Measuring dependence between dimensions of poverty in Spain: An approach based on copulas

Ana Pérez[1], Mercedes Prieto[2]

[1]Universidad de Valladolid
[2]Universidad de Valladolid

Abstract

Welfare and close related issues, like poverty and inequality, are multidimensional as they involve not only income, but also education, health or labour. This paper aims to measure the dependence among dimensions using copula-based coefficients. This approach focuses on the positions of the individuals across dimensions, rather than their specific values. We apply copula-based orderings of dependence to analyze how the concordance among dimensions has evolved in Spain over the last years. We also compute multivariate extensions of the Spearman's rank coefficient. The variables considered are those included in the AROPE rate: income, material needs and work intensity.

Keywords: copulas, concordance, Spearman, orthant dependence, poverty.

1. Motivation

There is a general agreement that welfare and closed related issues, like poverty and inequality, are multidimensional encompassing not only income, but also non-monetary dimensions such as education and health; see, for instance, Kolm (1977), Atkinson and Bourbignon (1982) and Sen (1985). In this framework, much attention has been driven to the selection of the dimensions involved (Ramos and Silber, 2005 and Anand *et al.*, 2009) and the way of aggregating the information contained in such dimensions (Maasumi, 1986; Tsui, 1995; Bourguignon and Chakravarty, 2003; Atkinson, 2003 and Alkire and Foster, 2011). Moreover, the level of welfare, poverty and inequality depends on the degree of dependence among dimensions (Bourbignon and Chakravarty, 2003 and Duclos *et al.*, 2006). However, the problem of measuring that dependence has been scarcely addressed in the literature and this is the main goal of this paper.

The information about the multivariate dependence among dimensions can be expressed either in terms of their joint distribution function or in terms of their implied copula. The copula approach focuses on the positions of the individuals across dimensions, rather than the specific values that the corresponding variables attain for such individuals. The advantage of this approach for continuous variables is that it enables the decomposition of the joint distribution function into its univariate marginals and the dependence structure captured by the copula. Moreover, copulas allow building

scaled-free measures of dependence and orderings of multivariate dependence; see Nelsen (2006) and Joe (2015) for a comprehensive review of copulas and related concepts.

In the bivariate case, there are several measures of dependence. The most well-known is Pearson's correlation coefficient, which is suitable for linear dependence and elliptical data. However, empirical research shows that, in welfare economics, the data seldom belongs to this class. In this case, measures based on ranks, such as Spearman's rho or Kendall's tau, would be more appropriate. These two coefficients measure a form of dependence known as concordance. Noticeably, they can be written in terms of copulas.

In a multivariate framework, neither the concept of concordance nor the generalization of the bivariate coefficients of concordance is unique. For instance, in the trivariate case, there are more than eight copula-based generalizations of Spearman's rho; see García *et al.* (2013) and the references therein.

In welfare contexts, copula-based methods have been recently employed by Quinn (2007a, 2007b) and Bø *et al.* (2012) in a two-dimensional setting and by Decancq (2013) in a multi-dimensional framework. This author introduces a copula-based ordering of dependence and characterizes two measures of multivariate dependence consistent with this order: the multivariate Kendall's tau proposed by Nelsen (1996) and a multivariate version of Spearman's rho in Nelsen (2002). He illustrates the results with a Russian data set including household income, self-assessed health and years of schooling.

In this paper, we extend the results in Decancq (2013) by considering: a) other two multidimensional versions of Spearman's rho, due to Wolff (1980), Joe (1990) and Nelsen (1996); b) the coefficients based on directional dependence recently proposed by Nelsen and Úbeda-Flores (2012) and García *et al* (2013) for trivariate distributions. These measures are capable of revealing some forms of dependences that the coefficients analysed in Decancq (2013) fail to detect. We also perform comparisons based on positive orthant dependent orderings using bootstrap methods.

Our empirical application is devoted to measuring how the dependence between the three dimensions included in the AROPE (At Risk Of Poverty or social Exclusion) rate, i.e. income, material needs and work intensity, has evolved in Spain over the last years. We focus on this

rate because it is the headline indicator to monitor and implement effective poverty-reduction policies in the framework of the EU 2020 Strategy.

The rest of the paper is organized as follows. Section 2 introduces the copula and summarizes its basic properties. Section 3 describes the orthant dependence orderings and the way to implement them in practise. Section 4 introduces ten multivariate copula-based measures of dependence and discusses its main properties. It also introduces its sample versions. Section 5 illustrates how these tools can be used to measure the evolution of the dependence between dimensions of poverty in Spain over the last years.

2. Preliminaries and notation

According to Nelsen (2006), a d-dimensional copula C is a function $C: I^d \rightarrow I$, with $I=[0,1]$, with the following properties:

(i) For every $u=(u_1,...,u_d)$ in I^d, $C(u) = 0$ if at least one coordinate of u is 0, and $C(\mathbf{u}) = u_j$ if all coordinates of $\mathbf{u}$ are 1 except u_j.

