Hacienda Pública Española/Review of Public Economics

Manuscript 2281

Changes in the dependence structure of AROPE components: evidence from the Spanish regions

César García-Gómez

Ana Pérez

Mercedes Prieto-Alaiz

Changes in the dependence structure of AROPE components: evidence from the Spanish regions

César García-Gómez[a]*, Ana Pérez[a,b] and Mercedes Prieto-Alaiz[a]

[a]Dpto. Economía Aplicada, Universidad de Valladolid.

[b]IMUVA, Universidad de Valladolid.

Abstract

The AROPE rate is a multidimensional indicator to monitor poverty in the European Union which combines income, work intensity and material deprivation. However, it misses the possible relationship between its components. To overcome this drawback, some authors proposed to complement the AROPE rate with measures of the dependence between its dimensions, since higher dependence can exacerbate poverty. In this paper, we follow this approach and measure that dependence in the Spanish regions over the period 2008-2018 using three multivariate versions of Spearman's rank correlation coefficient. Our results reveal an asymmetric effect of the economic cycle on the dependence between poverty dimensions, as this dependence, in many Spanish regions, substantially increased during the Great Recession but dropped little during the economic recovery. Moreover, regions with higher AROPE rates also tend to experience more dependence between its dimensions.

*Corresponding Author. Dpto. Economía Aplicada. Universidad de Valladolid. C/Avda. Valle Esgueva 6, 47011 Valladolid. Spain. E-mail: cesar.garciag@uva.es

JEL Classification: D63, I32, O52.

Keywords: Multivariate dependence; Spearman's rho; Multidimensional poverty; AROPE.

1 Introduction

The Sustainable Development Goal 1 in the 2030 Agenda for Sustainable Development adopted by the United Nations is "End poverty in all its forms everywhere". In the European Union (EU), the indicator adopted to monitor progress towards this goal is the AROPE (At Risk Of Poverty or Social Exclusion) rate. This rate captures somehow the multidimensional nature of poverty as it combines three facets of poverty: income, work intensity and material deprivation. However, it is blind to an inherent feature of any multidimensional phenomena, which is the interrelationship between its dimensions. To understand the importance of this feature when analysing multidimensional poverty, let us consider the following example. Imagine a region where one household is top-ranked in all dimensions, another household is second-ranked in all dimensions, and so forth, until the last household, which scores the lowest in all dimensions. In this region, there is arguably more concentration of deprivations than in another region with the same marginal distribution for each dimension but where some households score high in some dimensions and low in others. Hence, a higher degree of dependence means more clustering of disadvantages and thus a higher risk of cumulative deprivation; see Atkinson and Bourguignon (1982), Bourguignon and Chakravarty (2003), Chakravarty (2018), Decancq (2020, 2022), and the references therein. In this framework, it could be relevant to analyse the dependence between poverty dimensions when designing and evaluating policies whose objective is reducing the extent of cumulative deprivation.

There are different approaches to incorporate the dependence into the poverty analysis.

On one hand, one could think of designing poverty assessment tools which are to some extent sensitive to dependence; see, for instance, Alkire and Foster (2011), Ferreira and Lugo (2013), Seth (2013), Duclos and Tiberti (2016) and Seth and Santos (2019). On the other hand, one could quantify the degree of dependence between the dimensions of poverty, making easier the comparisons across different societies or over time (see Decancq (2014), Pérez and Prieto-Alaiz (2016a), García-Gómez et al. (2021), D 'Agostino et al. (2022), among others). This paper follows the second approach.

The aim of this paper is to quantify the multivariate dependence between the poverty dimensions in the Spanish regions and evaluate the effect of the Great Recession on the dependence structure of poverty. Now, the question arises on how to capture the dependence among the single components of a multidimensional concept, knowing that measuring pairwise dependence, over all distinct bivariate margins, is not enough and could conceal important aspects of multivariate dependence; see Durante et al. (2014). To address this goal, we complement the information given by the AROPE rate with measures of multivariate dependence between its three components and evaluate their evolution over the period 2008-2018. In particular, we use three multivariate extensions of the bivariate Spearman's rank correlation coefficient, proposed by Joe (1990) and Nelsen (1996, 2002), which are appropriate to measure multivariate dependence in non-Gaussian and possibly non-linear contexts, such as the ones we usually face in multidimensional poverty analyses. First, we consider the coefficient of average lower orthant dependence, which is the most relevant one in our context, as it captures, through a rescaled average, the probability of being simultaneously low-ranked in all poverty dimensions as compared to what this would be were the dimensions independent. In a similar fashion, we compute the coefficient of average upper orthant dependence, which measures departure from independence based on the propensity of cumulative affluence, and finally, we calculate the average of both of these coefficients.

Our main findings are as follows. First, low (high) values of income tend to occur simultaneously with low (high) values of the other two dimensions in all Spanish regions. However, the average probability of being simultaneously low-ranked in all dimensions tends to be higher than the mirrored probability of being simultaneously high-ranked. Second, there is evidence of an asymmetric effect of the economic cycle on the dependence between poverty dimensions, as this dependence, in many Spanish regions, substantially increased during the Great Recession but dropped little during the economic recovery. Third, the regions with more deprived people, as measured by the AROPE rate, tend to also experience a higher risk of clustering of disadvantages.

This paper builds on recent related work that look at the dependence between poverty dimensions. García-Gómez et al. (2021) analyse the evolution of the multivariate dependence between the three AROPE components over the period 2008-2014 for the EU-28 countries, using the same coefficients as ours, and find that, in most countries, the Great Recession entailed a significant increase in dependence between poverty dimensions. In comparison, our analysis is confined to Spain and compares regions while covering a longer period (2008-2018) that includes both the recession and recovery. Decancq (2020) relies on the notion of cumulative deprivation and derives measures of dependence from the so-called cumulative deprivation curve. In a follow-up paper, Decancq (2022) provides an illustration of these concepts with information about income, health and housing from the Belgian MEQIN data set. Our approach differs from these two papers in both methodological and empirical aspects. Whereas they use Spearman's footrule-type coefficients and confine its application to national Belgium data, we apply Spearman's rho-type coefficients and extend to sub-national data for the Spanish regions. However, there are some similarities between our contributions, as the type of measures we both use share some interesting properties. D'Agostino et al. (2022) and García-Gómez et al. (2022) employ tail dependence measures to study the evolution of poverty in the EU-28 countries. Like us, both papers rely on the dimensions of AROPE rate, but the former is limited to the bivariate case and two years (2009 and 2018) whereas the latter takes a multidimensional approach and covers a larger period (2008-2018), as we do. However, our paper differs from García-Gómez et al. (2022) in both the measures used and the scope of the empirical application. Noticeably, our conclusions are similar, as we both find that dependence between poverty dimensions is time-varying over 2008-2018 and the effect of the Great Recession is not homogeneous over the units (countries/regions) analysed. Another reference dealing with dependence between poverty dimensions in several European countries is Tkach and Gigliarano (2018). However, this paper uses a parametric approach and focuses only on bivariate dependence. By contrast, we use a non-parametric method and analyse multivariate dependence.

To the best of our knowledge, our paper is the first attempt to measure the multivariate dependence between poverty dimensions at a sub-national level. This regional analysis allows us to discover very different patterns that remain concealed when analysing aggregate data for Spain as a whole. Moreover, the long period analysed (2008-2018) and the methodology used lead to a better understanding of the effect of economic cycles on the regional poverty structure. This could be important for policy-makers, since the Spanish regions have an important degree of autonomy and their own public budgets to design and apply social policies.

The rest of the paper is structured as follows. In Section 2, we introduce the three multivariate extensions of Spearman's rho. Section 3 is devoted to the application of these measures to analyse the evolution of multivariate dependence between the AROPE components in the Spanish regions over the period 2008-2018. Finally, Section 4 summarises the main conclusions of the paper.

2 Multivariate generalisations of Spearman's rho

As we pointed out above, one key aspect of multidimensional poverty overlooked by the AROPE rate is the possible dependence between its three components. Moreover, as these components follow non-Gaussian distributions and we need to measure dependence in a multivariate framework, other coefficients beyond Pearson's linear correlation should be applied. To face this goal, we focus on three multivariate extensions of the well-known Spearman's rank coefficient, which are based on the positions of the households across variables, rather than on the specific values that the corresponding variables attain for such households. In this section, we first describe the population version of bivariate Spearman's rho and then we introduce the multivariate generalisations and we briefly discuss how to estimate these measures to be useful in practical applications.

Let X_1 and X_2 denote two continuous random variables with joint cumulative distribution function F and marginal distribution functions F_1 and F_2, respectively, and let $F_1(X_1)$ and $F_2(X_2)$ be the random variables defined by the probability integral transformations. These variables assign to each household in the population its relative position (rank) in the ith dimension, with $i = 1, 2$, and follow standard uniform distributions $U(0, 1)$. For instance, if X_1 and X_2 are two dimensions of poverty, say income and education, the variables $F_1(X_1)$ and $F_2(X_2)$ will transform the outcomes of each household in income and education into the positions that this household attains in both poverty dimensions as compared with others. Then, Spearman's rho is based on comparing such positions, so that the more aligned these positions are, the stronger the relationship between the variables X_1 and X_2.

Formally, the population version of bivariate Spearman's rho for X_1 and X_2 can be defined as the Pearson's correlation coefficient between the position variables $F_1(X_1)$ and $F_2(X_2)$, that becomes:

$$\rho_S = 12E(F_1(X_1)F_2(X_2)) - 3. \tag{1}$$

This coefficient can be alternatively written (Schweizer and Wolff, 1981) as

$$\rho_S = 12\int_{R^2}[F(x_1,x_2) - F_1(x_1)F_2(x_2)]dF_1(x_1)dF_2(x_2). \tag{2}$$

Hence, as Nelsen (2002) points out, ρ_S can be regarded as a measure of average quadrant dependence as it measures the "average distance" between the joint distribution of X_1 and X_2 (as represented by F) and independence.[1] Therefore, when X_1 and X_2 are independent, $\rho_S = 0$. Moreover, the maximum value of ρ_S is $+1$ for perfect positive dependence, that is, when one variable is almost surely a strictly increasing function of the other (the ranks in both variables coincide) and its minimum value is -1 for perfect negative dependence, that is, when one variable is almost surely a strictly decreasing function of the other (the ranks in both variables are reversed). Note also that, since the survival function of any variable X_i is defined as $\overline{F}_i(x_i) = p(X_i > x_i) = 1 - F_i(x_i)$, ρ_S can also be written as

$$\rho_S = 12E(\overline{F}_1(X_1)\overline{F}_2(X_2)) - 3. \tag{3}$$

When we move to a multivariate framework with more than two variables involved, there is not a unique multivariate version of ρ_S. In particular, we focus on two multivariate versions of ρ_S introduced by Joe (1990) that are further developed in Nelsen (1996) in a copula-based framework. To formally define these coefficients, let $\mathbf{X} = (X_1, ..., X_d)$ be a d-dimensional continuous random variable with joint distribution function F and margins $F_1, .., F_d$ such that the transformed variables $F_i(X_i)$, for $i = 1, 2, ..., d$, are uniform on $[0, 1]$. In our setting, the random vector $\mathbf{X}$ represents the relevant d dimensions of poverty for a population and the random vector $(F_1(X_1), ..., F_d(X_d))$ represents the relative positions

[1] See Appendix A for the definition of quadrant dependence.

of one household in all poverty dimension as compared to other households.

