ECINEQ

Society for the Study of Economic Inequality

Working Paper Series

Studying the evolution of cumulative deprivation among European countries with a copula-based approach.

Giovanna Scarcilli

ECINEQ 2024 667

ECINEQ 2024 667
2024 February
www.ecineq.org

Studying the evolution of cumulative deprivation among European countries with a copula-based approach.

Giovanna Scarcilli
University of Milan

Abstract

This paper adopts a novel copula-based technique to measure multidimensional dependence among facets of cumulative deprivation and provides empirical insights on this phenomenon from a cross-country and time perspective. Cumulative deprivation is a condition of simultaneous relative poverty across multiple dimensions of life. The dimensions taken into account are: disposable income, health status, housing quality, job conditions and educational attainment. Multidimensional deprivation is evaluated with the downward diagonal dependence index (DDDI). This index provides a measure of statistical dependence among the considered dimensions specifically for the bottom part of the overall joint distribution. The empirical application focuses on Belgium, France, Germany, Italy, Spain, Czech republic, Romania and Sweden from 2007 to 2019 using EU-SILC data. In the considered period cumulative deprivation and multidimensional dependence both show a growing trend. The rise in the proportion of people who are deprived in many different dimensions strengthens their statistical association. The growth of multidimensional dependence is concerning and requires reconsidering the actual welfare states and their policies.

Keyword: Cumulative deprivation, Copula function, Multidimensional dependence, Cross time and country comparisons

JEL Cassification: D63, I32

Studying the evolution of cumulative deprivation among European countries with a copula-based approach.*

Giovanna Scarchilli †

February 22, 2024

Abstract

This paper adopts a novel copula-based technique to measure multidimensional dependence among facets of cumulative deprivation and provides empirical insights on this phenomenon from a cross-country and time perspective. Cumulative deprivation is a condition of simultaneous relative poverty across multiple dimensions of life. The dimensions taken into account are: disposable income, health status, housing quality, job conditions and educational attainment. Multidimensional deprivation is evaluated with the downward diagonal dependence index (DDDI). This index provides a measure of statistical dependence among the considered dimensions specifically for the bottom part of the overall joint distribution. The empirical application focuses on Belgium, France, Germany, Italy, Spain, Czech republic, Romania and Sweden from 2007 to 2019 using EU-SILC data. In the considered period cumulative deprivation and multidimensional dependence both show a growing trend. The rise in the proportion of people who are deprived in many different dimensions strengthens their statistical association. The growth of multidimensional dependence is concerning and requires reconsidering the actual welfare states and their policies.

JEL-Classification: D63, I32

Keywords: Cumulative deprivation, Copula function, Multidimensional dependence, Cross time and country comparisons

*Acknowledgements: I am grateful to Koen Decancq for the supervision on this project during my visiting PhD period at the Herman Deleeck Centre for Social Policy, University of Antwerp. Many thanks to Francisco H. G. Ferreira, Paolo Brunori, Flaviana Palmisano, Chiara Gigliarano, and Marco Bee for their precious advice. Financial support from Fondazione Compagnia di San Paolo, the University of Trento and the University of Milan is gratefully acknowledged.

†Post-doc Research Fellow at University of Milan.
Author e-mail: giovanna.scarchilli@unimi.it

1 Introduction

In the literature on poverty measurement there is a well-established concern in studying it from a multidimensional perspective (Alkire and Foster, 2011; Atkinson and Bourguignon, 1982; Bourguignon and Chakravarty, 2003; Deaton, 2016). The overall well-being status of a person is the outcome of a complex interaction among a number of life aspects, or functionings à là Sen (1985) and Nussbaum and Sen (1993), which show strong complementarities. When deprivations occur across a wide range of these life facets, the condition of poverty requires to be addressed taking into account each life dimension as well as the relation within these dimensions (Bossert et al., 2013; Chakravarty and D'Ambrosio, 2006).

The wide literature on the measurement of multidimensional poverty proposes several technical ways to aggregate the multiple dimensions of poverty into one univariate index, a composite indicator, which evaluates poverty as a unique welfare score and provides distributional analysis within and between countries, (see among others, Atkinson, 2003; Bourguignon and Chakravarty, 1999, 2019; Deutsch and Silber, 2005; Duclos et al., 2006; Espinoza-Delgado and Silber, 2021; Tsui, 2002). The composite indicator approach requires to take several arbitrary decisions on the attributes of the poverty index which impact on the overall definition of the poverty status (Decancq and Lugo, 2013). More specifically, these attributes are the dimensions' aggregation function, the order of units' aggregation (whether firstly across individuals or across dimensions), the weights assigned to each dimension, and other parameters defining the interaction between the dimensions.[1]

The exercise of aggregating information for each dimension with given inter-dimensional relations assumption might overlook to provide information on the underlying association among the dimensions *per se* (Duclos et al., 2006; Tkach and Gigliarano, 2020). In the realm of multidimensional poverty, it is important to examine how the different dimensions being considered are interconnected. This pertains to both defining and measuring poverty, wherein a significant stage involves determining whether the inadequate outcomes across various aspects of deprivation among individuals are occurring by chance or not. Thus, the interrelation between these dimensions should be taken into account when studying multidimensional deprivation (Decancq, 2020; Duclos et al., 2006). In technical terms, this implies to be able to observe a multidimensional correlation among the considered welfare dimensions and to estimate their joint distribution. Terzi and Moroni (2022) stress the convenience in investigating statistical concordance in multidimensional well-being to shed some light on the complex phenomenon of social vulnerability. In the literature the recourse to copula-based techniques to empirically measure correlation, or statistical dependence, between the multiple dimensions of life is not new (see among others, Aaberge et al., 2018; Decancq, 2014; García-Gómez et al., 2020; Pérez, 2015; Pérez and Prieto-Alaiz, 2016; Quinn, 2007; Tkach and Gigliarano, 2020). Aiming at extending this literature, this paper proposes an empirical cross-country evaluation of the dependence in the distribution of the various deprivations relying on the concept of statistical dependence. Among the relevant theoretical contributions (Decancq, 2020; Dolati and Úbeda-Flores, 2006; Taylor, 2016), Decancq (2020) proposes a more comprehensive framework to adapt the copula-based dependence analysis to the concept of cumulative deprivation. Cumulative deprivation is represented by the recurrent low outcomes in several life dimension for the same individual, in other words, it represents a condition of multidimensional exposition to poverty. In the broader context of multidimensional poverty measurement, cumulative deprivation performs a relative intersection identification approach. Indeed, the single-dimension scores are the ranks of the individual in the society, and the poverty identification takes place when the ranks in all the dimensions are below a certain rank threshold. Inspired to Atkinson (2003), when posing the question "How far do countries differ in the extent of multiple deprivation?", the proposal of this paper is to evaluate the dependence of the cumulative deprivation status and to compare various societies in terms of this feature. When comparing the various countries it remains important to address the study of poverty at the individual level so that the poverty indicator can be associated with certain individual characteristics and maintain the link with policy proposals and perspectives (Bossert et al., 2013). The relevance of this

[1] An alternative to composite indicators is the fuzzy indicator approach (Lemmi and Betti, 2006), which require to take decision about the probabilistic assignment rule to the poverty status given by the set of observed conditions.

analysis is discussed throughout the following simplified example. Let assume a researcher wants to study the welfare conditions of two different societies through the observation of two socioeconomic dimensions: per capita yearly income and life expectancy. Furthermore, let assume that within each society the observation is done on two representative individuals. The following tables describe the two societies in terms of each individual i each dimension achievement j and provide two different ways to aggregate the information.

Table 1: Comparing two hypothetical societal multidimensional distributions

(a) Society A

Individual	**Income**	**Health**
i_1	5000	82
i_2	80000	76
$\mathbf{\bar{x}_j}$	**42500**	**79**

(b) Society B

Individual	**Income**	**Health**
i_1	5000	76
i_2	80000	82
$\mathbf{\bar{x}_j}$	**42500**	**79**

(c) Composite indicator

Individual	**Income Index**	**Health Index**	$I(x_i^J)$
Society A	0.914	0.907	0.911
Society B	0.914	0.907	0.911

When the researcher adopts a dashboard approach, she ends up with the result presented in the last rows of tables 1a and 1b, which show a simple descriptive parameter for each dimension distribution, i.e. the sample mean ($\mathbf{\bar{x}_j}$). The conclusion drawn from this analysis would be that the two societies are the same in terms of average income and health, and not much information is available for deciding which one is more equally distributed in multidimensional terms. Alternatively, Table 1c shows an hypothetical result achievable following the HDI multidimensional composite indicator approach.[2] More precisely, the researcher is aggregating the dimension-specific scores obtaining a synthetic result which depends on the aggregation formula and the weights assigned to each dimension. What emerges from the composite indicator analysis is a similar picture with respect to the other approach: the two societies are similar in terms income and health aggregated conditions. Although the two approaches presented through this example depict fundamental information to describe the socioeconomic status of the population of a country, they miss an element to enrich the comparison between two societies.