(ii) C is d-increasing, that is for every $\mathbf{u_1}=(u_{11},...,u_{d1})$ and $\mathbf{u_2}=(u_{12},...,u_{d2})$ in I^d such that $u_{j1} \leq u_{j2,}$ for all $j=1,...d$

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+....+i_d} C\left(u_{1i_1},....,u_{di_d}\right) \geq 0$$

From the statistical point of view, the importance of copulas comes up in the Sklar's theorem. This theorem establishes that, if $X=(X_1,...,X_d)$ is a d-dimensional random variable with joint distribution function F and marginals $F_1,…,F_d$, then, for all $x=(x_1...x_d)\in\Re^d$, F has a copula representation given by:

$$F\left(x\right) = F\left(x_1,...,x_d\right) = C\left[F\left(x_1\right),...,F\left(x_d\right)\right] \qquad (1)$$

If $F_1,…,F_d$ are all continuous, then C is unique; otherwise C is uniquely determined on $\mathrm{Ran}F_1\times\cdots\times\mathrm{Ran}F_d$. Conversely, if C is a d-copula and $F_1,…,F_d$ are univariate distribution functions, then the function F defined in (1) is a d-dimensional distribution function with margins $F_1,…,F_d$.

Moreover, when F has continuous margins $F_1,…,F_d$ and copula C satisfies (1), then, for any u in I^d,

$$C\left(u\right) = C\left(u_1,...,u_d\right) = F\left[F_1^{-1}\left(u_1\right),...,F_d^{-1}\left(u_d\right)\right] \qquad (2)$$

where F_j^{-1}, for $j=1,...,d$, denotes the inverse of F_j. The copula C itself, as defined in (2), is a multivariate distribution whose univariate marginals are all uniform on the interval (0,1).

Each copula function C is bounded by its so-called Fréchet-Hoeffding bounds, $W(u) \leq C(u) \leq M(u)$, where

$W(u)=max(u_1+...+u_d-d+1, 0)$ and $M(u)=min(u_1,...,u_d)$. M is always a copula but W is a copula only if $d=2$. Taylor (2007) points out that M can be thought of as a state of "maximal concordance", i.e. the state where each component X_i of X is an almost surely increasing function of every other component X_j. Unlike, if X is a vector of independent random variables, then its copula is the independent copula Π defined as $\Pi(u)=u_1\cdots u_d$.

Another important function associated with a copula C is the *survival function* $\overline{C}$ defined as:

$$\overline{C}(\mathbf{u})=p(\mathbf{U}>\mathbf{u})= p(U_1>u_1,…,U_d>u_d).$$

In general, $\overline{C}$ is not a copula.

3. Copula-based orderings

The concept of multivariate dependence can be defined in different ways (see, for example, Mari and Kotz, 2001). The one we handle in this paper is *positive dependence*, that is, when large (small) values in one component tend to be associated with large (small) values in the other components. Following Nelsen (2006) and Decancq (2013) we consider three copula-based orderings of dependence defined as follows.

Let X and Y be two d-dimensional random variables with copula functions C^X and C^Y, respectively, and with survival functions $\overline{C}^X$ and $\overline{C}^Y$, respectively.

- X is more positively lower-orthant dependent (*PLOD*) than Y if

$$C^X\left(u\right) \geq C^Y\left(u\right) \text{ for every } \mathbf{u} \text{ in } \mathbf{I^d} \qquad (3)$$

- X is more positively upper-orthant dependent (*PUOD*) than Y if

$$\overline{C}^X\left(u\right) \geq \overline{C}^Y\left(u\right) \text{ for every } \mathbf{u} \text{ in } \mathbf{I^d} \qquad (4)$$

- X is more positively orthant dependent (*POD*) than Y if both (3) and (4) hold[1].

The three definitions above were already introduced in Joe (1997) in terms of cumulative distribution functions. As this author points out, the expressions (3) and (4) mean that the components of X are more likely simultaneously to have small and large values, respectively, compared with the components of Y.

In the bivariate case, the three definitions above are equivalent.