The first generalisation of ρ_S that we consider was proposed by Joe (1990) as a scaled expected value of $F_1(X_1)\cdots F_d(X_d)$, namely:

$$\rho_d^+ = \frac{(d+1)}{2^d-(d+1)}\left[2^d E\left(\prod_{i=1}^{d} F_i(X_i)\right)-1\right]. \tag{4}$$

Alternatively, Joe (1990) proposed a second multivariate version of ρ_S which consists of replacing $F_1\cdots F_d$ in (4) by the corresponding survival functions $\overline{F}_1\cdots\overline{F}_d$ and is given by:

$$\rho_d^- = \frac{(d+1)}{2^d-(d+1)}\left[2^d E\left(\prod_{i=1}^{d} \overline{F}_i(X_i)\right)-1\right]. \tag{5}$$

Notice that equations (4) and (5) are the natural generalisations of (1) and (3), respectively. Additionally, Nelsen (2002) introduced a third multivariate Spearman's rho as the average of the two generalisations in (4)-(5), namely

$$\rho_d = \frac{\rho_d^- + \rho_d^+}{2}. \tag{6}$$

For radially symmetric distributions, the three coefficients above coincide. Moreover, when the variables $(X_1,...,X_d)$ are independent, they all take the value 0 and they take the maximum value +1 for perfect positive dependence, that is, when each of the random variables $X_1,...,X_d$ is almost surely a strictly increasing function of any of the others, and they all have a lower bound $[2^d-(d+1)!]/\{d![2^d-(d+1)]\}$, which is greater than -1; see Nelsen (1996). For $d=2$, the three coefficients in (4)-(6) reduce to bivariate Spearman's ρ_S in (1), and for $d=3$, the coefficient ρ_3 becomes the average of the three possible pairwise Spearman's rho, i.e.:

$$\rho_3 = \frac{\rho_{12}+\rho_{13}+\rho_{23}}{3},$$

where ρ_{ik} denotes the bivariate sample Spearman's rho for the pair (X_i, X_k), with $1 \leq i < k \leq 3$.

The advantage of ρ_d^+ and ρ_d^- over ρ_d is that the former could reveal some forms of multivariate dependence that ρ_d fails to detect. For instance, $\rho_d = 0$ could be erroneously interpreted as no dependence at all, whereas ρ_d^+ and ρ_d^- can be different from zero, indicating some degree of lower and upper dependence; see examples 1 and 6 in Nelsen (1996). But ρ_d^+ and ρ_d^- are not without their shortcomings either. For instance, Nelsen and Úbeda-Flores (2012) show that ρ_3^+ and ρ_3^- may fail to detect some forms of trivariate dependence when they take values around zero. However, García et al. (2013) show that, if the three pairwise ρ_{12}, ρ_{13} and ρ_{23} are all positive, there is no undetected positive trivariate dependence in ρ_3^+ and ρ_3^-.

As we mentioned before, ρ_d^+ and ρ_d^- were discussed in Nelsen (2002) as copula-based measures of average upper and lower orthant dependence, respectively.[2] Following this approach, we could say that ρ_d^+ captures the "similarity" between the multivariate distribution of $\mathbf{X}$ and independence from an upper perspective, by comparing, through a rescaled average, the probability that all variables take simultaneously high values with the value of this probability if the variables were independent. Hence, the more aligned the positions are around the upper corner of the joint distribution, the higher the value of ρ_d^+. By contrast, ρ_d^- captures the "similarity" between our multivariate data and the situation of independence from a lower perspective by comparing, through a rescaled average, whether the probability that all variables are simultaneously low is at least as large as in the case of independence. Hence, the more aligned the positions are around the lower corner of the joint distribution, the higher the value of ρ_d^-. In turn, the coefficient ρ_d can be regarded as a measure of average orthant dependence and it also fulfills the conditions to be a measure of multivariate concordance; see Dolati and Úbeda-Flores (2006).

[2]See Appendix A for a brief description of orthant dependence concepts.

In welfare economics, where the data usually do not exhibit symmetric distributions - see Kleiber and Kotz (2003) and the references therein - the coefficients ρ_d^+ and ρ_d^- are preferable to ρ_d, in general. For example, if the joint distribution of income, health and education, becomes more concentrated around its lower tail (more clustering of disadvantages) than around its upper tail (more concentration of advantages), such difference will be captured by ρ_d^- and ρ_d^+, respectively, but ρ_d will be blind to them. Even though, the coefficient ρ_d could still provide valuable information for welfare analysis. For instance, Decancq (2014) recalls that ρ_d can be interpreted as the normalized probability that a randomly drawn household from a given society outranks or is outranked by a randomly drawn household from a reference society with independent welfare dimensions. Hence, the higher this probability, the higher the dependence between the dimensions of welfare. In a poverty setting, the more relevant coefficient is ρ_d^- because it allows to capture how likely it is, in average, to be simultaneously "low ranked" in all poverty dimensions as compared to independence. Moreover, in a trivariate setting, the pairwise Spearman's could still be interesting, as they provide useful clues on whether ρ_3^+ and ρ_3^- are fully informative. But simple methods, like only averaging pairwise coefficients, will not be enough for a proper understanding of the dependence structure of multidimensional poverty.

The dependence measures described so far are only appropriate for measuring multivariate dependence between continuous variables. In order to account for possible discontinuities in the components of $\mathbf{X}$, Quessy (2009), Mesfioui and Quessy (2010) and Genest et al. (2013) proposed modified versions of the multivariate Spearman's rho which amounts to substituting the function F_i in (4)-(6), by other function $\widetilde{F}_i$, which is defined, for all $i \in \{1, ..., d\}$ and $x \in \mathbb{R}$, as

$$\widetilde{F}_i(x) = \frac{1}{2}\{F_i(x^-) + F_i(x)\},$$

where $F_i(x^-) = \Pr(X_i < x)$ and $F_i(x) = \Pr(X_i \leq x)$. Hence, the new coefficients proposed become[3]

$$\rho_d^{+\maltese} = \frac{(d+1)}{2^d-(d+1)}\left[2^d E\left(\prod_{i=1}^{d}\widetilde{F}_i(X_i)\right)-1\right], \tag{7}$$

$$\rho_d^{-\maltese} = \frac{(d+1)}{2^d-(d+1)}\left[2^d E\left(\prod_{i=1}^{d}(1-\widetilde{F}_i(X_i))\right)-1\right] \tag{8}$$

$$\rho_d^{\maltese} = \frac{\rho_d^{-\maltese}+\rho_d^{+\maltese}}{2}. \tag{9}$$

If all the components of $\mathbf{X}$ are continuous, then $\widetilde{F}_i = F_i$ for all i and the coefficients in (7)-(9) will reduce to those in (4)-(6). Moreover, the former inherit some of the properties of the latter. In particular, they all become 0 in the case of multivariate independence and attain their maximum value in the case of positive dependence, but such value is smaller than 1, when the probability of ties is positive for one or more of the variables; see Quessy (2009). Noticeably, when $d = 2$, the three coefficients above coincide with one version of bivariate Spearman's rho for non-continuous data proposed by Nešlehová (2007). Furthermore, when $d = 3$, $\rho_3^{\maltese}$ becomes the average of the three corresponding non-continuos pairwise Spearman's rho.

In practice, the coefficients previously described must be estimated from the data. Therefore, empirical versions of these coefficients are required. For continuous variables, Pérez and Prieto-Alaiz (2016b) derived nonparametric estimators of ρ_d^- and ρ_d^+ which are consistent and asymptotically normally distributed. For non-continuous variables, such as the ones usually encountered in welfare economics, estimates of $\rho_d^{-\maltese}$ and $\rho_d^{+\maltese}$ will be required. In order to do that, let $\mathbf{X}_1 = (X_{11}, ..., X_{d1}), ..., \mathbf{X}_n = (X_{1n}, ..., X_{dn})$ be a sample of n serially independent random vectors from the d-dimensional vector $\mathbf{X} = (X_1, ..., X_d)$, with joint distribution distribution function F and margins $F_1, ..., F_d$. For each $i = 1, .., d$,

[3] A detailed description of the motivation and concepts behind these coefficients is included in Appendix B.

the empirical counterpart of F_i at any $x \in R$, is defined by

$$F_{in}(x) = \frac{1}{n}\sum_{j=1}^{n} 1\{X_{ij} \leq x\},$$

where $\mathbf{1}\{A\}$ denotes the indicator function of a set A. For possibly non-continuous data, the empirical analogue $\widetilde{F}_{in}$ of $\widetilde{F}_i$ then satisfies, for $i = 1, ..., d$ and $j = 1, ..., n$,

$$\widetilde{F}_{in}(X_{ij}) = \frac{1}{2}\{F_{in}(X_{ij}^{-}) + F_{in}(X_{ij})\}. \tag{10}$$

Now, to estimate the coefficients in (7)-(9), Genest et al. (2013) proposed plug-in estimators obtained by replacing $\widetilde{F}_i(X_i)$ with its empirical analogue $\widetilde{F}_{in}(X_{ij})$ in (10). In doing so, the following expressions come up

$$\hat{\rho}_d^{+\maltese} = \frac{(d+1)}{2^d-(d+1)}\left[2^d\frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d}\widetilde{F}_{in}(X_{ij}) - 1\right], \tag{11}$$

$$\hat{\rho}_d^{-\maltese} = \frac{(d+1)}{2^d-(d+1)}\left[2^d\frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d}\Big(1-\widetilde{F}_{in}(X_{ij})\Big) - 1\right], \tag{12}$$

$$\hat{\rho}_d^{\maltese} = \frac{\hat{\rho}_d^{-\maltese} + \hat{\rho}_d^{+\maltese}}{2}. \tag{13}$$

Genest et al. (2013) worked out alternative expressions of the estimators above based on midranks. In particular, when $d = 2$, these three estimators coincide and are equal to:

$$\hat{\rho}_2^{\maltese} = \frac{12}{n^3}\sum_{j=1}^{n}(\widetilde{R}_{1j} - \widetilde{R}_{1\bullet})(\widetilde{R}_{2j} - \widetilde{R}_{2\bullet}), \tag{14}$$

where $\widetilde{R}_{ij}$ is the midrank of X_{ij} among $\{X_{i1}, ..., X_{in}\}$ and $\widetilde{R}_{i\bullet}$ is the average of the midranks for component i, with $i = 1, 2$ and $j = 1, ..., n$. Furthermore, when $d = 3$, the following relationship holds:

$$\hat{\rho}_3^{\maltese} = \frac{\hat{\rho}_{12}^{\maltese} + \hat{\rho}_{13}^{\maltese} + \hat{\rho}_{23}^{\maltese}}{3},$$

where $\hat{\rho}_{ij}^{\maltese}$ denotes the corresponding bivariate coefficient in (14) for the pair (X_i, X_k), with $1 \leq i < k \leq 3$. Genest et al. (2013) show that the estimators in (11)-(13) are asymptotically normally distributed and provide expressions of their limiting variances, thereby correcting previous errors in the formulas derived in Quessy (2009) and Mesfioui and Quessy (2010). Nevertheless, the asymptotic variances are complex and hence, in practice, they will be estimated by bootstrap methods, as we will see in the following section.