Recalling Duclos et al. (2006), these two approaches are not able to capture the joint incidence of deprivation in multiple dimensions.[3] Certainly, the interaction among different welfare facets, aiming to delineate conditions of multidimensional poverty, has already been thoroughly investigated using a range of approaches. However, what continues to present a challenge in the analyses of multidimensional poverty is precisely understanding the magnitude of correlation among these distinct dimensions, and observing how this correlation undergoes fluctuations over time and across countries. Therefore, the aim of this paper is to fill the missing information regarding the degree of dependence among deprivations in certain welfare dimensions.

The utility of this study is threefold. First, a novel technique is applied to investigate the dependence between the various conditions of deprivation following the theoretical formulation of Decancq (2020), the downward diagonal dependence index. Deprivations' dependence is measured for several countries at different points in time, thus it is possible to perform dominance comparisons in dependence trends among countries and to formulate considerations linking the dependence with the economic cycle. Second, an analysis of the structure of the statistical dependence between the

[2] In this example I use two out of three dimensions of the HDI, the Income Index (II) and the Life Expectancy Index (LEI) following the formulas used with the Human Development Index. $II = \frac{ln(Y_{pc})-ln(100)}{ln(75000)-ln(100)}$, where the Y is the average income per capita in a given society. $LEI = \frac{LE-20}{85-20}$, where LE is the average life expectancy in a given society.

[3] Namely, neither of the two techniques individuates that going from society a to society b there is a correlation increasing "switch".

various dimensions is proposed, evaluating which are the driving forces of a condition of cumulative deprivation. Third, a descriptive analysis on the evolution of cumulative deprivation and the socio-demographic profiles of the deprived people is presented. To the best of my knowledge, this is the first study applying this technique to a cross-country perspective, since only Decancq (2022) presents an empirical application of this methodology for Belgium. The empirical application is performed on EU-SILC data for France, Italy, Spain, Germany, Czech Republic, Romania, Belgium, and Sweden between 2007 and 2019. The cumulatively deprived individuals are identified with the intersection criterion, the reference population is the working age population, the dimensions of life considered are personal income, work intensity, educational attainment, health status and housing quality. The dimensions are conceived as individual welfare domains based on the recommendations of Stiglitz et al. (2009). Moreover, these dimensions are among the main pillars of the UN Sustainable Development Goals and the EU 2020 strategy poverty target measure 'At Risk of Poverty or Social Exclusion' (AROPE).

The paper is structured as follows. Section 2 illustrates the copula methodology and the diagonal dependence index. Section 3 describes the data and dimensions. Section 4 displays the empirical application results and a discussion. In section 5, conclusive thoughts are presented.

2 Methodological framework

2.1 The Copula

The copula function is a particular multivariate distribution function with uniform univariate margins (Joe, 1997). It is also described as a multivariate function which aggregates all its marginal univariate components (Nelsen, 2007).

Let the individual well-being in a society be described by the d-dimensional random vector $X = (X_1, ..., X_d)$; and let $F(x_1, ..., x_d)$ be the joint cumulative distribution of well-being in a society.

Let $F_j(x_j)$ be the j^{th} marginal distribution of $F(x_1, ..., x_d)$, for $j = (1, ..., d)$. A non parametric approach to estimation is based on the count of the proportion of people in the society who have less than or exactly x_j in every j^{th} dimension of well-being. Equivalently, the proportion of individuals in the society who have strictly more than x_j in all d-dimensions is given by the survival function $\bar{F}(x_1, ..., x_d)$, where $\bar{F}(x_1, ..., x_d)$ is the multidimensional complement to one of $F(x_1, ..., x_d)$.

The individual who exactly has the amount x_j in dimension j, has a position p_j on the ordered outcomes across the population. The ranking series for each dimension j is a non-parametric estimate of the j^{th} marginal distribution $F_j(x_j)$. As follows from the *probability integral transform theorem*, the normalised rankings of each dimension, are uniformly distributed as $U(0, 1)$.

The formal definition of copulas can be derived from the Sklar's Theorem (Sklar, 1996, 1959).

According to the Sklar's Theorem, for any d-dimensional distribution function F_X with univariate margins $F_1, ..., F_d$, there exist a copula $C : [0, 1]^d \longrightarrow [0, 1]$ such that, for the set $\boldsymbol{x} = (x_1, ..., x_d) \in \mathbb{R}^d$

$$F_X(x_1, ..., x_d) = C_X(F_1(x_1), ..., F_d(x_d)). \tag{1}$$

And, if all F_j for $j = (1, ..., d)$ are continuous and strictly increasing, then C_X is uniquely defined in the unit hypercube $[0, 1]^d$.

Given that the inverse of a continuous and strictly increasing *cdf* F is $F^{-1} = F^{\leftarrow}$, the copula function of $F(x_1, ..., x_d)$ can be uniquely defined as follows:

$$C(p) = F(F_1^{\leftarrow}(p_1), ..., F_d^{\leftarrow}(p_d)) \tag{2}$$

and it is determined on the positional set $p = (Rank(F_1) \times ... \times Rank(F_d))$.

The Sklar's Theorem combines precisely the univariate marginal densities to form a d-dimensional joint distribution. This theorem depicts the reason of the use of copulas in statistical applications to study dependence between components of a random vector.[4]

[4]A simplified illustration of the Sklar's Theorem is presented by Hofert et al. (2019).

Given the random vector X and the position vector P representing the set of distributions of the ranked outcomes observed, being $F_j(X_j)$ the marginal distribution of the j^{th} dimension of $F(X)$, its copula function is a multivariate distribution C_X defined as:

$$C_X(p_1, ..., p_d) = Pr[F_1(X1) \leq p_1 \text{ and ... and } F_d(X_d) \leq p_d] \tag{3}$$

Where C_X expresses the proportion of individuals in the society who are outranked by the specific position set $p = p_1, ..., p_d$.[5] Equivalently, the survival function $\bar{C}_X(p_1, ..., p_d)$ represents the proportion of individuals who are outranking the same position set.

Intuitively, in case of absence of any type of dependence among the marginals, the copula function would simply be the product of the d-margins.[6] For a random vector $\boldsymbol{P} = (P_1, ..., P_d)$ with $P_1, ..., P_d \sim^{ind} U(0, 1)$, the *independence copula* is

$$\Pi(\boldsymbol{p}) = \prod_{j=1}^{d} p_j, \;\; \boldsymbol{p} \in [0, 1]^d \tag{4}$$

There are other two types of copulas that need to be mentioned because they represent two extreme cases in terms of dependence. They are known as the Fréchet-Hoeffding bounds (F-H), and they represent the lower and upper bound of every copula. They are respectively $W(\boldsymbol{p}) = max\{\sum_{j=1}^{d} p_j - d + 1, 0\}$ and $M(\boldsymbol{p}) = min_{1 \leq j \leq d}\{p_j\}$, for $\boldsymbol{p} \in [0, 1]^d$. For any given d-dimensional copula C, the theorem of Hoeffding (1940) and Fréchet(1951) states that any copula C is pointwise bounded from below by a lower bound W, and from above by an upper bound M. We could interpret the lower bound as representing complete counter-monotonicity among the dimensions, and the upper bound as the complete co-monotonicity among the dimensions.

2.2 The diagonal dependence and the copula sections

The copula function is useful to investigate the dependence or association between random variables. This can be done taking into consideration two of the extreme cases of dependence introduced in the previous section. Namely, the co-monotonic case, which implies maximal dependence, and the independence case in which there is no dependence at all, i.e., a fully random interaction between the copula dimensions. The dependence can be investigated as a global aspect characterising the whole distribution, or as a phenomenon that varies across the distribution. In order to evaluate the multidimensional dependence among deprived individuals it is necessary to focus on a specific quadrant of the joint distribution, namely its left tail. The left tail dependence in the copula framework, hands over relevant insights for quantifying the phenomenon of cumulative deprivation in well-being.