In practise, the PLOD, PUOD and POD conditions may be checked once empirical versions of copulas and survival functions are known. To introduce these concepts,

---

[1] Note that this notion of *POD* is equivalent to the definition of C^X being *more concordant* than C^Y in Dolati and Úbeda-Flores (2006).

further notation is needed. Let $X_i=(X_{i1},...,X_{id})$, $i=1,...,n$, be a random sample of size n from the d-dimensional continuous random variable X. Its marginal distribution functions are estimated by their empirical counterparts:

$$\hat{F}_j(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_{ij} \leq x\}}, \text{ for } j=1...d \text{ and } x \in \Re$$

where $1_{\{A\}}$ denotes the indicator function on a set A. Let R_{ij} be the rank of X_{ij} among $\{X_{1j},...,X_{nj}\}$, with $i = 1,...,n$ and $j = 1,...,d$, and define $\overline{R}_{ij} = n+1-R_{ij}$. The copula C is estimated by

$$\hat{C}(\boldsymbol{u}) = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} 1_{\{\hat{U}_{ij} \leq u_j\}}, \text{ for } \mathbf{u}=(u_1,...,u_d) \text{ in } \boldsymbol{I^d}$$

where $\mathbf{1}_A$ denotes the indicator function of a set A and $\hat{U}_{ij} = \hat{F}_j(X_{ij}) = \frac{R_{ij}}{n}$. The empirical survival function is defined in a similar way. Fermanian *et al.* (2004) establish regularity conditions for $\hat{C}$ to converge towards a Gaussian process. As the asymptotic distribution is unfeasible in most cases, bootstrap methods are commonly used to approximate it.

Conditions (3) and (4) can then be checked out by comparing the values of both the empirical copula and the empirical survival function on the two d-dimensional distributions X and Y being compared. For this comparison to be feasible, a grid of points on the hypercube $\boldsymbol{I^d}$ should be defined. Let $0<p_1<...<p_k<1$ be k points in $\boldsymbol{I}$, the k^d points in $\boldsymbol{I^d}$ that make up the grid are $\left(p_{i_1},...,p_{i_d}\right)$, with $i_j=1,...,k$ and $j=1,...,d$. At each of these points, the values of the empirical copula of both X and Y are computed and it is checked if one of the copulas outranks the other at all points in the grid. The same procedure is carried out with the empirical survival function.

Obviously, this comparison is very restrictive as it is merely descriptive. However, the procedure can be significantly improved by bootstrapping. In doing so, we will be able to formally test the hypothesis $H_o : \boldsymbol{C^X} \geq \boldsymbol{C^Y}$ (PLOD condition). Similarly, the equivalent hypothesis for the survival function, i.e., the PUOD condition, could also be tested.

4. Copula-based multivariate measures of dependence

The concordance ordering defined in the preceding section produces an incomplete ranking. In contrast, a dependence measure allows a complete ranking of all the distributions being compared. In this section, we first discuss three dependence measures which can be regarded as multivariate generalizations of the well-known Spearman's rho. These measures are related to the three multivariate dependence ordering conditions

introduced in Section 3. A complete description of these and other copula-based measures of multivariate dependence can be found in Wolff (1980), Joe (1990), Nelsen (1996, 2002), Dolati and Úbeda-Flores (2006) and Schmid and Schmidt (2007). The problem of estimating such measures is addressed in Joe (1990), Schmid and Schmidt (2007) and Pérez and Prieto (2015).

The first copula-based multivariate extension of Spearman's rho that we consider, ρ_d^-, is due to Wolff (1980) and Nelsen (1996) and it is defined as follows:

$$\rho_d^- = \frac{\int_{I^d} C(\boldsymbol{u})d\boldsymbol{u} - \int_{I^d} \Pi(\boldsymbol{u})d\boldsymbol{u}}{\int_{I^d} M(\boldsymbol{u})d\boldsymbol{u} - \int_{I^d} \Pi(\boldsymbol{u})d\boldsymbol{u}} \tag{5}$$

Since the denominator of this expression represents the maximum value of its own numerator, i.e. its value at the maximal copula $C = M$, it is guaranteed that the maximum value of ρ_d^- is 1. By working out some of the integrals in (5), the following alternative expression turns out:

$$\rho_d^- = \frac{(d+1)}{2^d-(d+1)}\left[2^d\int_{I^d} C(\boldsymbol{u})d\boldsymbol{u} - 1\right] = \frac{(d+1)}{2^d-(d+1)}\left[2^d\int_{I^d} \overline{\Pi}(\boldsymbol{u})dC(\boldsymbol{u}) - 1\right]$$

where $\overline{\Pi}(\boldsymbol{u}) = \prod_{j=1}^{d}\left(1-u_j\right)$. Following Nelsen (1996), ρ_d^- can be regarded as a multivariate measure of average lower orthant dependence.

The second multivariate version of Spearman's rho considered in this paper, ρ_d^+, was originally proposed by Nelsen (1996) as a multivariate measure of average upper orthant dependence and it is defined as follows:

$$\rho_d^+ = \frac{\int_{I^d} \Pi(\boldsymbol{u})dC(\boldsymbol{u}) - \int_{I^d} \Pi(\boldsymbol{u})d\boldsymbol{u}}{\int_{I^d} \Pi(\boldsymbol{u})dM(\boldsymbol{u}) - \int_{I^d} \Pi(\boldsymbol{u})d\boldsymbol{u}} \tag{6}$$

Again, the denominator of this expression resembles its own numerator evaluated at the maximal copula $C = M$ and so the maximum value of ρ_d^+ is 1. The expression (6) can be alternatively written as:

$$\rho_d^+ = \frac{(d+1)}{2^d-(d+1)}\left[2^d\int_{I^d} \Pi(\boldsymbol{u})dC(\boldsymbol{u}) - 1\right]$$

When the copula of X is the upper bound M, both ρ_d^- and ρ_d^+ attain their maximum value, 1, and they become zero when the components of X are independent, i.e. when $C = \Pi$. A lower bound for both ρ_d^- and ρ_d^+ is $\left[2^d-(d+1)!\right]/\left\{d!\left[2^d-(d+1)\right]\right\}$; see Nelsen (1996).