3 Multivariate dependence between AROPE components in Spanish regions (2008-2018)

As we pointed out in the Introduction, the AROPE rate is the key indicator adopted by the EU to monitor poverty and living conditions. Nevertheless, as a multidimensional poverty measure, it misses an important part of the picture, which is the possible interactions between its dimensions. To overcome this drawback, in this section, we complement the information provided by the AROPE rate with measures of multivariate dependence between its components. In particular, we apply the three multivariate extensions of Spearman's rho discussed in Section 2 to analyse how the multivariate dependence between the three variables defining the AROPE rate (income, work intensity and material deprivation) has evolved in the Spanish regions over the period 2008-2018.

3.1 Data and variables

The data used comes from the cross-sectional waves of the EU-SILC survey of all years of the period 2008-2018. This survey constitutes the reference source of information on income and living conditions in the EU countries. Moreover, it provides territorial disaggregation by NUTS-2 or regions, which is essential to study cross-regional differences in countries like Spain. EU-SILC provides the three variables characterising the AROPE rate, which are defined as follows.

- The measure of income is the equivalised disposable income of the household using the OECD modified equivalence scale.[4]
- The work intensity of a household is the ratio of the total number of months that all working-age household members have worked during the income reference year and the total number of months they could have theoretically worked during the same period.[5]
- Number of deprivations from a list of the following nine items: 1) the capacity of facing unexpected expenses; 2) one-week annual holiday away from home; 3) a meal involving meat, chicken or fish every second day; 4) an adequately warm dwelling; 5) a washing machine; 6) a colour television; 7) a telephone; 8) a car; 9) the capacity to pay their rent, mortgage or utility bills.

In our empirical application, we use the number of "no-deprivations" as a measure of material needs. Thus, this variable takes the following values: 0 (having all the 9 possible deprivations), 1 (having eight out of the nine aforementioned deprivations), . . . , 9 (having no deprivations). In doing so, the three variables considered keep the same relationship with poverty, that is, high values of these three variables (equivalised disposable income, work intensity, and number of no-deprivations) indicate lower chance to be poor, whereas low values of each variable indicate higher chance to be poor.

The unit of analysis is the household and we only work with subsamples of households

[4]The modified OECD scale gives a weight of 1 to the first adult, 0.5 to other household members aged 14 or over and 0.3 to household members aged less than 14.

[5]Eurostat considers that a working-age person is a person aged 18-59 years, excluding also the students aged 18-24 years.

for which we have complete information for all the three variables. In particular, in these subsamples, households composed only of children, of students aged 18-24 and/or people aged 60 or more are excluded, due to their missing values in the work intensity variable. In these subsamples, the sample sizes range from 68 households (Melilla, 2012) to 2063 households (Catalonia, 2017).

3.2 Estimation results

In this Section, we analyse the patterns of multivariate dependence between poverty dimensions in the Spanish regions in the period 2008-2018. To do so, and given the non-continuous nature of some of the variables used, we apply the non-parametric estimators of the tie-corrected multivariate extensions of Spearman's rho defined in (11)-(13). To calculate all point estimates, we compute the empirical cumulative distribution function thereof using weighted data with cross-sectional weights.

Since the asymptotic variances of these estimators are complex, we rely on a non-parametric bootstrap method to compute the bootstrap standard errors (see Mashreghi et al. (2016) for a review of bootstrap methods in the context of survey data). As Goedemé (2013) and Osier et al. (2013) point out, the implementation of bootstrap should take into account as much as possible the EU-SILC's complex sample design, otherwise the method could yield an inconsistent variance estimator.[6] In our case, the bootstrap procedure comprise three steps. First, we draw an independent sample of households with replacement from the original sample in each Autonomous Community. Second, the cross-sectional weights are adjusted as Rao et al. (1992) and Rust and Rao (1996) proposed. For instance, the

[6]For each Autonomous Community, a two-stage random sampling design with stratification of the first stage units is used. The primary sampling units are census sections and the last sample units are the main family dwellings. The final sample includes all private households resident in the main family dwellings selected.

adjusted weight for household i in Autonomous Community j, w_{ij}^*, is given by

$$w_{ij}^* = w_{ij} r_i \frac{n_j}{n_j - 1}$$

where w_{ij} is the original cross-sectional weight, r_i is the number of times the $i-th$ household in Autonomous Community j is selected in the bootstrap sample and n_j is the original sample size of Autonomous Community j. Third, we compute point estimates of multivariate extensions of Spearman's rho, using the adjusted weights. These three steps are repeated 1000 times. At the end, the bootstrap standard error estimator is approximated by the sample standard deviation of 1000 point estimates of each multivariate extensions of Spearman's rho.

Figure 1 displays, for Spain and its regions and over the whole period analysed, the values of $\hat{\rho}_3^{-\maltese}$ (in blue), $\hat{\rho}_3^{+\maltese}$ (in green) and $\hat{\rho}_3^{\maltese}$ (in red). The first conclusion that can be highlighted is that, both for Spain and all its regions, and for all years, the three coefficients are always positive, indicating a positive multivariate association between poverty dimensions both from a downwards and an upwards perspective. This means that households with low (high) incomes tend to suffer simultaneously from low (high) work intensity and many (few) deprivations.

- INSERT FIGURE 1 HERE -

Another relevant feature that emerges from Figure 1 is that $\hat{\rho}_3^{-\maltese}$ is greater than $\hat{\rho}_3^{+\maltese}$ in Spain and most of its regions. This indicates that average lower orthant dependence between poverty dimensions tends to be higher than average upper orthant dependence. That is, the average normalised probability of being simultaneously low-ranked in all poverty dimensions tends to be higher than the average normalised probability of being simultaneously high-ranked in all poverty dimensions. The exceptions to this pattern are Andalusia, the Canary Islands, Ceuta, Murcia and Extremadura, where $\hat{\rho}_3^{-\maltese}$ and $\hat{\rho}_3^{+\maltese}$ are very similar. Although the methodology used does not allow to provide causal explanations for this latter pattern, it is important to remark that these regions traditionally present a combination of high levels of inequality and poverty; see, for instance, Ayala et al. (2011) and Ayala and Jurado (2020). Hence, it seems plausible that inequality affects the dependence structure. Future research should throw more light into the relationship between regional characteristics and the dependence structure of poverty dimensions.

Figure 1 also suggests the existence of important cross-regional differences in the level of multivariate dependence between poverty dimensions. To better visualise these differences, Figure 2 displays, for each Spanish region, the levels of $\hat{\rho}_3^{-\maltese}$ for years 2008, 2014 and 2018.[7] As we can see in these figures, the lowest values for $\hat{\rho}_3^{-\maltese}$ in 2008 were found in northern regions such as Asturias, the Basque Country, Navarre and Aragon. By contrary, multivariate dependence between poverty dimensions seemed to be higher in the south of Spain (Extremadura, Andalusia or the Canary Islands), with relatively high levels also in some northern regions such as Cantabria, Galicia or Catalonia. If we look at the situation in 2014, as compared to 2008, we notice a remarkable increase in average lower orthant dependence between poverty dimensions. This means that, after the Financial Crisis of 2008, Spain experienced an increase in the overall tendency of households to be simultaneously poor in the three dimensions of the AROPE rate, and this increase is found in the vast majority of its regions. By contrast, between 2014 and 2018, there was a decrease in average lower orthant dependence in some of the regions which had the lowest levels of dependence in 2008, namely the Basque Country, Navarre or Aragon, whereas other regions do not seem to have experienced a relevant change in the level of dependence over that period; see, for instance, Madrid, Extremadura or Andalusia. Furthermore, if we compare the situation in 2018 with that of 2008, average lower orthant dependence was still higher in 2018 than in 2008 in many Spanish regions, which suggests the existence

[7] We only show here the results of $\hat{\rho}_3^{-\maltese}$, but the figure is very similar if we consider $\hat{\rho}_3^{+\maltese}$ or $\hat{\rho}_3^{\maltese}$. These figures are available upon request.

of an asymmetric response of dependence between poverty dimensions to the economic cycle. Noticeably, the cross-regional differences observed in 2018 are very similar to those previously described for 2008.

- INSERT FIGURE 2 HERE -

To get a better insight into the evolution of multivariate dependence between poverty dimensions, Tables 1 and 2 report, for the Spanish regions and also for Spain as a whole, point estimates of $\hat{\rho}_3^{-\maltese}$ and $\hat{\rho}_3^{+\maltese}$, respectively, as well as their bootstrap standard errors (in parenthesis), for years 2008, 2014 and 2018. Moreover, these tables also display the results of three one-sided two-independent sample t-tests with unequal variances to determine the significance of the variation of the corresponding coefficient over the sub-periods 2008-2014 (in column 4) and 2014-2018 (in column 5), and over the whole period 2008-2018 (in column 6). The t-statistics are computed using bootstrap standard errors and the corresponding p-value (in parenthesis) is computed assuming asymptotic normality of the t-statistic.