To provide an intuitive geometric interpretation of tail dependence, it is necessary to refer to the sections of a bi-dimensional copula function. The sections of a bi-dimensional copula are three: horizontal, vertical and diagonal (Hofert et al., 2019). In mathematical terms, the sections of a bi-dimensional copula, $C(u_1, u_2)$, are two-dimensional planes that slice the surface of the copula density function and fall perpendicularly on the (u_1, u_2) plane. While the vertical and the horizontal sections are planes that slice the copula density at a fixed point $a \in [0, 1]$ in one of the two dimensions, the diagonal section is the function defined as $\delta_C(p) = C(u_1 \leq p, u_2 \leq p)$ where p can be any value in the $[0, 1]$ interval. For the issue of this study it is only necessary to observe the diagonal section of the copula. The diagonal section show us the density of the copula function when all the dimension-specific ranks are lower or equal than p.

Decancq (2020) proposes a simple and intuitive way to observe the diagonal section of the copula function with the construction of the *diagonal dependence diagram*. In the diagonal dependence diagram there are two curves, the downward diagonal dependence curve representing the diagonal

[5]Recalling one of the definitions of the copula function provided by Nelsen (2007) (page 1), "*[..] copulas are functions that join or "couple" multivariate distribution functions to their one-dimensional marginal distribution functions.*"

[6]This case represents the simplest copula, i.e., the independence copula.

section of the copula function and the upward diagonal dependence curve representing the diagonal section of the survival function, i.e., the multidimensional complement to one of the copula. They are respectively showing the population share outranked and the share outranking a specific positional combination between 0 and 1. This paper focus on the downward diagonal dependence curve.

Given a d-dimensional random vector $\boldsymbol{X}$ with copula function C_X, the downward diagonal dependence curve $D_X(p)$ is defined as the proportion of people who have a lower or equal position than p in all d dimensions:

$$D_X(p) = C_X(p, ..., p); \quad \text{for all } p \in [0, 1] \tag{5}$$

A graphical representation of the downward diagonal dependence curve is provided by Decancq (2020). The copula diagonal section is represented on a two-dimensional plot having on its x-axis the set of rank positions for all the d dimensions, $\boldsymbol{p} = (p_1, ..., p_d)$, combined such that all the rank positions are the same ($p_1 = p_2 = ... = p_d$). The y-axis represents the proportion of population that is outranked by each position combination in the set. In other words, the y-axis of the downward diagonal dependence curve coincides with the copula density at a specific multidimensional point. Figure 11 in the Appendix is an example of how a downward diagonal dependence curve of a d-dimensional copula function looks like. The downward diagonal dependence curve is meant to compare the diagonal section of the empirical copula estimated on the multiple dimensions of well-being, with the case of a co-monotonic copula. The co-monotonic copula describes a "feudal or cast society" [7] as being poor in one dimension automatically implies being poor in all the other dimensions. Given that the highest density of a two-dimensional co-monotonic copula function lies exactly on the points in which individual positions on each dimension are the same, its diagonal section coincides with the 45° line.

The dependence comparisons can be done as well between copulas. When considering the global dependence of two different multidimensional copulas, one of the requirements to be fulfilled could be similar to a first-order stochastic dominance ordering (Decancq, 2014). However, in order to assess whether a $d-$dimensional copula $C_X(p_1, ..., p_d)$ is more diagonally dependent than $C_Y(p_1, ..., p_d)$ only for those positional vectors in which ($p_1 = p_2 = ... = p_d$), the requirement is weaker and regards only the copula diagonal section. As defined by Decancq (2020), $D_X(p)$ is more diagonally dependent than $D_Y(p)$ if

$$D_X(p) \geq D_Y(p); \quad \text{for all } p \in [0, 1] \tag{6}$$

With the diagonal dependence orderings it is possible to operate a pair-wise comparison of two different copula sections with respect to their proximity to the co-monotonic case.

In order to measure the level of diagonal dependence among the components of a copula C_X, Decancq (2020) proposes to calculate the area underlying the Downward Diagonal Dependence Curve and derive the index measuring the downward dependence (Eq. 7).

$$\boldsymbol{\delta}_d^-(X) = \frac{2(d+1)\int_I D_X(p)dp - 2}{d-1} \tag{7}$$

The DDDI represents the degree of concentration of population being deprived in all dimensions.

In the results that will follow, I empirically compute the DDDI for each country and year to measure how the dependence among deprivations has evolved. In order to complete the diagonal dependence picture I present the downward diagonal dependence diagrams for each year and country; and I provide a cross-country and time diagonal dependence partial ordering.

2.3 Empirical copula section computation

Let q be the number of quantiles and d the number of dimensions of the copula function. When working with empirical copulas the density function is approximated by calculating the proportion of society which is falling in each of the $N = q^d$ groups of dimension-specific quantile combinations.

[7] As stated by Decancq (2020)

The number of positional groups quantifies the grid size belonging to the $\boldsymbol{I}^d = [0,1]^d$ set. In other words, N represents all the possible combinations of the given quantiles among d dimensions.

Intuitively, if there was no dependence among $d = 5$ random variables, the probabilities for each of the N combinations of the quantiles would always be the same. More precisely, if $q = 3$, the probability of the independence case is 0.004 for every positional group.[8] If $q = 5$, the probability for the independence case of every single positional group is 0.00032, i.e. $= 1/3125$. For q being equal to 10 or 100, the independence case will assign respectively a probability of 1^{-5} and e^{-10} to each positional group. The overall empirical copula density estimation requires to have a rule for ordering the various positional combinations. This task is not necessary when working with the diagonal section of the copula function. More simply, the diagonal section of the empirical copula density function, for every country and year, is derived through the following steps:[9]

1. compute the ranking series of the observed sample for each of the five dimensions (ties randomly sorted, ranking process replicated 100 times)
2. normalise the ranking series on a $(0,1)$ scale dividing each ranking position by $(n+1)$, where n is the sample size
3. for every individual, take the highest normalised ranking score among the five scores associated to the various dimensions
4. the diagonal section of the copula is evaluated at $p = (0.01, 0.02, ..., 0.99, 1)$ ranking positions by counting the proportion of population which is outranked by each rank

3 Data source and preparation

The empirical application uses EU-SILC cross-sectional data for selected European countries from 2007 to 2019. The countries are Belgium, France, Germany, Italy, Spain, Czech republic, Romania and Sweden. Due to the lack of availability of full information for every country in every year, 2007 is missing for Belgium, Sweden and Germany and 2008 is missing for Belgium and Germany. Furthermore, throughout all the years, Germany lacks information regarding housing conditions. As a result, our analysis for Germany is conducted using only four dimensions. The sample used for the study is composed by people in the working age, from 25 to 65 years old.

The following five welfare dimensions have been selected and constructed to analyse deprivation: income, work intensity, educational attainment, general health status, housing conditions.[10] For all these dimensions the relevant information is each individual's rank position with respect to the others within the single dimensional ordering. It is well-known in the copula-based literature on multidimensional well-being and poverty studies, that many dimension's ranking suffer for a very high amount of ties. Generally, the ties are randomly ranked in all the dimensions. In this study, in order to reduce the dependence on the single random ranking of the tied observations, the final ranking series are derived by replicating all the random rankings a hundred times and weighting the computation results at each iteration. At each replication all dimension-specific ties are randomly ranked and a DDDI is computed.[11] Each dimension and its specific ranking choices are illustrated in the rest of this section.

Housing conditions is observed at the household level, therefore, the ties are not only appearing between households but also within the household. The housing condition dimension is an counting

[8] for $N = 3^5 = 243$ combinations, the independence case requires the probability of each combination to be $P = 1/243$

[9] Steps one and two are equivalent to the computation of the empirical cumulative distribution function within each dimension.

[10] The exclusion of younger than 25 years old individuals is motivated by the necessity of avoiding the count of those who are still in education as people deprived in education.

[11] The presented DDDI is therefore bootstrapped on a hundred random rankings. Since the confidence intervals are all very small and vary only at the fourth decimal of the resulting DDDI value, the results are plotted without showing the CI for each index value. The appendix contains the SE values for each year and country's DDDI averaged on 100 random rankings of the ties.

variable summing all the deprivations related with dwelling-related quantitative and qualitative features. In the housing quality indicator are included the following aspects: the presence of damp walls, leaking roof or rot in window frames or floors, ability to keep adequately warm, bath or shower in dwelling, indoor flushing toilet for sole use of the household, possession of colour TV, computer, washing machine, capacity to replace worn-out furniture, problems with dwelling: too dark, not enough light, position quality: noise from neighbours or from street, health quality of the location of dwelling: pollution, or other environment problems, social conditions in the geographic location (crime, violence, or vandalism in the area), overcrowded household rate. All the counted deprivations are assigned the same weight and are simply summed up in the final counting variable.