Nelsen (2002) proposes another multivariate version of Spearman's rho, ρ_d, defined as the average of ρ_d^- and ρ_d^+, i.e.:

$$\rho_d = \frac{\rho_d^- + \rho_d^+}{2} = \frac{(d+1)}{2^d - (d+1)}\left\{2^{d-1}\left[\int_{I^d} C(\boldsymbol{u})d\Pi(\boldsymbol{u}) + \int_{I^d}\Pi(\boldsymbol{u})dC(\boldsymbol{u})\right] - 1\right\}$$

(7)

This measure is further discussed in Dolati and Úbeda-Flores (2006), where it is shown that it is a measure of average orthant dependence. It has also been used in Decancq (2013).

In the bidimensional case (d=2) the three coefficients defined above reduce to bivariate Spearman's rho, ρ_S, that is, $\rho_2^- = \rho_2^+ = \rho_2 = \rho_S$.

In the trivariate case (d=3), the three coefficients defined above become:

$$\rho_3^- = 8\int_{I^3} C(u_1,u_2,u_3)du_1du_2du_3 - 1$$ (8a),

$$\rho_3^+ = 8\int_{I^3} \overline{C}(u_1,u_2,u_3)du_1du_2du_3 - 1$$ (8b),

$$\rho_3 = \frac{\rho_3^- + \rho_3^+}{2} = \frac{\rho_{12} + \rho_{13} + \rho_{23}}{2}$$ (8c),

where ($\rho_{12}, \rho_{13}, \rho_{23}$) are the three possible pairwise Spearman's rho coefficients.

The advantage of ρ_3^- and ρ_3^+ is that they are capable of revealing some forms of dependences that ρ_3 fails to detect. Nelsen (1996) illustrates this feature with an example where $\rho_3 = 0$, presumably indicating no dependence at all, while ρ_3^+ is positive and ρ_3^- is negative, indicating some degree of positive upper orthant dependence and negative orthant dependence, that is missed by ρ_3.

In spite of this advantage, there are still some forms of multivariate dependence that the coefficients ρ_3^- and ρ_3^+ may fail to detect when they take values near 0. To overcome this drawback, a general class of directional ρ-coefficients has recently been proposed by Nelsen and Úbeda-Flores (2012). Let ($\alpha_1,\alpha_2,\alpha_3$), with α_i={1,-1}, denote the eight vertexes of the cube $\boldsymbol{I^3}$ determining the eight directions in which we could measure dependence in the 3-dimensional case. For each direction ($\alpha_1,\alpha_2,\alpha_3$), a directional ρ-coefficient is defined as:

$$\rho_3^{\alpha_1\alpha_2\alpha_3} = \frac{\alpha_1\alpha_2\rho_{12} + \alpha_1\alpha_3\rho_{13} + \alpha_2\alpha_3\rho_{23}}{3} + \alpha_1\alpha_2\alpha_3\frac{\rho_3^+ - \rho_3^-}{2}$$ (9)

Hence, each of these eight directional coefficients is a simple linear combination of the pairwise measures and the two measures ρ_3^- and ρ_3^+ of trivariate association. Noticeably, $\rho_3^{(1,1,1)} = \rho_3^+$ and $\rho_3^{(-1,-1,-1)} = \rho_3^-$.

Finally, we consider the new index of maximal dependence $\rho_3^{\max}$ proposed by García et al. (2013) as the largest of the eight directional ρ-coefficients defined in (9). These authors proved that their index can also be calculated using the three pairwise Spearman's rho coefficients and the three common 3-dimensional versions of Spearman's rho in (8) as follows:

$$\rho_3^{\max} = \frac{2}{3}\max\left\{\rho_{12},\rho_{13},\rho_{23},3\rho_3\right\} - \min\left\{\rho_3^-,\rho_3^+\right\}$$ (10)

The non-parametric estimator of the coefficient ρ_d^+ in (6) was already given in Joe (1990). For d=3, this estimator becomes:

$$\hat{\rho}_3^+ = \frac{8}{n(n+1)(n^2-1)}\sum_{i=1}^n R_{i1}R_{i2}R_{i3} - \frac{n+1}{n-1}$$

(11)