- INSERT TABLE 1 HERE -

- INSERT TABLE 2 HERE -

These tables confirm the patterns depicted in Figures 1 and 2. First, in Spain as a whole, both $\rho_3^{-\maltese}$ and $\rho_3^{+\maltese}$ significantly increased between 2008 and 2014 and significantly decreased between 2014 and 2018, but the level of multivariate dependence between poverty dimensions was still higher in 2018 than in 2008. However, when we analyse the evolution of the coefficients in the Spanish regions, several cross-regional differences arise, depending on the period analysed:

- *Period 2008-2014* (column 4 of Tables 1 and 2). In all the Spanish regions, except in Castilla La Mancha, La Rioja, Extremadura, Andalusia, Ceuta and the Canary Islands, there was an statistically significant increase in both $\rho_3^{-\maltese}$ and $\rho_3^{+\maltese}$ over this period. Moreover, over this period, $\rho_3^{-\maltese}$ also significantly increased in Andalusia and $\rho_3^{+\maltese}$ in Castilla La Mancha and La Rioja. Hence, these results suggest that, after the Financial Crisis of 2008, there was a generalised increase in the likelihood of a household to occupy simultaneously bottom (top) positions in all poverty dimensions.

- *Period 2014-2018* (column 5 of Tables 1 and 2). Different situations can be distinguished. On one hand, we find that, in most of the Spanish regions, both coefficients did not significantly change, i.e, the average probability of being simultaneously low (high)-ranked in the three dimensions of the AROPE rate remained rather stable. This is the case of Asturias, Cantabria, La Rioja, Madrid, Castilla La Mancha, Extremadura, the Balearic Islands, Andalusia, Murcia, Melilla and the Canary Islands. On the other hand, we find regions where the coefficients significantly decreased over that period. This is the case of some of the regions which started with a low level of dependence in 2008, namely the Basque Country, Navarre, Aragon, Castile and Leon, Catalonia and the Valencian Community. Moreover, $\rho_3^{+\maltese}$ (Table 2) also decreased in Galicia. Remarkably, we only find region, namely Ceuta, where multivariate dependence between poverty dimensions significantly increased in the post-crisis period of 2014-2018.

- *Period 2008-2018* (column 6 of Tables 1 and 2). In many Spanish regions, namely Galicia, Asturias, Basque Country, Madrid, Castilla La Mancha, Balearic Islands, Andalusia and Murcia, $\rho_3^{-\maltese}$ was still significantly higher in 2018 than in 2008. Thus, in these regions it was still more likely, on average, to be simultaneously low-ranked in all dimensions of poverty in 2018 than in 2008. This suggest the existence, in these regions, of an asymmetric effect of the economic cycle on the dependence between poverty dimensions, with this dependence increasing substantially during the crisis but decreasing little during the economic recovery. Moreover, in the case of $\rho_3^{+\maltese}$, this coefficient was still significantly higher in 2018 than in 2008 in Asturias, Madrid, Extremadura, Andalusia and Murcia. There are also some regions, namely Cantabria, Navarre, La Rioja, Aragon, Castile and Leon, Catalonia, the Valencian Community, Ceuta, Melilla and the Canary Islands, where $\rho_3^{-\maltese}$ and $\rho_3^{+\maltese}$ in 2018 were not statistically different from those in 2008.

As expected, the analysis of the evolution of $\hat{\rho}_3^{\maltese}$ (Table 3) leads to the same conclusions, since this coefficient is the average of $\hat{\rho}_3^{-\maltese}$ and $\hat{\rho}_3^{+\maltese}$. However, since $\rho_3^{\maltese}$ is also the average of the three possible pairwise Spearman's rho, we complement our analysis by measuring all possible pairwise relationships between the three dimensions of the AROPE rate. These results are depicted in Figure 3, which shows, for Spain and its regions, the estimated values of the pairwise dependence coefficients between income and work intensity (in red), income and no-material deprivation (in green), and work intensity and no-material deprivation (in blue) over the whole period analysed (2008-2018). Several conclusions can be drawn from this figure. First, both for Spain and all its regions, the three pairwise coefficients are always positive over the period analysed; recall that this ensures that there is no trivariate dependence left undetected by $\rho_3^{-\maltese}$ and $\rho_3^{+\maltese}$. Second, both in Spain and all its regions, the highest dependence is always between income and either work intensity or no-material deprivation, whereas the lowest dependence is found between work intensity and no-material deprivation (except for some years in Extremadura and Ceuta). Third, the evolution of these coefficients is very different from one region to another, which makes it difficult to reach general conclusions. Nevertheless, we generally observe, in Spain and most of its regions, an increase in the three coefficients between 2008 and 2014, followed by a stabilisation or decrease afterwards.

- INSERT TABLE 3 HERE -

- INSERT FIGURE 3 HERE -

To close this section, we wonder whether those regions with higher incidence of poverty are also regions with higher risk of cumulative deprivation. To address this issue, Figure 4 contains three scatter plots showing the relationship between the AROPE rate and the coefficient $\hat{\rho}_3^{-\maltese}$ for the Spanish regions in 2008, 2014 and 2018.[8] In all graphs, the horizontal and vertical reference lines represent the average values of $\hat{\rho}_3^{-\maltese}$ and the AROPE rate, respectively. We focus on $\hat{\rho}_3^{-\maltese}$ because it is arguably the most relevant concept when analysing multidimensional poverty, but similar patterns are found with $\hat{\rho}_3^{+\maltese}$ and $\hat{\rho}_3^{\maltese}$. The main conclusions that emerge from this figure are the following. First, in 2008 there was a clear positive relationship between the AROPE rate and lower orthant dependence, that is, those regions with a high proportion of households at risk of poverty and social exclusion tended to be also the regions where, on average, a household was more likely to be simultaneously poor in all dimensions (income, work intensity and material deprivation). Second, between 2008 and 2014, there was a remarkable increase in both the AROPE rate and $\hat{\rho}_3^{-\maltese}$ in the vast majority of the Spanish regions. Interestingly, the increase in $\hat{\rho}_3^{-\maltese}$ was particularly important in those regions which started with relatively low levels of both incidence of poverty and multivariate dependence between its dimensions. As a result, in 2014 we do not find that clear positive relationship between the AROPE rate and $\hat{\rho}_3^{-\maltese}$ found in 2008. Finally, we observe that, between 2014 and 2018, both the AROPE rate and $\hat{\rho}_3^{-\maltese}$ remained stable or decreased in most of the Spanish regions, without reaching, in most cases, the values of 2008. Furthermore, this decrease was, again, particularly important in the regions with relatively low levels of incidence of poverty and multivariate dependence between it dimensions. Then, in 2018 we observe again a clear

[8] The AROPE rate is calculated here as the proportion of households in our sample that are poor in at least one of the three dimensions considered.

positive relationship between the AROPE rate and $\hat{\rho}_3^{-}$✠.

To sump up, our results suggest that, in Spain as a whole and in many of its regions, the dependence between poverty dimensions as well as the incidence of poverty substantially increased when the economy was in recession but dropped little during the economic recovery. Moreover, this pattern is particularly clear for the coefficient of average lower orthant dependence, which captures the average probability of being simultaneously low-ranked in all dimensions and thus becomes specially relevant in a multidimensional poverty analysis. These findings are consistent with the asymmetric effect of the economic cycle on poverty and inequality in Spain recently reported in Ayala and Cantó (2022). As these authors point out, "if this dynamic is not reversed, poverty in Spain, which historically was characterized by being recurrent but transitory, runs the risk of becoming chronic, which would cause the effects of transitory shocks persist over time".

We believe that a detailed study of the link between social policy and dependence between poverty dimensions constitutes a promising avenue for further research, but goes beyond the scope of this study. Nevertheless, we hope that the results of this paper are a wake-up call for a rethinking of the public policy interventions to tackle the risk of cumulative deprivation and effectively fight the chronification of poverty in Spain. In this sense, given the clear association between income an the other two AROPE dimensions, it seems reasonable to think of an adequate minimum income scheme as a key instrument. Analyses of the effectiveness of these programs both in the EU (Almeida et al., 2022) and in Spain in particular (Ayala et al., 2022) can help to understand their shortcomings and generate ideas to overcome them. Moreover, effective minimum income benefits should be accompanied by an increase in the redistributive capacity of the fiscal system and by the extension of the countercyclical instruments; Ayala and Cantó (2022). Finally, in a highly decentrilised country such as Spain, it is crucial to improve the coordination between central, regional and local public administration in the implementation of social policies.

- INSERT FIGURE 4 HERE -

4 Conclusions

In this paper, we complement the information given by the AROPE rate with measures of the dependence between its components. In particular, we focus on three multivariate extensions of Spearman's rho linked to the concept of orthant dependence. We apply these coefficients to analyse the evolution of multivariate dependence between the three dimensions of the AROPE rate (income, work intensity and material needs) in the Spanish regions over the period 2008-2018. Due to the non-continuous nature of some of the variables included in the AROPE rate, we use tie-corrected versions of these coefficients. Among the coefficients used, the coefficient of average lower orthant dependence becomes especially relevant when analysing multidimensional poverty, as it captures the rescaled average probability of being simultaneously low-ranked in all dimensions of poverty as compared to what this would be were those dimensions independent.

Several noteworthy results emerge from our analysis. First, regardless of the coefficient used, there is a positive multivariate association between poverty dimensions in all Spanish regions and over the whole period analysed. This means that low (high) values of income tend to occur simultaneously with low (high) values of the other two dimensions. Second, average lower orthant dependence tends to be higher than average upper orthant dependence, that is, the simultaneous occurrence of bad rankings in the three AROPE components -i.e., households with simultaneously low incomes, low work intensity and few no-deprivations- is more likely than the simultaneous occurrence of good rankings in the three components. Third, when we analyse the temporal evolution over the period 2008-2018, some cross-regional differences arise. In 2008, we find lower dependence in northern regions than in southern regions. Between 2008 and 2014, the Financial Crisis of 2008 entailed a generalised and significant increase in the average probability of being simultaneously low-ranked (high-ranked) in all poverty dimensions. By contrast, over the post-crisis period (2014-2018), the multivariate dependence between the three dimensions of the AROPE rate remained rather stable in many regions and even decrease in others. Nevertheless, in many regions, multivariate dependence was still significantly higher in 2018 than in 2008, because such dependence dropped little during the economic recovery. This pattern is specially clear for the coefficient of lower orthant dependence, which is particularly relevant when analysing multidimensional poverty. This finding suggests the existence of an asymmetric effect of the economic cycle on the dependence between poverty dimensions. Finally, we find that the regions with higher incidence of multidimensional poverty tend to also experience a higher degree of multivariate dependence between its dimensions.

These findings should be a wake-up call for the need to complement the AROPE rate with other measures of the dependence structure of poverty and also to perform analysis at the sub-national level to reveal patterns that remain concealed when analysing aggregate data for the country as a whole. Following this approach could help to adequately design effective poverty-reduction policies at both country and regional levels.

Bibliographical references

Alkire, S. and Foster, J. E. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7-8):476–487.