The income dimension is gathering both information on equivalent household disposable income and individual income. This choice is based on the necessity to perform an individual-based analysis which takes into account for both for the household income and the inequalities which may take place within the household. Moreover gender disparities in income are better visible when considering individual income and not household equivalent disposable income. In practice, at a first stage individuals are ranked using equivalent household income. At a second stage individuals are ranked within the same household using individual labour market incomes and the remaining ties are randomly ranked.

The education dimension is derived relying on a double level of ranking, firstly on the ISCED levels, secondly on the years of schooling.

The job condition dimension is ranks individuals with respect to work intensity index scores, which are provided by the Eurostat and included in the EU-SILC data. The work intensity is defined as the ratio of the total number of months that a working-age person have worked during the income reference year and the total number of months the person theoretically could have worked in the same period. This measure embeds a conversion of part-time work into full-time equivalent in order not to assign low scores to part-time but to individuate the cases in which there is fragmentation in working path of the considered year.

The health dimension is derived from the estimated latent general health status extrapolated from the Self-Assessed Health status distribution conditional on some personal characteristics and health-related behaviours available in the data set. Due to the absence of objective health measures in survey data, the use SAH in social studies is very common.[12] A pitfall of such a measure is that it provides a very low variability among the respondents and difficulty to observe it from some distributional perspective. Many empirical studies adopted non-linear regression techniques to translate the subjective health categories into a cardinal measure representing the estimated underlying latent health status taking a continuous form.[13] The health dimension in this study represents a latent general health status estimated using data on the self-assessed general health (a factor which takes five levels indicating increasing health conditions) from the EU-SILC cross-sectional individual data set. The estimation takes into account other health habits/features, and personal characteristics (presence of any chronic disease, limitations in everyday activity, age). Calling $\boldsymbol{X'}$ the vector of the regressors presented, the estimated ordered logit model of latent health status is: $h^* = \boldsymbol{X'}\beta + \epsilon$. Then, the predicted latent general health status for each individual is $h_i^* = \boldsymbol{X_i'}\beta$, also named *z-score.*

The health, education and income measures are correlated with age. The achievements in these dimensions should represent a deprivation in relation to the people of the same age group. Therefore, before deriving the ranking series, these dimensions have been standardised with respect to the age-group means and standard deviations.[14] Table 2 provides a description and a synthetic illustration of the empirical copula dimensions.

[12] Idler and Benyamini (1997) have demonstrated that the SAH could be a good predictor of other health measures such as life expectancy and use of healthcare.

[13] These approaches have been validated by the contribution of Van Doorslaer and Jones (2003).

[14] The age-group intervals are: 24-35, 36-45, 46-55, 56-65.

Table 2: Description and methods of construction of the selected well-being dimensions

Dimension	Data used	Method to derive the outcome variable
Individual income	Equivalent household disposable income, Individual net income	*Unit:* Individual *Method:* Two levels of ranking: 1) ranking of people with respect to equivalent household disposable income 2) ranking with respect to individual income *Outcome variable:* Income dimension
Health status	General health status, presence of chronic illness, limitations in everyday activities, age, gender	*Unit:* Individual *Method:* Ordered logit estimation *Outcome variable: z-score*/latent general health
Housing quality	Overcrowing rate, material conditions of dwelling, house location, social conditions	*Unit:* Household *Method:* Counting approach summing all the dwelling-related no-deprivations *Outcome variable:* Index of quality of housing
Educational attainment	ISCED level, years of schooling	*Unit:* Individual *Method:* Two levels of ranking: 1) ranking of people with respect to ISCED level 2) ranking with respect to the years of schooling *Outcome variable:* Educational attainment by time
Working condition	work intensity indicator	*Unit:* Individual *Outcome variable:* Working intensity, EU-SILC built-in indicator

For describing cumulative deprivation incidence across the population, I have selected the portion of population falling in the bottom 40% in all dimensions (i.e., those whose maximum position among the one-dimensional rankings is found to be ≥ 0.4).