Pérez and Prieto (2015) propose a non-parametric estimator of ρ_d^-, which reduces to the following expression for d=3:

$$\hat{\rho}_3^- = \frac{8}{n(n+1)(n^2-1)}\sum_{i=1}^n \overline{R}_{i1}\overline{R}_{i2}\overline{R}_{i3} - \frac{n+1}{n-1}$$

(12)

Note that Schmid and Schmidt (2007) propose other nonparametric estimators of ρ_d^- and ρ_d^+ but, as Pérez and Prieto (2015) show up, these are unfeasible, since they can take values out of the parameter space. Note also that there is a *typo* in García et al. (2015), since their expressions for the estimators $\hat{\rho}_3^-$ and $\hat{\rho}_3^+$ are the other way round, being their equation (12) the correct expression of $\hat{\rho}_3^+$ and their equation (13) the proper expression of $\hat{\rho}_3^-$.

Finally, the parameter ρ_3 in (8c) can be estimated as:

$$\hat{\rho}_3 = \frac{\hat{\rho}_3^- + \hat{\rho}_3^+}{2} = \frac{\hat{\rho}_{12} + \hat{\rho}_{13} + \hat{\rho}_{23}}{2}$$

For the bidimensional case (d=2), all expressions above collapse to the well-known sample bivariate Spearman's rank correlation.

To estimate the directional coefficients in (9) and the index of maximal dependence in (10), García et al. (2013) suggest using plug-in estimators, that is, replacing each pairwise correlation by its well known sample counterpart and replacing ρ_3^- and ρ_3^+ by their non-parametric estimators in (12) and (11), respectively.

5. Empirical application

Several recent works (Atkinson and Morelli, 2011; Foessa, 2013 and Prieto *et al.* 2015) point out that poverty in Spain has increased significantly over the last years of economic crisis. As we said before, poverty depends on the degree of interdependence among the poverty dimensions. In this section, we apply the copula-based methods described in Section 3 and Section 4 to measure the evolution of the dependence among the poverty dimensions in Spain over the period 2009-2013. This analysis will give us a general picture of how the degree of poverty has evolved in Spain in the years after the crisis.

5.1. Data and variables

The dimensions of poverty we have selected for our empirical application are those included in the AROPE rate[2], namely income, material needs and work intensity. We select these variables because the AROPE rate is the headline indicator to monitor and implement effective poverty-reduction policies in the framework of the Europe 2020 Strategy.

The income of a household is the total income of the household, after tax and other deductions divided by the equivalent household size[3].

The material deprivation is originally defined as the enforced inability to: 1) pay unexpected expenses; 2) afford a one-week annual holiday away from home; 3) have a meal involving meat, chicken or fish every second day: 4) heat adequately a dwelling; 5) have a washing machine; 6) have color television; 7) have telephone; 8) have car; 9) manage with payment arrears (mortgage or rent, utility bills, hire purchase instalments or other loan payments). For the sake of simplicity and easy of interpretation, we consider a transformation of this variable that takes its complementary values, namely: 0 (having all the 9 possible deprivations), 1 (having eight out of the nine aforementioned deprivations), …, 9 (having no deprivations). Hence, this new variable indicates the number of no-privations out of the nine possible.

The work intensity of a household is the ratio of the total number of months that all working-age[4] household members have worked during the year and the total number of months the same household members theoretically could have worked in the same period.

With our transformation of the variable measuring material needs we ensure that the three variables considered keep the same relationship with phenomenon being measured, poverty. That is, high values of each variable convey lower chance to be poor, while low values of each variable convey higher chance to be poor.

The data comes from the EU-Statistics on Income and Living Conditions (EU-SILC) survey, which is the EU reference source for comparative statistics on income distribution and social inclusion at the European level. To face our goal, we base our analysis on data from surveys launched in 2009 and 2013.

The unit of analysis is the household. The samples sizes of the EU-SILC survey are 13360 and 12139 in 2009 and 2013, respectively. However, we only work with subsamples of 10863 households in 2009, and 9695 households in 2013, for which we have complete information for all the three ranking variables. In particular, in these subsamples, households composed only of children, of students aged less than 25 and/or people aged 65 or more are excluded, due to their missing values in the variable work intensity. Although the full samples are constructed to be representative for the entire Spain, this restricted subsamples might not be. Nevertheless, the difference between some statistics obtained with the entire sample and those obtained with the subsamples are tiny. For example, the means household income obtained with the entire sample are 16945 and 16087 in 2009 and 2013 respectively, while both means with the reduced sample are 17842 and 16489.

In the copula-based framework we deal with how the individuals are ranked in the three dimensions, rather than with the exact outcome levels of the individuals in these dimensions. Hence, in each dimension, households are ranked according to the variable defining such dimension. If a tie occurs, the households tied are ranked randomly. As we will illustrate later, this random assignment do not noticeably change the results.