Almeida, V., De Poli, S., and Martín, A. H. (2022). The effectiveness of minimum income schemes in the eu. Working Papers 627, ECINEQ, Society for the Study of Economic Inequality.

Atkinson, A. B. and Bourguignon, F. (1982). The comparison of multi-dimensioned distributions of economic status. *The Review of Economic Studies*, 49(2):183–201.

Ayala, L. and Cantó, O. (2022). *Radiografía de medio siglo de desigualdad en España*. Palma: Observatorio Social de la Fundación La Caixa.

Ayala, L. and Jurado, A. (2020). Diferencias de desigualdad y bienestar en las regiones españolas. In *4º Informe sobre la Desigualdad en España. Una perspectiva territorial*, pages 121–150. Fundación Alternativas, Madrid.

Ayala, L., Jurado, A., and Pérez-Mayo, J. (2011). Income poverty and multidimensional deprivation: Lessons from cross-regional analysis. *Review of Income and Wealth*, 57(1):40–60.

Ayala, L., Jurado, A., and Pérez Mayo, J. (2022). El ingreso mínimo vital: adecuación y cobertura. *Papeles de Economía Española*, (172):155–172.

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

Chakravarty, S. R. (2018). *Analyzing multidimensional well-being: A quantitative approach*. Wiley, New York.

D 'Agostino, A., Deluca, G., and Guégan, D. (2022). Estimating lower tail dependence between pairs of poverty dimensions in Europe. *Review of Income and Wealth*.

Decancq, K. (2014). Copula-based measurement of dependence between dimensions of well-being. *Oxford Economic Papers*, 66(3):681–701.

Decancq, K. (2020). Measuring cumulative deprivation and affluence based on the diagonal dependence diagram. *Metron*, (78):103–117.

Decancq, K. (2022). Cumulative deprivation: identification and aggregation. Working Papers 604, ECINEQ.

Dolati, A. and Úbeda-Flores, M. (2006). On measures of multivariate concordance. *Journal of Probability and Statistical Science*, 4(2):147–163.

Duclos, J.-Y. and Tiberti, L. (2016). Multidimensional poverty indices: A critical assessment. In *The Oxford Handbook of Well-Being and Public Policy*. Oxford University Press, Oxford.

Durante, F., Nelsen, R. B., Quesada-Molina, J. J., and Úbeda-Flores, M. (2014). Pairwise and global dependence in trivariate copula models. In Laurent, A., Strauss, O., Bouchon-Meunier, B., and Yager, R. R., editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 243–251, Cham. Springer.

Ferreira, F. H. G. and Lugo, M. A. (2013). Multidimensional poverty analysis: Looking for a middle ground. *World Bank Research Observer*, 28(2):220–235.

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

García-Gómez, C., Pérez, A., and Prieto-Alaiz, M. (2021). Copula-based analysis of multivariate dependence patterns between dimensions of poverty in Europe. *Review of Income and Wealth*, 67(1):165–195.

García-Gómez, C., Pérez, A., and Prieto-Alaiz, M. (2022). The evolution of poverty in the EU-28: a further look based on multivariate tail dependence. Working Papers 605, ECINEQ, Society for the Study of Economic Inequality.

Genest, C., Nešlehová, J., and Rémillard, B. (2013). On the estimation of Spearman's rho and related tests of independence for possibly discontinuous multivariate data. *Journal of Multivariate Analysis*, 117:214 – 228.

Goedemé, T. (2013). How much confidence can we have in EU-SILC? Complex sample designs and the standard error of the Europe 2020 poverty indicators. *Social Indicators Research*, 110(1):89–110.

Joe, H. (1990). Multivariate concordance. *Journal of Multivariate Analysis*, 35(1):12–30.

Joe, H. (2014). *Dependence Modeling with Copulas*. Chapman and Hall, London, UK.

Kleiber, C. and Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*. John Wiley & Sons, Hoboken, NJ.

Mashreghi, Z., Haziza, D., and Léger, C. (2016). A survey of bootstrap methods in finite population sampling. *Statistics Surveys*, 10:1 – 52.

Mesfioui, M. and Quessy, J.-F. (2010). Concordance measures for multivariate non-continuous random vectors. *Journal of Multivariate Analysis*, 101(10):2398 – 2410.

Nelsen, R. B. (1996). Nonparametric measures of multivariate association. In Rüschendorf, L., Schweizer, B., and Taylor, M. D., editors, *Distributions with Given Marginals and Related Topics*, volume 28, pages 223–232. Institute of Mathematical Statistics.

Nelsen, R. B. (2002). Concordance and copulas: A survey. In Cuadras, C. M., Fortiana, J., and Rodriguez-Lallena, J. A., editors, *Distributions With Given Marginals and Statistical Modelling*, pages 169–177. Springer, Dordrecht.

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

Nelsen, R. B. and Úbeda-Flores, M. (2012). Directional dependence in multivariate distributions. *Annals of the Institute of Statistical Mathematics*, 64(3):677–685.

Nešlehová, J. (2007). On rank correlation measures for non-continuous random variables. *Journal of Multivariate Analysis*, 98(3):544–567.

Osier, G., Berger, Y. G., and Goedeme, T. (2013). Standard error estimation for the eu-silc indicators of poverty and social exclusion. In *Proceedings of the Eurostat Methodologies and Working Papers 13-024*, Eurostat, Luxembourg.

Pérez, A. and Prieto-Alaiz, M. (2016a). Measuring the dependence among dimensions of welfare: A study based on Spearman's footrule and Gini's gamma. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 24(Suppl. 1):87–105.

Pérez, A. and Prieto-Alaiz, M. (2016b). A note on nonparametric estimation of copula-based multivariate extensions of Spearman's rho. *Statistics & Probability Letters*, 112:41–50.

Quessy, J.-F. (2009). Tests of multivariate independence for ordinal data. *Communications in Statistics - Theory and Methods*, 38(19):3510–3531.

Rao, J., Wu, C., and Yue, K. (1992). Some recent work on resampling methods for complex surveys. *Survey methodology*, 18(2):209–217.

Rust, K. F. and Rao, J. (1996). Variance estimation for complex surveys using replication techniques. *Statistical methods in medical research*, 5(3):283–310.

Schweizer, B. and Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics*, 9(4):879–885.

Seth, S. (2013). A class of distribution and association sensitive multidimensional welfare indices. *The Journal of Economic Inequality*, 11(2):133–162.

Seth, S. and Santos, M. (2019). On the interaction between focus and distributional properties in multidimensional poverty measurement. *Social Indicators Research*, 145:503–521.

Tkach, K. and Gigliarano, C. (2018). Multidimensional poverty measurement: dependence between well-being dimensions using copula function. *Rivista Italiana di Economia Demografia e Statistica*, 72(3).

Appendix A

Let (X_1, X_2) be a random vector with joint cumulative distribution function F and continuous marginal distribution functions F_1, F_2, so that $F(x_1, x_2) = p(X_1 \leq x_1, X_2 \leq x_2)$, for all $(x_1, x_2) \in \mathbb{R}^2$ and $F_i(x_i) = p(X_i \leq x_i)$, for $i = 1, 2$. Let $\overline{F}, \overline{F}_1, \overline{F}_2$ be their corresponding survival functions, namely, $\overline{F}(x_1, x_2) = p(X_1 > x_1, X_2 > x_2)$ and $\overline{F}_i(x_i) = p(X_i > x_i) = 1 - F_i(x_i)$, for $i = 1, 2$. We say that X_1 and X_2 are positively quadrant dependent (PQD) - see Nelsen (2006) and Joe (2014) - if, for all $(x_1, x_2) \in \mathbb{R}^2$,

$$F(x_1, x_2) \geq F_1(x_1)F_2(x_2). \tag{A1}$$

Note that, since $\overline{F}(x_1, x_2) = 1 - F(x_1) - F(x_2) + F(x_1, x_2)$, the inequality above is equivalent to

$$\overline{F}(x_1, x_2) \geq \overline{F}_1(x_1)\overline{F}_2(x_2). \tag{A2}$$

Intuitively, X_1 and X_2 are PQD if the probability that they are simultaneously small (or simultaneously large) is at least as great as it would be were they independent. So, in a sense, the difference $F(x_1, x_2) - F_1(x_1)F_2(x_2)$ (or equivalently, the difference $\overline{F}(x_1, x_2) - \overline{F}_1(x_1)\overline{F}_2(x_2)$) measures "local" quadrant dependence at each point $(x_1, x_2) \in \mathbb{R}^2$; see Nelsen (1996). Because of that, the formula of the Spearman's rho in (2) allows to interpret this coefficient as a measure of "average" quadrant dependence.

When we move to a multivariate framework with more than two variables involved, we have "orthants" rather than quadrants and the two concepts in (A1) and (A2) are no longer equivalent, so other concepts are required. To introduce such concepts, we need further notation.

Let $\boldsymbol{X} = (X_1, \ldots, X_d)$ be a $d-$dimensional random vector with joint cumulative distribution function F and continuous marginal distribution functions $F_1, \ldots, F_d$, so that

$F(\boldsymbol{x}) = p(X_1 \leq x_1, \ldots, X_d \leq x_d)$ for any real vector $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $F_i(x_i) = p(X_i \leq x_i)$, $i = 1, \ldots, d$. Let $\overline{F}, \overline{F}_1, \cdots, \overline{F}_d$ be the corresponding survival functions, namely $\overline{F}(\boldsymbol{x}) = p(X_1 > x_1, \ldots, X_d > x_d)$ and $\overline{F}_i(x_i) = p(X_i > x_i) = 1 - F_i(x_i)$, for $i = 1, ..., d$. We say that (Nelsen, 2006; Joe, 2014)

(1) $\mathbf{X}$ is positively *lower orthant dependent* (PLOD) if $F(\boldsymbol{x}) \geq F_1(x_1) \cdots F_d(X_d)$, for all $\boldsymbol{x} \in \mathbb{R}^d$, that is, if the probability that the variables $X_1, ..., X_d$ are simultaneously small is at least as great as it would be in the case of independence;

(2) $\mathbf{X}$ is positively *upper orthant dependent* (PUOD) if $\overline{F}(\boldsymbol{x}) \geq \overline{F}_1(x_1) \cdots \overline{F}_d(X_d)$, for all $\boldsymbol{x} \in \mathbb{R}^d$, that is, if the probability that the variables $X_1, ..., X_d$ are simultaneously large is at least as great as it would be in the case of independence;

(3) $\mathbf{X}$ is positively *orthant dependent* (POD) if both previous inequalities hold.