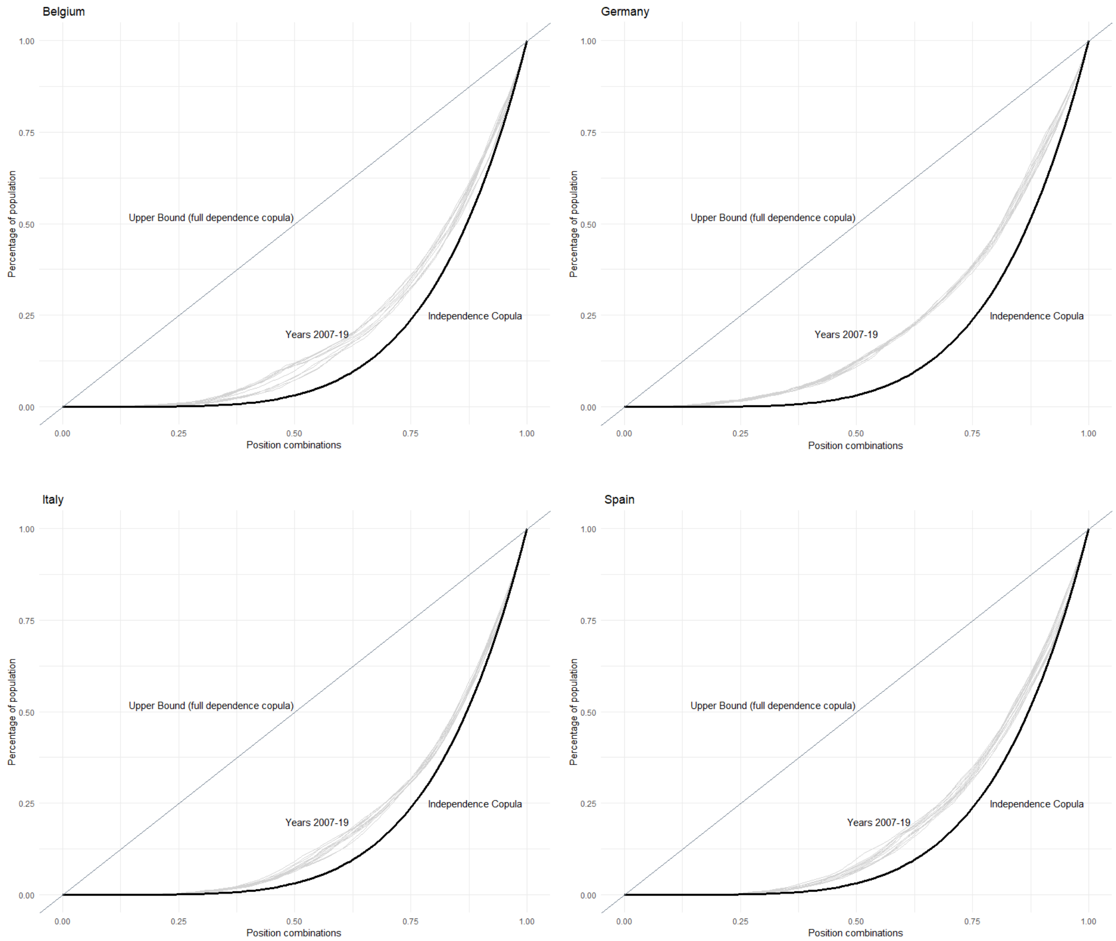
4 Empirical application and results

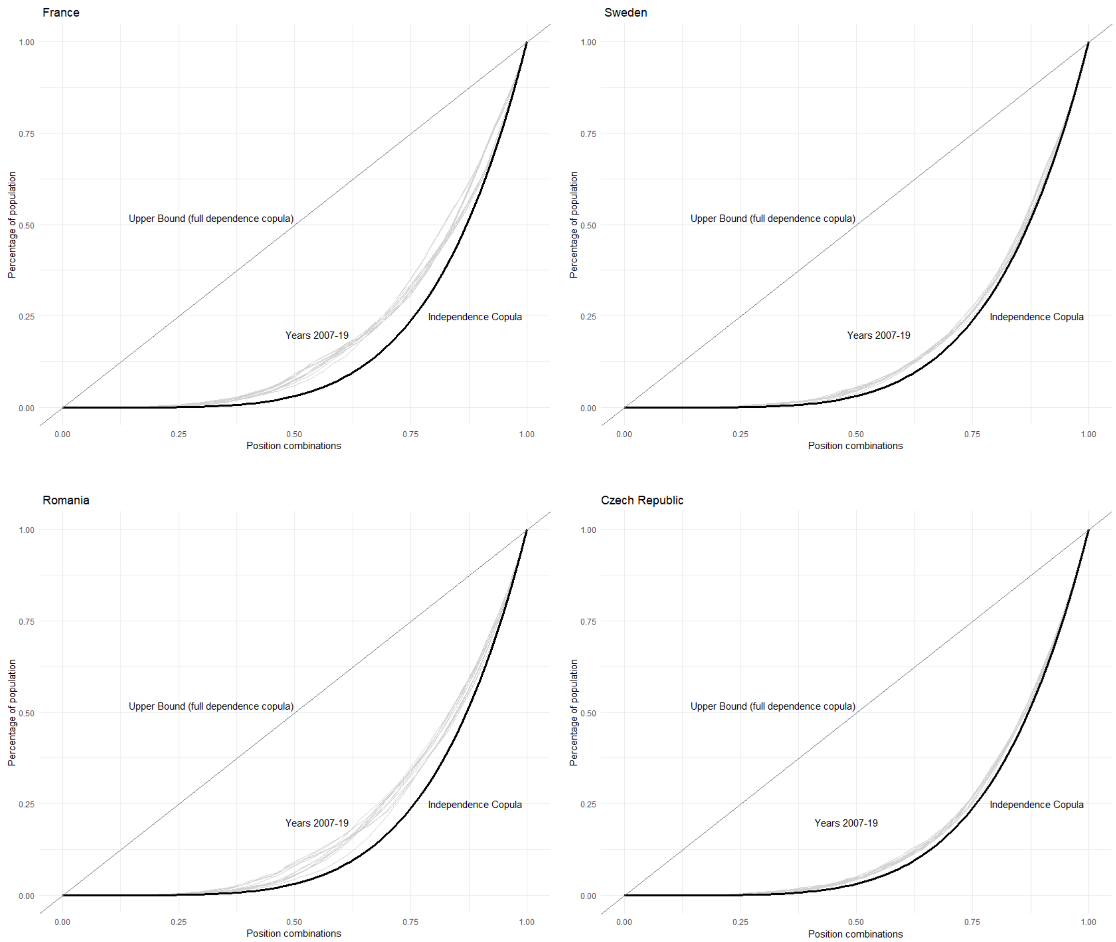
4.1 Multidimensional diagonal dependence across the various dimensions of poverty

Figures 2a-b-c-d and 3a-b-c-d show the downward diagonal dependence curves for the single country and year copula sections. Each country-specific plot shows the densities of the yearly multidimensional copula sections observed when the individual ranking positions across all the dimensions are below a certain percentage value indicated in the x-axis; i.e., the diagonal sections. With this diagram it is possible to represent how far is the multidimensional dependence structure for each country from the two extreme cases of total independence and full dependence. In each plot are shown, in light-grey, the country-specific copula diagonal sections computed at each year, the independence and full dependence copula sections, respectively the black curve and the 45-degrees line.

Since the yearly curves intersect with each other at different density values for each country, it is impossible to make dominance comparisons on dependence levels through this representation. However, even if it is not possible to understand whether one year dominates the others, these plots provide a descriptive and synthetic tool to compare cumulative multidimensional distributions with the two full dependence or independence cases.

The different patterns emerging from the plots may guide the interpretation in two directions. On the one hand, an heterogeneous dispersion in the yearly dependence curves emerges for each country implying a different cross-country relation between the degree of association among dimensions of poverty and the economic cycles. On the other hand, the average dependence is also heterogeneous across the various countries. This phenomenon may be driven by the differences among the various economic contexts and welfare systems characterising the linkages between the dimensions of poverty in a specific country. The strong interconnection across income, health, education, work condition and housing is not distributed in the same way across the selected countries. Similar dependence structures are observed among the continental and southern European countries, Belgium, Germany,


Belgium
Germany
Italy
Spain
Upper Bound (full dependence copula)
Years 2007-19
Independence Copula
Percentage of population
Position combinations
0.00
0.25
0.50
0.75
1.00



France, Italy and Spain. Romania and Czech republic, the eastern European countries, instead behave quite differently. Finally, Sweden shows some similarities with Czech republic.

While the cross-country variation could be interpreted as institutional differences in the interplay among the considered spheres of life, the time variation within a certain country may be attributed to policies, reforms and economic cycles. Hence, it is possible to contextualise the time variation in dependence as a comparison between the years of austerity policies post-financial crisis (2009-2012), with the years immediately preceding. These considerations are purely descriptive and should not be interpreted as a causal effect of different welfare systems or economic cycles on the dependence patterns.

4.2 The trend in diagonal dependence across the various countries

Figure 3 presents the evolution of the downward diagonal dependence index for each country in each year. Table B.1 in the Appendix shows the bootstrapped standard errors. With this figure, it is possible to notice two country clusters in terms of average DDDI values, Sweden and Czech Republic belong to the group with a DDDI ranging below 5% for most of the time. The other countries more than double this value and show more increasing trends over time. The two countries showing highest degree of diagonal dependence in 2019 are Belgium and France. The former presents high values from the 2012 onward, whilst the latter did a rapid catching up within the last years. Apart from the within-period oscillations, almost all countries present a positive variation between the initial and the last available year. Despite the heterogeneous patterns shown by the various countries, we see a common characteristic among all of them. The steeper increase of every country's DDDI takes place in the aftermath of the sovereign debt crisis, during a period which have seen a deep recession and a generalised implementation of austerity policies across the EU. In these times, despite the different reactions of each country, public policies across all EU were adopting austerity measures and public expenditure cuts. It is therefore plausible that the economic conjuncture impacted in determining the rise in the degree of dependence among income poverty and deprivation on the side of health, education, housing and work intensity.

Figure 3: Downward Diagonal Dependence Index for selected countries and years

4.3 Diagonal dependence dominance comparisons

The downward diagonal dependence index represents the average distance of a certain distribution from the extreme cases of diagonal independence or dependence, yet, it does not provide a strict dominance ordering across different distributions. In order to assess whether a country dominates another country in a specific year with respect to the diagonal multidimensional dependence, it is necessary to compare the multidimensional copula and survival density of each country and year at all points of the copula diagonal section (Decancq, 2020). This analysis is equivalent to that done in the standard distributional approaches which relies on stochastic dominance criteria to compare couples of Lorenz or concentration curves (Levy and Robinson, 2006; Yitzhaki and Schechtman, 2013).

Intuitively, we see whether the copula diagonal sections for each country and year are falling above another copula section and never cross each other in all their points. This exercise checks whether any proportion of the population in a given society experiencing a certain combination of positional rankings in each of the considered dimensions is always greater or equal to the proportion observed for another society. Table 3 shows all pair-wise comparisons across diagonal copula sections for all countries and years (between 2009 and 2019, given that the 2007 and 2008 are missing for some countries). The test is performed on a finite grid of positions 10^5. Therefore, the dominance is assessed over 5 points of the copula function. The comparison is performed between pairs of countries in a specific year. The country in the column is dominating the country in the row in the indicated year. Zeroes are placed when the outcome of the comparison is indecisive, hence there is at least one position out of the grid which does not respect the partial dominance requirement.[15] Given that the comparison takes place across pairs of countries in a specific year, we can have that country A is dominated by country B in year x while country A dominates country B in a different year y, therefore the table is not necessarily symmetric. As it emerges from the table, the more dependent countries in terms of poverty dimensions are Germany, Belgium and France, while Eastern European countries are characterised by weaker dependence across the considered life facets when compared singularly with all the other countries. The years in which we find more countries strictly dominating the others are 2019, 2015 and 2011.

Table 3: Pair-wise dominance comparisons

	ES	**IT**	**FR**	**DE**	**BE**	**SE**	**CZ**	**RO**
ES	-	0	2018	2009 - 2016, 2019	2019	0	0	0
IT	0	-	0	2009- 2015, 2019	2019	0	0	0
FR	0	0	-	2009 - 2016	2016, 2017	0	0	0
DE	0	0	0	-	0	0	0	0
BE	0	0	0	2009 - 2015, 2017	-	0	0	0
SE	0	0	2019	2009 - 2019	2013, 2019	-	0	0
CZ	2012, 2014, 2015, 2019	2019	2011 - 2015, 2017 - 2019	2009 - 2019	2012, 2014 - 2019	2014 - 2018	-	2012, 2015
RO	2011	2011, 2019	2011, 2019	2009 - 2016, 2019	2015, 2016, 2019	0	0	-

The evaluation of the dependence within cumulative deprivation dimensions and dependence characterisation in relative terms, with the partial dominance orderings, are two ways to compare countries and years. However, to conduct a more comprehensive analysis of the status of multidimensional poverty, it is also important to assess the single-dimension contribution in defining the levels of interdependence. In the next paragraph, the analysis is focused on assessing exactly the role of each dimension in shaping the overall diagonal dependence.