5.2. Results

We first test the hypothesis of increased dependence in Spain over the period 2009-2013 by performing copula based-orderings comparisons, as described in Section 3.

In order to do that, for each year, we compute the empirical copula function over a grid of 27 points in I^3 defined by (0.25, 0.5, 0.75) and then, for each point in the grid, we compare the values of the empirical copula for both years as described in section 3. Then, we approximate the sampling distribution of each empirical copula by resampling with replacement repeatedly (1000 subsamples) from the original sample. Afterwards, on the basis of the bootstrap distributions obtained, we test for the difference between the two years considered to be significant.

The results are displayed in Table 1 (see the Appendix). Several conclusions emerge from this table. First, the empirical copula in year 2013 outranks the empirical

copula in year 2009 at all the grid points. Second, the differences between 2013 and 2009 are significantly non-negative at 5% in all cases. Hence, the PLOD condition holds at all grid points at 5% significant level. This means that the three poverty dimensions (income, no material needs and work intensity) are more likely to have simultaneously small values in 2013 than in 2009; that is, being simultaneously poor in all the three dimensions is more likely to occur in 2013 than it was in 2009. Therefore, the situation for the poors has become even worse in 2013 than it was in 2009.

We next illustrate the multivariate dependence among income, no material needs and work intensity through the multivariate versions of Spearman's rho introduced in Section 4. In order to do that, we compute the estimators of the ten 3-dimensional indexes introduced in Section 4 for the two years considered. Again, we approximate the sampling distribution of these estimators by bootstrap methods, using the resampling scheme explained above. The results are displayed in Table 2 in the Appendix. This table also includes the three possible pairwise Spearman's rho correlation coefficients, namely, the correlation between: income and material needs ($\hat{\rho}_{12}$), income and work intensity ($\hat{\rho}_{13}$) and material needs and work intensity ($\hat{\rho}_{23}$).

As expected, most coefficients have increased its magnitude in 2013 as compared to 2009. Actually, the differences between 2013 and 2009 are all significantly non-negative at 5% level, except for the directional coefficients $\hat{\rho}_3^{(-1,1,1)}$, $\hat{\rho}_3^{(1,1,-1)}$, $\hat{\rho}_3^{(1,-1,-1)}$. Hence, the dependence among dimensions of poverty has increased significantly over the period 2009-2013.

Moreover, in both years, the largest coefficient is $\hat{\rho}_3^-$, indicating that the direction of maximal dependence is (-1,-1,-1). This means that small values of the three poverty dimensions (income, no material needs and work intensity) tend to occur together, and this simultaneous occurrence is more likely in 2013 than in 2009. Noticeably, the second largest coefficient of multivariate dependence is $\hat{\rho}_3^+$, which measures dependence in direction (1,1,1). This means that high values of the three poverty dimensions tend to occur together, i.e., households with high income are more likely to have simultaneously high work intensity and few material needs. Moreover, this simultaneous occurrence of good rankings in the three dimensions is more likely in 2013 than in 2009.

To summarize, our results show that not only the dependence among dimensions of poverty has increased significantly from 2009 to 2013, but also the society has become more polarized, with poverty dimensions (income, no material needs and work intensity) being more likely to have simultaneously small (large) values in 2013 than in 2009.

6. Conclusions

In this paper we analyse how the dependence among the dimensions of poverty has evolved in Spain over the period 2009-2013. The dimensions considered are those included in the AROPE rate: income, material needs and work intensity. In order to do that, we have applied copula-based orderings of dependence over a grid of 27 points to test whether such dependence has increased over the period considered. We have also computed several multivariate extensions of the Spearman's rank coefficient for the two years considered and we have tested, using bootstrap methods, whether the differences between both years are significant.

Our main conclusion is that the degree of dependence among income, material needs and work intensity has noticeably increased over the period considered. Moreover, in both years, the maximal dependence among the three dimensions occur in direction (-1,-1,-1) and the direction of the second largest dependence is (1,1,1). This means that small (high) values of the three poverty dimensions tend to occur together, and this simultaneous concentration of small (large) values of income, no material privations and work intensity, is more likely to occur in 2013 than in 2009. Therefore, after the crisis, the Spanish society has become more polarized, with the poor being more likely to keep being poor and the rich being more likely to keep being rich.

Our results also highlight the relevance of income as the main factor leading poverty, since the other variables considered are highly correlated with it.

References

[1] S. Alkire and J. Foster, Counting and multidimensional poverty measurement, *Journal of Public Economics*, 95(7): 476-487, 2011.

[2] P. Anand, G. Hunter, I. Carter, K. Dowding, F. Guala, M. Van Hees, The development of capability indicators, *Journal of Human Development and Capabilities*, 10 (1): 125–152, 2009.

[3] A.B. Atkinson, Multidimensional deprivation: contrasting social welfare and counting approaches, *Journal of Economic Inequality,* 1:51–65, 2003.