Intuitively, PLOD (PUOD) means that $(X_1, \ldots, X_d)$ are more likely simultaneously to take small (large) values, compared with a vector of independent random variables with the same corresponding univariate margins. For $d = 2$, PLOD and PUOD are equivalent to POD, but this is not the case for $d > 2$. For poverty analysis, PLOD will be the more relevant concept, as it focuses on the proportion of households in the population that are simultaneusly "low ranked" in all poverty dimensions.

Based on the definitions above, the differences $[F(\mathbf{x}) - F_1(x_1) \cdots F_d(x_d)]$ and $[\bar{F}(\mathbf{x}) - \bar{F}_1(x_1) \cdots \bar{F}_d(x_d)]$ can be regarded as measures of "local" lower and upper orthant dependence, respectively, at each point $\mathbf{x} \in R^d$. What Nelsen (1996) proposed was to take a rescaled average of these differences with respect to the marginal distributions to obtain measures of average lower and upper orthant dependence, respectively. Noticeably, these two measures are the coefficients ρ_d^+ and ρ_d^- in (4) and (5), respectively. In turn, Dolati and Ubeda Flores (2006) considered a rescaled average of the sum of the two differences above and worked out the coefficient ρ_d in (6) as an average orthant dependence measure.

Appendix B

In this Appendix, we shed light on the concepts behind the non-continuous versions of Spearman's rho coefficients in equations (7)-(9) and their estimators in equations (11)-(13). To simplify the notation and make the explanation clearer, discussions below are restricted to the bivariate case only ($d = 2$). The extension to a multivariate framework comes up in a natural way.

To start with, let us recall that, given two continuous random variables, X_1 and X_2, with marginal distribution functions F_1 and F_2, respectively, the random variables defined by the probability integral transformations, $F_1(X_1)$ and $F_2(X_2)$, follow standard uniform distributions $U(0,1)$. Hence,

$$\begin{aligned} E(F_1(X_1)) &= E(F_2(X_2)) = 1/2, \\ Var(F_1(X_1)) &= Var(F_2(X_2)) = 1/12. \end{aligned}$$

In this setting, the population version of bivariate Spearman's rho for X_1 and X_2 can be defined as the Pearson's correlation coefficient between the position variables $F_1(X_1)$ and $F_2(X_2)$ - see Section 2 - that is:

$$\rho_S = \frac{Cov(F_1(X_1), F_2(X_2))}{\sqrt{Var(F_1(X_1))Var(F_2(X_2))}} = 12E(F_1(X_1)F_2(X_2)) - 3. \tag{B1}$$

When the variables involved are non-continuous, the transformed variables $F_1(X_1)$ and $F_2(X_2)$ are no-longer uniform and the identity above no-longer holds. To overcome this drawback, Nešlehová (2007) and Genest et al. (2013) rely on an alternative transformation of an arbitrary (possibly non-continuous) random variable that would lead to the uniform distribution $U(0,1)$. Using this technique, a modified version of Spearman's rho for non-continuous random variables is worked out. The procedure is as follows.

Let U_1 be a uniform random variable $U(0,1)$ independent of X_1, defined on some common probability space (Ω, A, P), and consider the transformation $\psi : [-\infty, \infty] \times [0,1] \to [0,1]$ given by

$$\psi(x_1, u_1) = P[X_1 < x_1] + u_1 P[X_1 = x_1] = F_1(x_1^-) + u_1 \Delta F_1(x_1),$$

where $\Delta F_1(x_1) = F_1(x_1) - F_1(x_1^-)$. In a similar way, we define the transformed variable $\psi(X_2, U_2)$, where U_2 is independent of U_1. Nešlehová (2007) shows that the random variables $\psi(X_1, U_1)$ and $\psi(X_2, U_2)$ are both uniformly distributed $U(0,1)$ and so, we have

$$\begin{aligned} E(\psi(X_1, U_1)) &= E(\psi(X_2, U_2)) = 1/2 \\ Var(\psi(X_1, U_1)) &= Var(\psi(X_2, U_2)) = 1/12 \end{aligned}$$

Therefore, we can define the analogous version of the Spearman's rho in equation (B1) for the transformed variables, namely

$$\rho^{\maltese} = 12E(\psi(X_1, U_1)\psi(X_2, U_2)) - 3. \quad \text{(B2)}$$

To figure out the relationship of this coefficient with the distribution functions of the original variables X_1 and X_2, first note that, upon conditioning on X_1, one deduces

$$\begin{aligned} E(\psi(X_1, U_1)|X_1{=}x_1) &= E(\psi(x_1, U_1)|X_1{=}x_1) = E(\psi(x_1, U_1)) \\ &= E(F_1(x_1^-) + U_1 \Delta F_1(x_1)) = F_1(x_1^-) + E(U_1)\Delta F_1(x_1) \\ &= \frac{F_1(x_1^-) + F_1(x_1)}{2} = \widetilde{F}_1(x_1). \end{aligned} \quad \text{(B3)}$$

Using a similar argument and taking into account that U_1 and U_1 are independent, one finds, upon conditioning on $\mathbf{X} = (X_1,X_2)$,

$$E(\psi(X_1,U_1)\psi(X_2,U_2)|\{X_1{=}x_1,X_2=x_2\})) = E(\psi(x_1,U_1))E(\psi(x_2,U_2))$$
$$= \frac{F_1(x_1^-)+F_1(x_1)}{2}\frac{F_2(x_2^-)+F_2(x_2)}{2} = \widetilde{F}_1(x_1)\widetilde{F}_2(x_2) \tag{B4}$$

Now, applying the Law of Iterated Expectations on (B4), one deduces

$$E(\psi(X_1,U_1)\psi(X_2,U_2)) = E(E(\psi(X_1,U_1)\psi(X_2,U_2)|\mathbf{X})) = E(\widetilde{F}_1(X_1)\widetilde{F}_2(X_2)). \tag{B5}$$

Putting (B5) back into (B2), a new formula of the corrected Spearman's rho arises

$$\rho^{\maltese} = 12E(\widetilde{F}_1(X_1)\widetilde{F}_2(X_2)) - 3. \tag{B6}$$

Furthermore, from the uniformity of $\psi(X_1,U_1)$ and appying the Law of Iterated Expectations on equation (B3), the following identity will come up:

$$\frac{1}{2} = E(\psi(X_1,U_1)) = E(E(\psi(X_1,U_1)|X_1))) = E(\widetilde{F}_1(X_1)).$$

In a similar fashion, it is proved that $E(\widetilde{F}_2(X_2)) = \frac{1}{2}$. Therefore, it immediately follows that

$$E(1-\widetilde{F}_1(X_1))(1-\widetilde{F}_2(X_2)) = E(\widetilde{F}_1(X_1)\widetilde{F}_2(X_2)). \tag{B7}$$

On combining now (B6) and (B7), the Spearman's coefficient $\rho^{\maltese}$ can be alternatively written in terms of the survival functions as

$$\rho^{\maltese} = 12E(1-\widetilde{F}_1(X_1))(1-\widetilde{F}_2(X_2)) - 3. \tag{B8}$$

The multivariate tie-corrected versions of Spearman's rho given in equations (7)-(8) in

Section 2, arise by generalising (B6) and (B8), respectively, for $d \geq 2$, as

$$\rho_d^{+\maltese} = \frac{(d+1)}{2^d-(d+1)}\left[2^d E\left(\prod_{i=1}^{d}\widetilde{F}_i(X_i)\right)-1\right],$$
$$\rho_d^{-\maltese} = \frac{(d+1)}{2^d-(d+1)}\left[2^d E\left(\prod_{i=1}^{d}(1-\widetilde{F}_i(X_i))\right)-1\right].$$

The third multivariate Spearman's rho in equation (9) in Section 2 is just the average of the two coefficients above. Now, a plug-in estimator of the coefficients above is obtained by replacing $\widetilde{F}_i(X_i)$ with its empirical analogue $\widetilde{F}_{in}(X_{ij})$, for $i = 1, .., d$ and $j = 1, ..., n$, as explained in Section 2.

Table 1: Estimated value of $\rho_3^{-\maltese}$ in Spain and its regions for years 2008, 2014 and 2018 and t-tests for its variation over the periods 2008-2014, 2014-2018 and 2008-2018.

	$\hat{\rho}_3^{-\maltese}$					
Region	**2008**	**2014**	**2018**	**t-test 08-14**	**t-test 14-18**	**t-test 08-18**
GALICIA	0.331	0.480	0.446	3.414**	-0.706	2.508**
	(0.028)	(0.033)	(0.036)	(0.000)	(0.240)	(0.006)
ASTURIAS	0.218	0.405	0.403	3.164**	-0.035	3.291**
	(0.034)	(0.048)	(0.045)	(0.001)	(0.486)	(0.000)
CANTABRIA	0.324	0.487	0.443	2.143*	-0.600	1.619
	(0.054)	(0.054)	(0.050)	(0.016)	(0.274)	(0.053)
BASQUE COUNTRY	0.271	0.487	0.374	4.021**	-2.097*	1.911*
	(0.038)	(0.038)	(0.038)	(0.000)	(0.018)	(0.028)
NAVARRE	0.279	0.482	0.264	2.978**	-2.882**	-0.189
	(0.049)	(0.048)	(0.059)	(0.001)	(0.002)	(0.425)
LA RIOJA	0.334	0.460	0.428	1.994*	-0.480	1.454
	(0.043)	(0.047)	(0.049)	(0.023)	(0.315)	(0.073)
ARAGON	0.282	0.465	0.348	3.063**	-1.839*	1.076
	(0.040)	(0.044)	(0.046)	(0.001)	(0.033)	(0.141)
MADRID	0.323	0.435	0.435	2.914**	0.000	2.999**
	(0.029)	(0.026)	(0.024)	(0.002)	(0.500)	(0.001)
CASTILE AND LEON	0.313	0.457	0.362	3.242**	-2.036*	1.161
	(0.028)	(0.034)	(0.031)	(0.001)	(0.021)	(0.123)
CASTILLA LA MANCHA	0.385	0.494	0.493	2.135*	-0.009	2.032*
	(0.035)	(0.038)	(0.041)	(0.016)	(0.496)	(0.021)
EXTREMADURA	0.403	0.446	0.438	0.647	-0.138	0.547
	(0.052)	(0.042)	(0.039)	(0.259)	(0.445)	(0.292)
CATALONIA	0.327	0.485	0.356	4.513**	-3.803**	0.830
	(0.025)	(0.024)	(0.024)	(0.000)	(0.000)	(0.203)
VALENCIAN COMMUNITY	0.364	0.503	0.367	3.124**	-2.820**	0.061
	(0.031)	(0.031)	(0.037)	(0.001)	(0.002)	(0.476)
BALEARIC ISLANDS	0.271	0.466	0.433	2.957**	-0.433	2.173*
	(0.044)	(0.049)	(0.060)	(0.002)	(0.333)	(0.015)
ANDALUSIA	0.424	0.460	0.494	1.118	1.066	2.207*
	(0.022)	(0.023)	(0.023)	(0.132)	(0.143)	(0.014)
MURCIA	0.286	0.481	0.449	3.429**	-0.438	2.072*
	(0.044)	(0.037)	(0.065)	(0.000)	(0.331)	(0.019)
CEUTA	0.572	0.423	0.643	-1.269	2.113*	0.764
	(0.076)	(0.090)	(0.054)	(0.102)	(0.017)	(0.222)
MELILLA	0.417	0.634	0.536	2.078*	-1.082	1.075
	(0.086)	(0.060)	(0.069)	(0.019)	(0.140)	(0.141)
CANARY ISLANDS	0.434	0.399	0.494	-0.585	1.279	0.861
	(0.039)	(0.048)	(0.058)	(0.279)	(0.100)	(0.195)
SPAIN	0.368	0.500	0.457	10.336**	-3.251**	6.613**
	(0.009)	(0.009)	(0.010)	(0.000)	(0.001)	(0.000)

Note: Standard errors for the coefficients and p-values for the one-side t-test are displayed in parentheses.