4.4 What is the contribution of each dimension to the diagonal dependence?

In order to disentangle the contribution to the overall multidimensional dependence from each single dimension, the dependence is evaluated removing one dimension at a time, and the empirical copula density is computed for the $d-1$ dimensions case. For all the countries except of Germany, the initial amount of dimensions is $d = 5$, while for Germany is $d = 4$. Figure 4 presents the puzzle of the single dimension removal[16].

From figure 4 it appears that the removal of a single dimension at a time does cause some variations in the downward diagonal dependence ranking between countries and in the trends.

[15] Note: single years are separated by comma, intervals of years are indicated with dash, i.e., 2012-2014 includes 2012, 2013 and 2014.

[16] Table B.2, in the Appendix, shows the t-test of mean differences between the d dimensions case and the $d-1$ dimensions case. Overall, these variations are statistically significant at the 95% of confidence level considering the single country paths across the years.

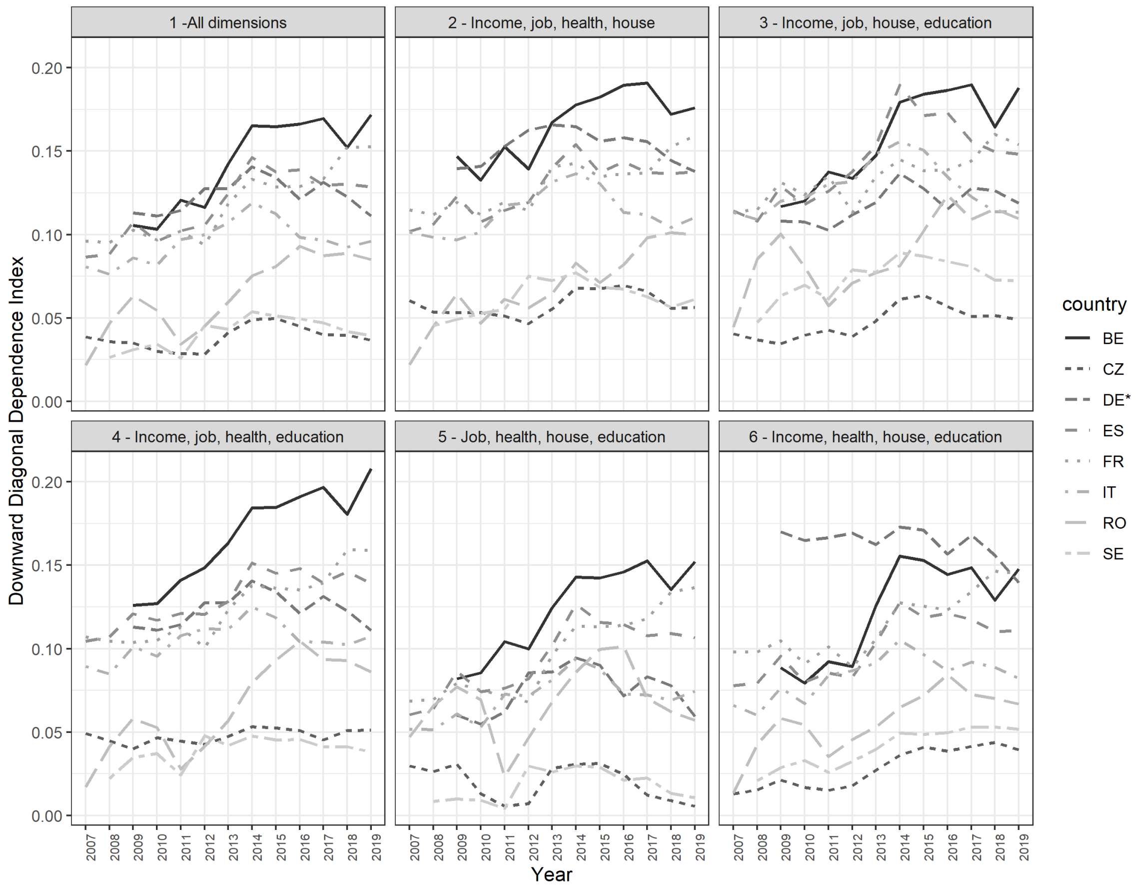


Figure 4: Downward Diagonal Dependence removing one dimension at a time

Starting from the second panel at the top, the downward diagonal dependence undergoes a slight generalised growth from the removal of the education dimension. A similar trend is observed with the removal of the health dimension. A different behaviour is represented by the removal of the housing conditions dimension, which leaves almost unchanged the series of all countries except Belgium. The most interesting results are undoubtedly provided by the last two panels, box 5 and 6 of the same figure. From box 5, it is noted that the removal of the income dimension significantly reduces the value of dependence, also changing the trend on some occasions. This is a sign that, although multidimensional poverty must be understood considering many aspects, income remains one of its main determinants and therefore, when this is not considered, the general interdependence between the various facets of well-being decreases. Finally, the exclusion of the work intensity dimension leads to an greater variation in the ranking between countries. Germany appears to have the highest multidimensional dependence score when not considering the working intensity, a possible explanation is due to the high levels of German employment rate compared to other European countries. The heterogeneity in the dependence scores across the different dimension selection could provide valid information for the definition of a country-specific poverty status.

4.5 Deprivation conditions beyond diagonal dependence

Cumulative deprivation, as conceived in this paper, is experienced by people when their maximal ranking score among all the considered dimensions is below the bottom 40%. The 40% serves as a threshold for counting people experiencing cumulative deprivation. The descriptive statistics and quantile distribution of the maximal ranking score in 2019 is provided in the Appendix, with Table

B.3. It emerges that the maximal position is in average really high and that the cumulatively deprived individuals fall roughly within the bottom 5% of the maximal positions distribution.

A more detailed picture on the evolution of cumulative deprivation is provided by Figure 5. The figure panel on the left shows cumulative deprivation incidence across countries considering all dimensions. The panel on the right side is instead the incidence of a subset of dimensions excluding the housing component in order to have a correct comparison with Germany.

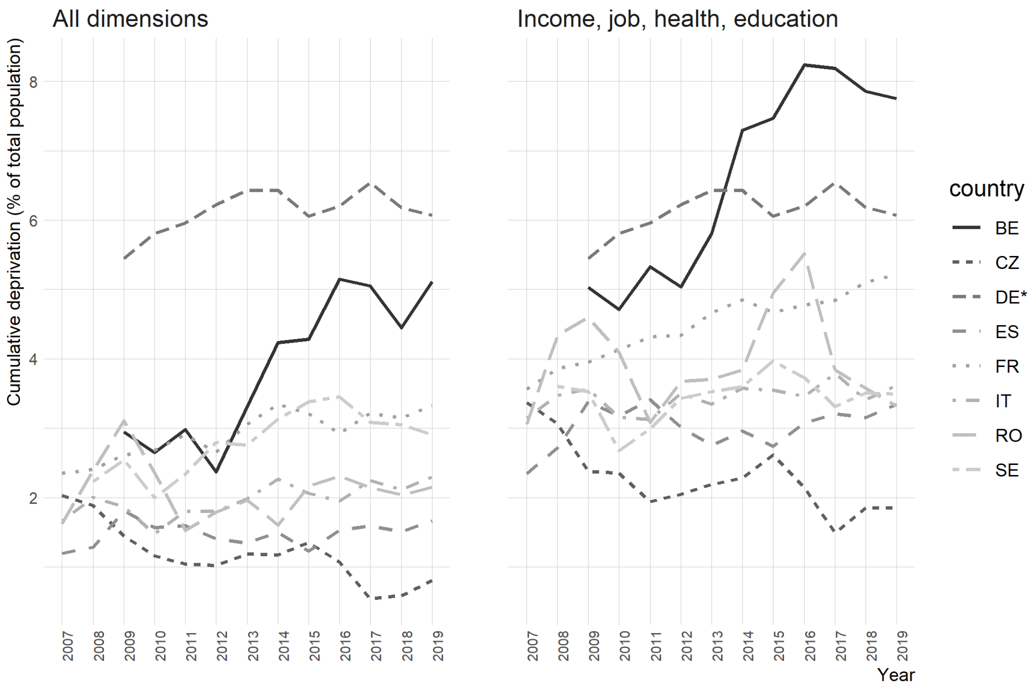


Figure 5: Cumulative deprivation incidence: proportion of total population which is falling in the bottom 40% in all the indicated dimensions of poverty.