[4] A.B. Atkinson and F.Bourguignon, The comparison of multi-dimensioned distributions of economic status, *Review of Economic Studies*, 12:183-201, 1982.

[5] A.B. Atkinson and S. Morelli, Economic crisis and Inequality, Human Development Research Paper, United Nations Development Programme, 2011.

[6] F. Bourguignon, and S.R. Chakravarty, The measurement of multidimensional poverty, *Journal of Economic Inequality*, 1(1):25–49, 2003.

[7] E.E. Bø, P.J. Lambert and T.O. Thoresen, Horizontal equity under a dual income tax system: principles and measurement, *International Tax and Public Finance,* 19: 625-640, 2012.

[8] K. Decancq, Copula-based measurement of dependence between dimensions of well-being, *Oxford Economic Papers*, 66(3):681-701, 2013.

[9] A. Dolati, M. Úbeda-Flores, On measures of multivariate concordance *Journal of Probability and Statistical Sciences*, 4:147–163, 2006.

[10] J.Y. Duclos, D. Sahn and S. Younger, "Robust multidimensional poverty comparisons", *Economic Journal,* 116(514): 943-968, 2006.

[11] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap,* Chapman and Hall, London, 2003.

[12] J.D. Fermanian, D. Radulovic, M. Wegkamp, Weak convergence of empirical copula processes, *Bernoulli*, 10 (5):847–860, 2004.

[13] FOESSA, Análisis y perspectivas 2013, Desigualdad y Derechos Sociales, FOESSA, Madrid, 2013.

[14] J. E. García, V.A. González-López and R.B. Nelsen. A new index to measure positive dependence in trivariate distributions, *Journal of Multivariate Analysis*, 115: 481–495, 2013

[15] H. Joe, Multivariate concordance. *Journal of Multivariate Analysis*, 35:12–30, 1990.

[16] H. Joe, *Multivariate models and dependence concepts*, Chapman & Hall, London, 1997.

[17] H. Joe, *Dependence Modeling with copulas*, Chapman & Hall, London, 2015.

[18] S. Kolm, Multidimensional equalitarianisms, *Quaterly Journal Economic*, 91(1):1-13, 1977.

[19] E. Maasoumi, The measurement and decomposition of multi-dimensional inequality, *Econometrica*, 54:771-779, 1986.

[20] D. Mari and S. Kotz, *Correlation and Dependence*. Imperial College Press, London, S. 2001.

[21] R.B. Nelsen, Nonparametric measures of multivariate association. In L. Rüschendorf, B. Schweizer, M. D. Taylor (Eds.), *Distributions with given marginals and related topics. IMS Lecture Notes-Monograph Series* (Vol. 28, pp. 223–232). Hayward, CA: Institute of Mathematical Statistics, 1996.

[22] R. B. Nelsen, Concordance and copulas: A survey. In C. M. Cuadras, J. Fortiana, & J. A. Rodriguez-Lallena (Eds.), Distributions with Given Marginals and Statistical Modelling (pp. 169-177), Kluwer Academic Publishers, Dordrecht, 2002.

[23] R. B. Nelsen, *An introduction to copulas* (2nd edn.), Springer, New York, 2006.

[24] R.B. Nelsen and M. Úbeda-Flores, Directional dependence in multivariate distributions, *Annals of the Institute of Statistical Mathematics,* 64:677-685, 2012.

[25] A. Pérez and M. Prieto-Alaiz, A note on nonparametric estimation of two copula-based multivariate extensions of Spearman's rho, submitted

[26] M. Prieto-Alaiz, Y. González-González and C. García-Pérez, La pobreza en España desde una perspectiva multidimensional, Revista de Economía Aplicada, forthcoming, 2015.

[27] C. Quinn, Using copulas to estimate reduced-form systems of equations, Health, Econometrics and Data Group, HEDG Working Papers 07/25, University of York, 2007a.

[28] C. Quinn, The health-economic applications of copulas: methods in applied econometric research, Health, Econometrics and Data Group (HEDG) Working Paper 07/22, University of York, 2007b.

[29] X. Ramos, J. Silber, On the application of efficiency analysis to the study of the dimensions of human development, *Review of Income and Wealth*, 51(2):285–309, 2005.

[30] F. Schmid and R. Schmidt, Multivariate extensions of Spearman's rho and related statistics, *Statistics and Probability Letters*, 77:407–416, 2007.

[31] A.K. Sen, *Commodities and capabilities*. North Holland, Amsterdam, 1985.

[32] M.D. Taylor, Multivariate measures of concordance, *Annals of the Institute of Statistical Mathematics*, 59:789–806, 2007.

[33] K.Y. Tsui, Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson-Kolm-Sen approach, *Journal of Economic Theory*, 67:251-265, 1995.