Table 2: Estimated value of $\rho_3^{+\maltese}$ in Spain and its regions for years 2008, 2014 and 2018 and t-tests for its variation over the periods 2008-2014, 2014-2018 and 2008-2018.

			$\hat{\rho}_3^{+\maltese}$			
Region	**2008**	**2014**	**2018**	**t-test 08-14**	**t-test 14-18**	**t-test 08-18**
GALICIA	0.328	0.465	0.391	3.605**	-1.747*	1.513
	(0.027)	(0.027)	(0.032)	(0.000)	(0.040)	(0.065)
ASTURIAS	0.187	0.341	0.365	2.671**	0.405	3.524**
	(0.033)	(0.047)	(0.038)	(0.004)	(0.343)	(0.000)
CANTABRIA	0.318	0.451	0.373	1.860*	-1.030	0.850
	(0.042)	(0.057)	(0.049)	(0.031)	(0.151)	(0.198)
BASQUE COUNTRY	0.247	0.411	0.311	3.721**	-2.223**	1.459
	(0.030)	(0.032)	(0.032)	(0.000)	(0.013)	(0.072)
NAVARRE	0.257	0.400	0.217	2.404**	-2.962**	-0.648
	(0.042)	(0.043)	(0.045)	(0.008)	(0.002)	(0.259)
LA RIOJA	0.300	0.399	0.379	1.642*	-0.324	1.344
	(0.042)	(0.044)	(0.042)	(0.050)	(0.373)	(0.089)
ARAGON	0.274	0.420	0.310	2.894**	-2.009*	0.701
	(0.033)	(0.038)	(0.039)	(0.002)	(0.022)	(0.242)
MADRID	0.308	0.377	0.374	1.968*	-0.094	1.987*
	(0.025)	(0.024)	(0.022)	(0.025)	(0.463)	(0.023)
CASTILE AND LEON	0.290	0.413	0.307	3.067**	-2.498**	0.427
	(0.026)	(0.030)	(0.030)	(0.001)	(0.006)	(0.335)
CASTILLA LA MANCHA	0.391	0.467	0.422	1.606	-0.853	0.611
	(0.031)	(0.036)	(0.039)	(0.054)	(0.197)	(0.271)
EXTREMADURA	0.359	0.448	0.453	1.493	0.097	1.661*
	(0.045)	(0.039)	(0.035)	(0.068)	(0.461)	(0.048)
CATALONIA	0.299	0.451	0.300	4.929**	-5.073**	0.035
	(0.022)	(0.021)	(0.021)	(0.000)	(0.000)	(0.486)
VALENCIAN COMMUNITY	0.341	0.474	0.304	3.199**	-3.545**	-0.774
	(0.029)	(0.030)	(0.037)	(0.001)	(0.000)	(0.220)
BALEARIC ISLANDS	0.232	0.411	0.344	2.752**	-0.923	1.616
	(0.043)	(0.049)	(0.054)	(0.003)	(0.178)	(0.053)
ANDALUSIA	0.401	0.459	0.481	1.861*	0.664	2.525**
	(0.021)	(0.023)	(0.023)	(0.031)	(0.253)	(0.006)
MURCIA	0.266	0.481	0.475	3.942**	-0.112	3.694**
	(0.041)	(0.036)	(0.039)	(0.000)	(0.455)	(0.000)
CEUTA	0.560	0.464	0.597	-0.847	1.227	0.356
	(0.076)	(0.084)	(0.068)	(0.198)	(0.110)	(0.361)
MELILLA	0.405	0.582	0.460	1.807*	-1.322	0.552
	(0.075)	(0.064)	(0.067)	(0.035)	(0.093)	(0.291)
CANARY ISLANDS	0.421	0.388	0.425	-0.545	0.514	0.053
	(0.036)	(0.050)	(0.051)	(0.293)	(0.304)	(0.479)
SPAIN	0.343	0.471	0.402	10.870**	-5.699**	4.858**
	(0.009)	(0.008)	(0.009)	(0.000)	(0.000)	(0.000)

Note: Standard errors for the coefficients and p-values for the one-side t-test are displayed in parentheses.

Table 3: Estimated value of $\rho_3^{\maltese}$ in Spain and its regions for years 2008, 2014 and 2018 and t-tests for its variation over the periods 2008-2014, 2014-2018 and 2008-2018.

			$\hat{\rho}_3^{\maltese}$			
Region	**2008**	**2014**	**2018**	**t-test 08-14**	**t-test 14-18**	**t-test 08-18**
GALICIA	0.329	0.472	0.418	3.620**	-1.216	2.083*
	(0.026)	(0.029)	(0.033)	(0.000)	(0.112)	(0.019)
ASTURIAS	0.203	0.373	0.384	2.999**	0.180	3.498**
	(0.033)	(0.046)	(0.040)	(0.001)	(0.429)	(0.000)
CANTABRIA	0.321	0.469	0.408	2.070*	-0.836	1.293
	(0.046)	(0.054)	(0.049)	(0.019)	(0.202)	(0.098)
BASQUE COUNTRY	0.259	0.449	0.342	3.970**	-2.195**	1.742*
	(0.033)	(0.035)	(0.034)	(0.000)	(0.014)	(0.041)
NAVARRE	0.268	0.441	0.241	2.766**	-2.985**	-0.401
	(0.044)	(0.044)	(0.050)	(0.003)	(0.001)	(0.344)
LA RIOJA	0.317	0.430	0.404	1.863*	-0.415	1.434
	(0.041)	(0.044)	(0.044)	(0.031)	(0.339)	(0.076)
ARAGON	0.278	0.443	0.329	3.062**	-1.965*	0.928
	(0.036)	(0.040)	(0.042)	(0.001)	(0.025)	(0.177)
MADRID	0.315	0.406	0.405	2.530**	-0.046	2.590**
	(0.026)	(0.024)	(0.022)	(0.006)	(0.482)	(0.005)
CASTILE AND LEON	0.302	0.435	0.335	3.249**	-2.325**	0.831
	(0.026)	(0.031)	(0.030)	(0.001)	(0.010)	(0.203)
CASTILLA LA MANCHA	0.388	0.480	0.458	1.962*	-0.435	1.394
	(0.031)	(0.035)	(0.039)	(0.025)	(0.332)	(0.082)
EXTREMADURA	0.381	0.447	0.445	1.074	-0.026	1.101
	(0.047)	(0.039)	(0.035)	(0.142)	(0.489)	(0.135)
CATALONIA	0.313	0.468	0.328	4.838**	-4.520**	0.470
	(0.023)	(0.022)	(0.022)	(0.000)	(0.000)	(0.319)
VALENCIAN COMMUNITY	0.352	0.488	0.335	3.245**	-3.266**	-0.361
	(0.029)	(0.030)	(0.036)	(0.001)	(0.001)	(0.359)
BALEARIC ISLANDS	0.251	0.439	0.388	2.939**	-0.687	1.947*
	(0.043)	(0.047)	(0.056)	(0.002)	(0.246)	(0.026)
ANDALUSIA	0.412	0.460	0.488	1.526	0.889	2.436**
	(0.021)	(0.022)	(0.023)	(0.063)	(0.187)	(0.007)
MURCIA	0.276	0.481	0.462	3.791**	-0.314	2.856**
	(0.041)	(0.035)	(0.051)	(0.000)	(0.377)	(0.002)
CEUTA	0.566	0.443	0.620	-1.093	1.730*	0.573
	(0.074)	(0.084)	(0.057)	(0.137)	(0.042)	(0.283)
MELILLA	0.411	0.608	0.498	2.001*	-1.244	0.843
	(0.079)	(0.059)	(0.066)	(0.023)	(0.107)	(0.200)
CANARY ISLANDS	0.428	0.393	0.459	-0.583	0.926	0.489
	(0.036)	(0.047)	(0.054)	(0.280)	(0.177)	(0.312)
SPAIN	0.356	0.486	0.430	10.898**	-4.549**	5.932**
	(0.009)	(0.008)	(0.009)	(0.000)	(0.000)	(0.000)

Note: Standard errors for the coefficients and p-values for the one-side t-test are displayed in parentheses.

Figure 1: Evolution of $\hat{\rho}_3^{-✠}$ (blue), $\hat{\rho}_3^{+✠}$ (green) and $\hat{\rho}_3^{✠}$ (red) in Spain and its regions over the period 2008-2018.