The percentage of population hit by cumulative deprivation is derived relying on the EU-SILC cross-sectional weights. The countries showing the highest incidence in cumulative deprivation are Belgium and Germany, reaching approximately the 5.5% of total population in recent times. As shown in the figure, the incidence of cumulative deprivation is higher for those countries for which the downward diagonal dependence index resulted high. However, the relation between the DDDI and cumulative deprivation incidence is not linear and the cross-country comparison shows a slightly different clustering.

4.6 The single-dimension achievements of cumulatively deprived people

Cumulative deprivation adopts a relative identification technique, it does not identify those experiencing poverty as people whose achievements fall below a given threshold. The following paragraph provides insights on the levels achieved at each dimension, and allows to depart from the fully relative perspective to describe the conditions of those cumulatively deprived. Figures 6, 7, 8, 9a and 9b, show the trends of the actual values achieved by people falling in the bottom 40% for all dimensions. Figure 6 compares the average yearly disposable income in a country with the average computed within the cumulatively deprived people, in this way we can summarise how distant in terms of the income dimension are the cumulatively poor with respect to the overall population.

The gap between the poor and the average population is wider in northern European countries, signalling that the condition of multidimensional exclusion is experienced by people which are far below the mean values. The trends are also very heterogeneous, showing a notable gap widening in Belgium, France Germany and Spain, implying a worsening of the income conditions with respect to the average income trends in the country.

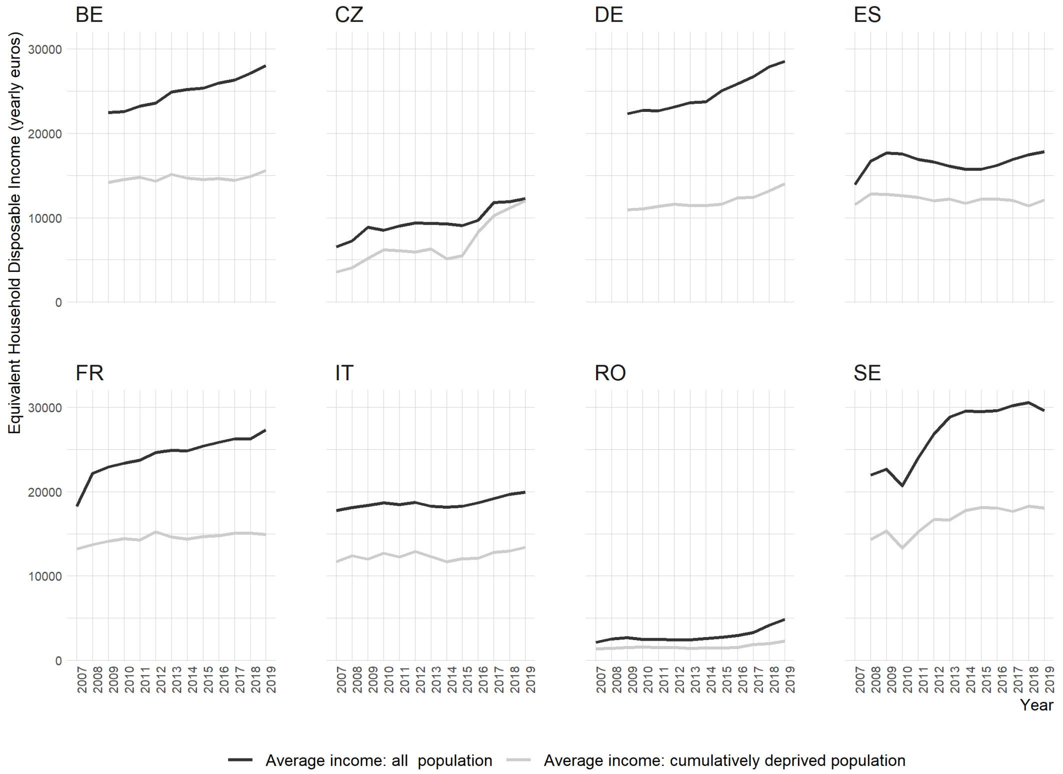


Figure 6: Average equivalent household disposable income of the multidimensionally deprived

On the side of work intensity (Figure 7), the gap results generally very high, meaning that cumulatively deprived people are less present in the labour markets for all the considered countries. Low work intensity is generally associated with discontinuous employment careers, intermittent work, part-time jobs. Hence, by definition, low work intensity is related to more unstable jobs with respect to labour market shocks. Looking at the trends across time, there is a fairly uniform decrease in work intensity among cumulatively poor, indicating that multidimensional poverty is being increasingly associated with a high level of labour deprivation. Moreover, the oscillations in work intensity are wider among the cumulatively deprived, meaning that, the cumulatively deprived population is more intensely subject to labour market shocks such as the financial crisis and the sovereign debt crisis.

The housing deprivation (Figure 8) is particularly affecting cumulatively deprived people in Romania, France and Belgium. The countries being compared all have very high enrolment rates. All of these countries have compulsory and state-subsidised education systems; therefore, educational deprivation can affect individuals who still complete the compulsory school cycle.

Although the definition of the educational dimension formulated for multidimensional deprivation includes years of education, the representation in Figure 9a refers only to the highest level of education attained. More precisely, Figure 9a shows which are the various year-specific median educational attainment for the cumulatively deprived people of a country. It emerges that in in

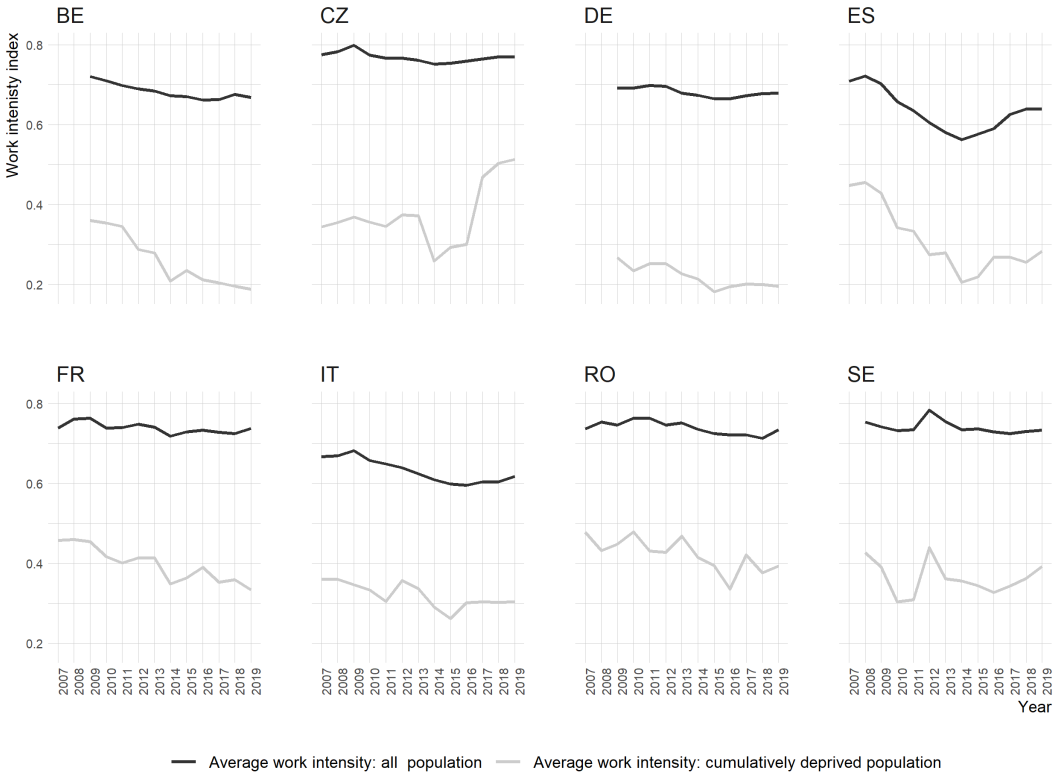


Figure 7: Average work intensity of the multidimensionally deprived

many countries people with upper secondary school educational attainment fall within cumulative deprivation. An exception is Spain, where poor people have mainly lower secondary education. With respect to self-assessed health (Figure 9b)[17], most of deprived people declare to be in fair or good general health status.

[17] Which is the original variable used to estimate the latent health condition.

