α_{13}	α_{23}	v	ρ_{12}	ρ_{13}	ρ_{23}	χ^2	χ^2_{adj}
Skew-t copula									
Al As Pa	0.0049	−0.1667	0.0461	53	0.2851	0.7097	0.2698	0.3188	0.3188
AA Mag OD	0.0429	0.0706	0.0904	52	−0.2711	−0.3517	−0.0370	0.3792	0.3792
Al NP CI	0.0021	2.8271	0.4688	48	−0.1348	0.6537	0.2647	0.3752	0.3752
M F OD	0.1702	0.0032	0.0368	53	−0.4176	−0.3760	0.7888	0.6062	0.6062
M Pa CI	0.0884	−0.1167	0.0538	53	−0.3235	0.3350	0.0390	0.4904	0.4904
As F H	0.5611	0.0312	−0.0392	53	0.0435	−0.0206	0.6223	0.4269	0.4269
As F OD	0.4458	−0.0333	−0.0884	52	0.0373	−0.1125	0.7801	0.4971	0.4971
As H OD	3.4029	0.5674	0.3054	48	0.3381	0.1293	0.5255	0.5235	0.5235
AA CI OD	1.4354	−1.901	0.5299	7	−0.4618	−0.0688	−0.4762	0.4463	0.4463

$$C(u_1, ..., u_d) = \left[1 + \sum_{i=1}^{d} (u_i^{-\frac{1}{\epsilon}} - 1) + \sum_{i<j} \alpha_{ij} (u_i^{-\frac{1}{\epsilon}} - 1)(u_j^{-\frac{1}{\epsilon}} - 1) \right.$$
$$\left. + \sum_{i<j<k} \alpha_{ijk} (u_i^{-\frac{1}{\epsilon}} - 1)(u_j^{-\frac{1}{\epsilon}} - 1)(u_k^{-\frac{1}{\epsilon}} - 1) + \cdots + \alpha \prod_{i=1}^{d} (u_i^{-\frac{1}{\epsilon}} - 1) \right]^{-\epsilon}. \tag{16}$$

This function possesses $2^d - d - 1$ interaction parameters. These are given by $\binom{d}{2}$ second-order interaction parameters α_{ij}, $\binom{d}{3}$ third-order interaction parameters α_{ijk} ... etc. If the value of these parameters are appropriately chosen function (16) is a copula that can be applied to fit multivariate phenomena. However, the number of parameters in (16) is very high and, in practice, it could be sufficient to consider only the interaction parameters up to the third order obtaining the model:

$$C(u_1, ..., u_d) = \left[1 + \sum_{i=1}^{d} (u_i^{-\frac{1}{\epsilon}} - 1) + \sum_{i<j} \alpha_{ij} (u_i^{-\frac{1}{\epsilon}} - 1)(u_j^{-\frac{1}{\epsilon}} - 1) \right.$$
$$\left. + \sum_{i<j<k} \alpha_{ijk} (u_i^{-\frac{1}{\epsilon}} - 1)(u_j^{-\frac{1}{\epsilon}} - 1)(u_k^{-\frac{1}{\epsilon}} - 1) \right]^{-\epsilon}. \tag{17}$$

Intuition suggests that the conditions provided by Lemma 1 could be applied to model (17) too. However, a preliminary analysis shows that the admissible values of each third-order interaction parameter α_{ijk} depend on all the second-order interaction parameters and not only on ϵ, α_{ij}, α_{ik}, and α_{jk}. It will be very hard to exploit a full characterization of the parametric space for model (17) but, if the results obtained in the future research on the trivariate case will be encouraging, it will be certainly worth to continue the study of the multivariate model.

References

1. Arcagni, A., Porro, F.: On the parameters of Zenga distribution. Stat. Method Appl. **22**(3), 285–303 (2013)
2. Arcagni, A., Zenga, M.M.: Application of Zenga's distribution to a panel survey on household incomes of european member states. Statistics & Applications X **I**(1), 79–102 (2013)
3. Azzalini, A.: The Skew-Normal and Related Families. IMS Monographs Series. Cambridge University Press, Cambridge (2014)
4. Burr, I.W.: Cumulative frequency functions. Ann. Math. Stat. **13**(2), 215–232 (1942)
5. Banca di Italia (2012) Survey on household income and wealth 2012
6. Dagum, C.: A new model of personal income distribution: specification and estimation. Econ. Appl. **XXX**(3), 413–437 (1977)
7. De Capitani, L., Zini, A.: The confluent hypergeometric-mixture of Polisicchio distributions: a generalized Zenga model. Statistics & Applications **XI**(1), 23–48 (2013a)
8. De Capitani, L,. Zini, A.: A comparison among two generalized Beta-Mixture of Polisicchio distributions and the Zenga model. Statistics & Applications **XI**(2), 193–230 (2013b)
9. Domma, F.: Some properties of the bivariate Burr type III distribution. Statistics **44**(2), 23215 (2010)
10. Durrleman, V., Nikeghbali, A., Roncalli, T.: Which Copula is the Right One? Groupe de Recherche Operationnelle Credit Lyonnais, France (2000)
11. Forina, M. et al.: PARVUS—An Extendible Package for Data Exploration, Classification and Correlation. Institute of Pharmaceutical and Food Analysis and Technologies, Via Brigata Salerno, vol. 16147 Genoa, Italy (1988)
12. Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)

13. Joe, H., Xu, J.J.: The estimation method of inference functions for margins for multivariate models, Department of Statistics, University of British Columbia, Technical Report, 166 (1996)
14. Kleiber C (2008) A guide to the Dagum distribution. In: Modelling Income Distributions and Lorenz Curves: Essays in Memory of Camilo Dagum, pp. 97–119. Springer, New York
15. Rodriguez, R.N.: Multivariate Burr-III distributions, Part I. Theoretical Properties. Research Publication GMR-3232, General Motors Research Laboratories, Warren, Michigan (1980)
16. Zenga, M.: Mixture of Polisicchio's truncated Pareto distributions with Beta weights. Statistics & Applications **VIII**(1), 3–25 (2010)