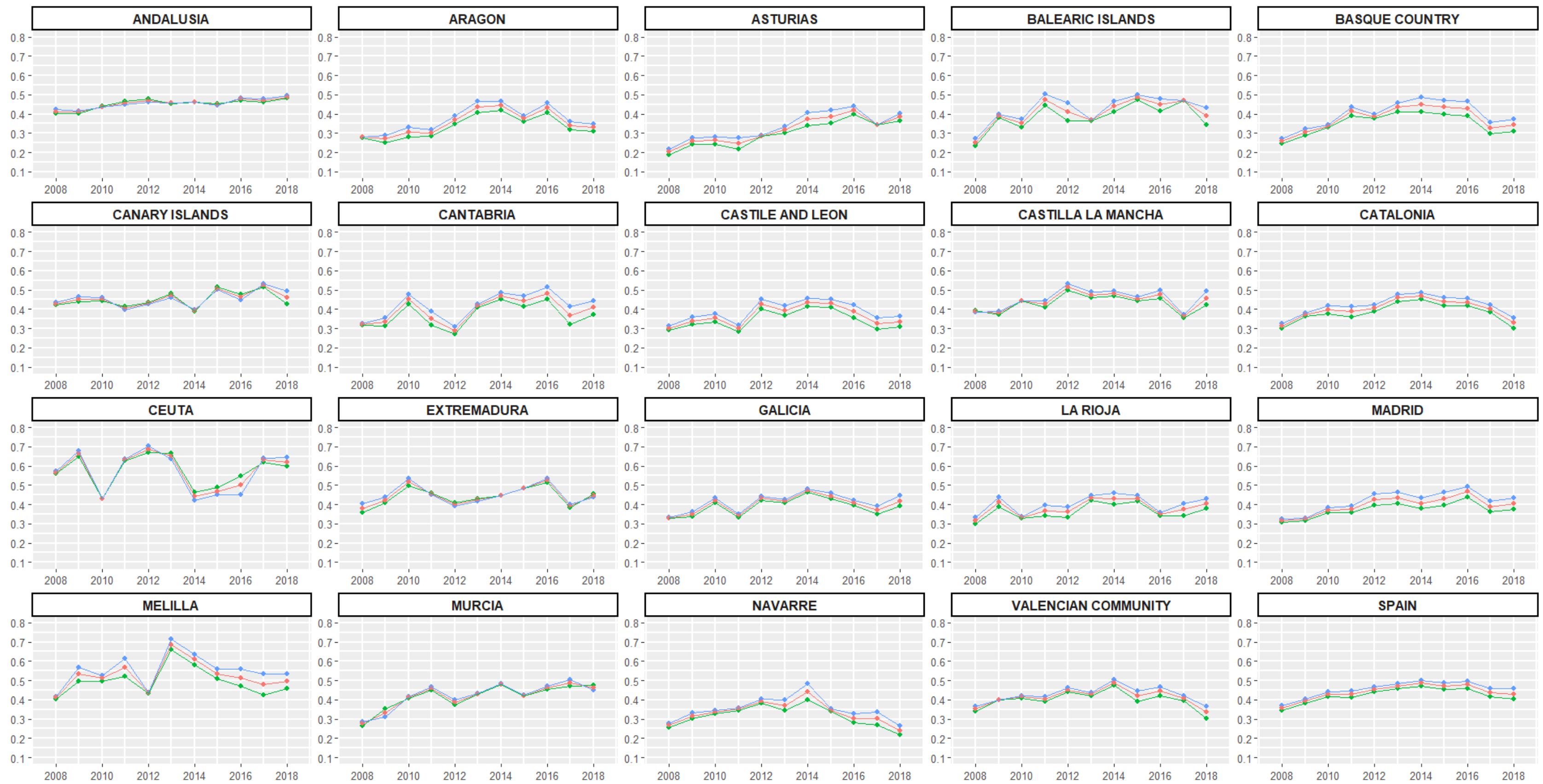

Figure 2: Cross-regional differences in the level of $\hat{\rho}_3^{-\maltese}$ for years 2008, 2014 and 2018.

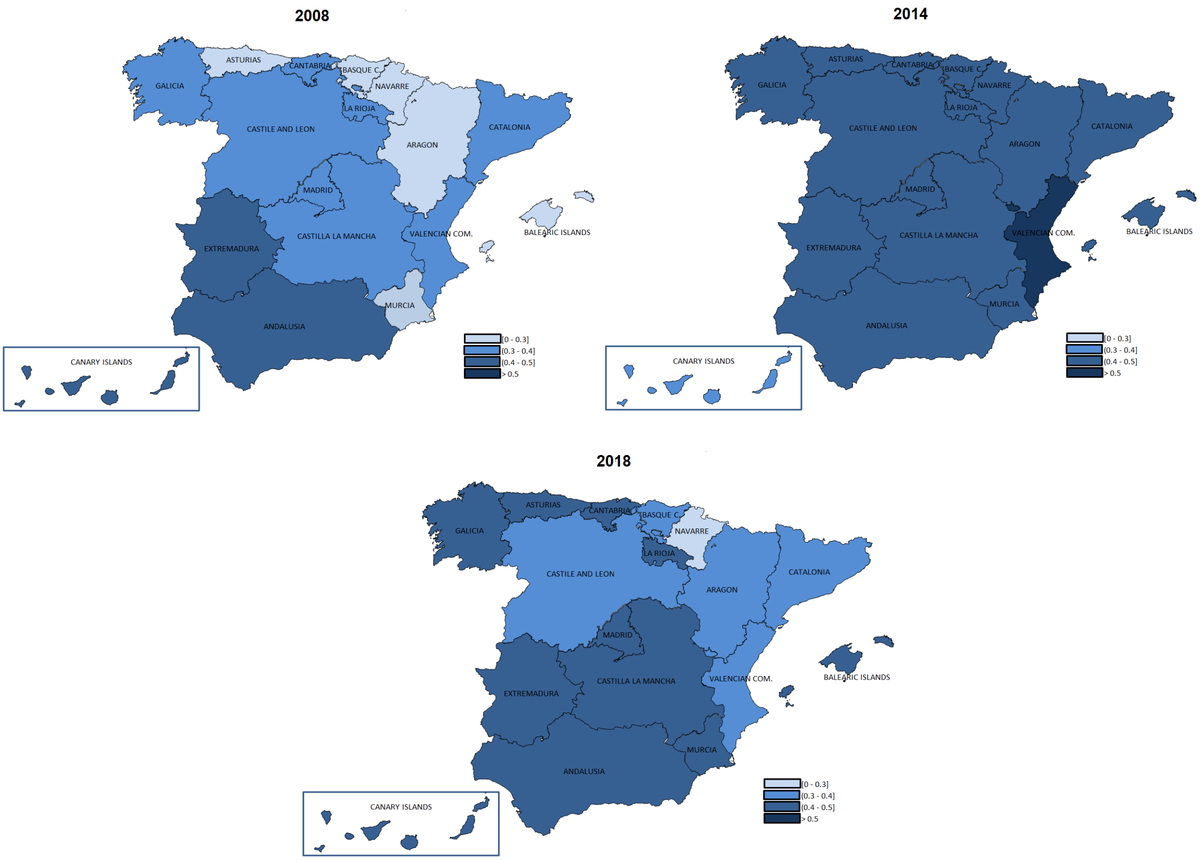

Figure 3: Evolution of $\hat{\rho}^{\maltese}_{income,work}$ (green), $\hat{\rho}^{\maltese}_{income,no-deprivation}$ (red) and $\hat{\rho}^{\maltese}_{work,no-deprivation}$ (blue) in Spain and its regions over the period 2008-2018.

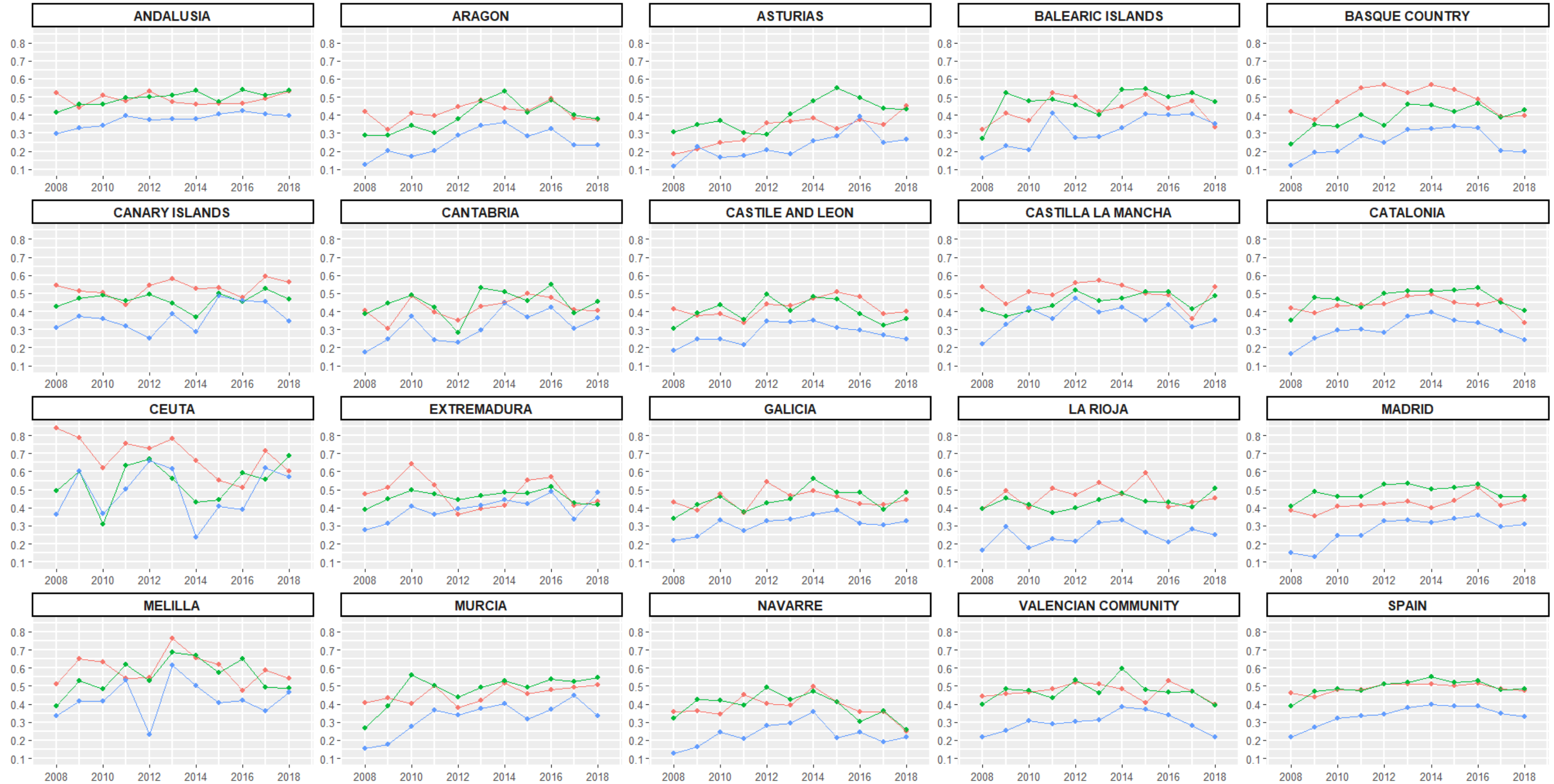

Figure 4: Relationship between AROPE rate and $\hat{\rho}_3^{-\maltese}$ for Spanish regions and years 2008 (left), 2014 (center) and 2018 (right).