[34] E. F. Wolff, N-dimensional measures of dependence, *Stochastica,* *4*(3):175-188, 1980.

Appendix: Tables

Table 1. Estimated copula (standard errors in parenthesis) for the three AROPE dimensions evaluated at 27 grid points $u=(u_1,u_2,u_3)$ for the two years 2009 and 2013, together with t-statistics to test for significant differences between the two years considered

$u=(u_1,u_2,u_3)$	2009	2013	t-statistic
(0.25, 0.25, 0.25)	0.064 (0.002)	0.078 (0.002)	5.133
(0.50, 0.25, 0.25)	0.088 (0.002)	0.101 (0.002)	4.570
(0.75, 0.25, 0.25)	0.096 (0.002)	0.107 (0.002)	3.953
(0.25, 0.50, 0.25)	0.095 (0.002)	0.112 (0.002)	5.715
(0.50, 0.50, 0.25)	0.141 (0.002)	0.156 (0.002)	5.102
(0.75, 0.50, 0.25)	0.160 (0.002)	0.172 (0.002)	3.832
(0.25, 0.75, 0.25)	0.107 (0.002)	0.121 (0.002)	5.196
(0.50, 0.75, 0.25)	0.163 (0.002)	0.177 (0.002)	4.706
(0.75, 0.75, 0.25)	0.195 (0.002)	0.202 (0.002)	2.889
(0.25, 0.25, 0.50)	0.095 (0.002)	0.106 (0.002)	4.104
(0.50, 0.25, 0.50)	0.141 (0.002)	0.154 (0.002)	4.065
(0.75, 0.25, 0.50)	0.160 (0.002)	0.168 (0.002)	2.544
(0.25, 0.50, 0.50)	0.148 (0.002)	0.164 (0.002)	5.252
(0.50, 0.50, 0.50)	0.240 (0.003)	0.260 (0.003)	5.187
(0.75, 0.50, 0.50)	0.285 (0.003)	0.299 (0.003)	3.809
(0.25, 0.75, 0.50)	0.170 (0.002)	0.186 (0.002)	5.496
(0.50, 0.75, 0.50)	0.286 (0.003)	0.308 (0.003)	5.973
(0.75, 0.75, 0.50)	0.360 (0.002)	0.373 (0.003)	3.749
(0.25, 0.25, 0.75)	0.109 (0.002)	0.117 (0.002)	2.903
(0.50, 0.25, 0.75)	0.175 (0.002)	0.185 (0.002)	3.331
(0.75, 0.25, 0.75)	0.207 (0.002)	0.211 (0.002)	1.771
(0.25, 0.50, 0.75)	0.171 (0.002)	0.184 (0.002)	4.425
(0.50, 0.50, 0.75)	0.302 (0.003)	0.320 (0.003)	4.717
(0.75, 0.50, 0.75)	0.380 (0.003)	0.394 (0.002)	3.702
(0.25, 0.75, 0.75)	0.199 (0.002)	0.211 (0.002)	4.647
(0.50, 0.75, 0.75)	0.365 (0.002)	0.388 (0.003)	6.439
(0.75, 0.75, 0.75)	0.493 (0.003)	0.512 (0.003)	5.065

Table 2. Estimated parameters (standard errors in parenthesis) of the trivariate correlation coefficients between the three AROPE dimensions for the two years 2009 and 2013, together with t-statistics to test for significant differences between the two years considered

Estimator	2009	2013	t-statistic
$\hat{\rho}_{12}$	0.431 (0.008)	0.503 (0.008)	6.487
$\hat{\rho}_{13}$	0.424 (0.009)	0.455 (0.009)	2.539
$\hat{\rho}_{23}$	0.239 (0.009)	0.340 (0.009)	7.788
$\hat{\rho}_3^+$	0.357 (0.006)	0.417 (0.007)	6.571
$\hat{\rho}_3^-$	0.372 (0.007)	0.449 (0.007)	8.121
$\hat{\rho}_3$	0.365 (0.006)	0.433 (0.006)	7.607
$\hat{\rho}_3^{(-1,1,1)}$	-0.197 (0.004)	-0.190 (0.004)	1.218
$\hat{\rho}_3^{(1,-1,1)}$	-0.074 (0.005)	-0.114 (0.004)	-6.055
$\hat{\rho}_3^{(1,1,-1)}$	-0.070 (0.005)	-0.081 (0.005)	-1.589
$\hat{\rho}_3^{(-1,-1,1)}$	-0.085 (0.005)	-0.113 (0.005)	-3.880
$\hat{\rho}_3^{(-1,1,-1)}$	-0.090 (0.005)	-0.145 (0.005)	-7.684
$\hat{\rho}_3^{(1,-1,-1)}$	-0.213 (0.004)	-0.222 (0.004)	-1.637
$\hat{\rho}_3^{max}$	0.372 (0.007)	0.449 (0.007)	8.121