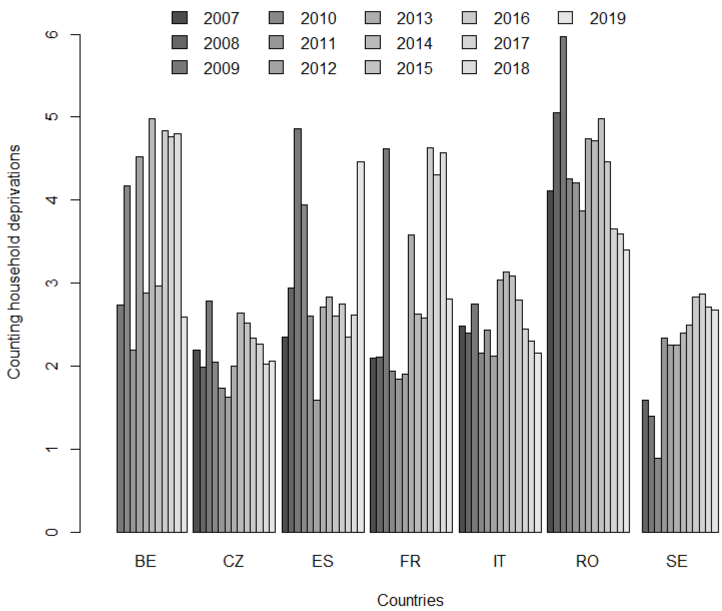
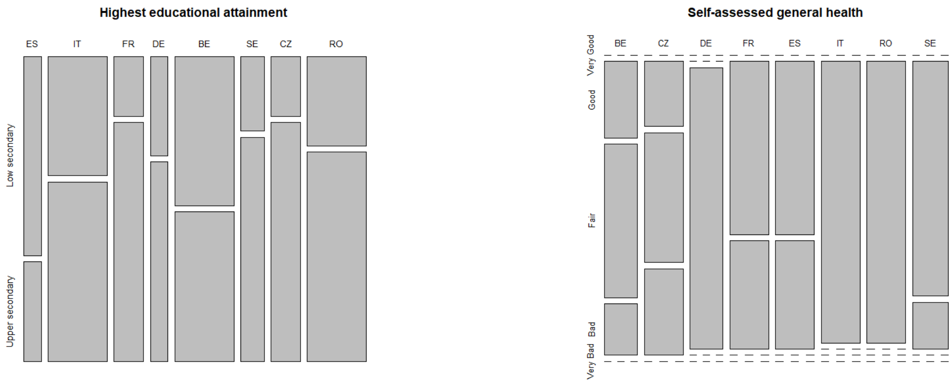


Figure 8: Average housing-related deprivations of the multidimensionally deprived person



(a) Median educational attainment of the multidimensionally deprived person

(b) Median self-assessed general health of the multidimensionally deprived person

Figure 9

4.7 Who are the cumulatively deprived people?

Cumulatively deprivation's socio-demographic characteristics are examined in this section. The analysis is developed with a descriptive non-linear *probit* regression using cumulative deprivation as the dependent variable and a set of socio-demographic characteristics as the explanatory variables. Recalling Table B.3, the maximal position of 0.4, which is the threshold here chosen to identify the cumulative deprivation status, falls in average within the bottom 5% of the maximal positions distribution for all countries. At this stage the probability of being falling in this small set of society, that is being ranked less than the 40^{th} percentile across all the five dimensions rankings, is evaluated conditional to given socio-demographic characteristics. The dependent variable choice is motivated by the fact that the use of the maximal ranking position is not able to describe the deprivation status as far as the score rises. Hence, a high maximal position do not necessarily describes total absence of deprivation since it indicates that the individual is not deprived in at least in one dimension (there could be deprivation in 1 out of 5 dimensions, or even in 4 dimensions out of 5). The explanatory variables are: gender, age, quadratic age, living in an area whose average income is below the mean national income, citizenship (whether local, EU or Extra-EU), number of children in household, activity status (employee, self-employed, unemployed, inactive)[18], tenancy status (owner, owner with mortgage, tenant at full market price, tenant at reduced price, tenant for free). The estimated model aims at describing the socio-demographic profile of cumulatively deprived individuals, the regression outputs are presented for each country and progressively add up regressors in order to grasp the evolution of single correlations.

Table 4: Population social and economic characteristics by cumulative deprivation status - 2019

Country	ES		IT		FR		DE		BE		SE		CZ		RO	
Cumulative Deprivation	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Gender																
Male	47.78	33.76	48.54	23.67	47.81	35.24	45.87	42.28	48.72	37.82	49.51	38.76	50.09	33.71	50.76	31.64
Female	52.22	66.24	51.46	76.33	52.19	64.76	54.13	57.72	51.28	62.18	50.49	61.24	49.91	66.29	49.24	68.36
Age class																
(23,35]	21.12	10.82	21.28	25.57	24.66	28.15	20.09	6.64	29.67	18.65	25.30	20.16	25.22	0.00	22.82	16.95
(35,45]	30.16	37.64	28.01	27.84	29.16	28.60	23.02	17.32	26.99	26.42	28.64	29.46	31.54	14.86	31.99	41.81
(45,55]	31.95	37.16	32.67	27.08	33.14	33.87	36.40	42.71	28.80	33.42	29.26	32.56	27.29	50.86	30.31	41.24
(55,67]	16.77	14.38	18.05	19.51	13.03	9.38	20.49	33.33	14.54	21.50	16.80	17.83	15.96	34.29	14.88	0.00
Activity Status																
Empl	62.09	10.98	60.21	14.18	75.19	29.91	81.27	35.36	70.03	13.80	80.93	49.61	74.69	21.84	63.98	31.02
Inactive	15.27	44.09	15.01	56.73	10.46	34.80	10.75	42.32	16.26	65.10	7.55	33.34	1.26	58.62	18.72	52.87
Self_empl	11.50	3.77	18.03	3.07	9.48	3.04	5.31	1.45	9.34	1.82	9.81	6.20	12.16	7.47	16.28	10.92
Unempl	11.74	41.15	6.87	25.67	6.07	33.18	2.68	20.87	4.82	19.27	2.12	12.40	1.89	12.07	1.02	5.17
Citizenship																
EU	2.49	2.42	3.65	3.17	2.20	2.52	NA	NA	9.13	12.18	2.88	2.33	1.59	0.57	0.03	0.00
Local	90.75	81.58	90.40	91.42	92.79	87.19	95.96	95.41	84.93	75.91	90.26	89.15	97.42	98.29	99.96	100.00
Extra-EU	6.16	15.99	5.93	5.41	3.77	9.38	4.04	4.59	5.50	11.92	6.44	6.98	0.99	1.14	0.01	0.00
Nr. children																
0	50.23	47.66	59.84	56.82	45.94	48.51	63.49	74.46	47.00	55.44	44.67	63.57	50.17	71.43	61.04	59.32
1	16.81	15.19	16.95	15.91	20.55	17.85	17.37	12.55	19.56	20.21	17.84	15.50	18.27	12.57	14.14	8.47
2	29.26	29.40	20.56	25.57	24.52	21.05	15.29	10.53	25.67	14.25	27.20	13.18	26.58	14.86	23.17	29.38
≥3	3.70	7.75	2.65	1.70	8.99	12.59	3.85	2.45	7.77	10.10	10.30	7.75	4.97	1.14	1.65	2.82
Tenancy																
Owner	43.76	36.35	56.64	57.39	27.75	13.04	24.17	16.59	22.13	15.80	8.25	4.72	57.92	54.29	95.65	96.61
Owner mtg	32.99	22.29	14.43	12.50	40.36	14.65	29.86	13.85	50.79	13.99	65.64	29.92	24.13	14.86	0.89	0.00
Tenant free	5.20	8.24	7.35	7.58	1.97	1.60	1.48	1.44	1.01	1.81	0.50	0.79	3.00	1.71	1.49	1.13
Tenant mkt	15.34	26.17	20.12	19.13	17.28	44.62	39.72	56.13	20.72	44.04	25.61	64.57	14.09	24.57	0.98	1.13
Tenant red	2.71	6.95	1.46	3.41	12.64	26.09	4.77	11.98	5.35	24.35	0.00	0.00	0.86	4.57	1.00	1.13

Table 4 depicts the incidence of certain individual characteristics across individuals experiencing cumulative deprivation and those who are not in this state. The proportions are obtained using cross-sectional weights and the totals are respectively non-multidimensionally-deprived people and multidimensionally-deprived people. This table provides an intuition of what might be the mean

[18]The retired people have been deleted from the sample. They are are a very small sample which varies from country to country according to retirement-access laws (below 10 units in Germany, France and Sweden; below 100 units in Czech Rep., Belgium, Spain and Italy; below 300 units in Romania).

relation between cumulative deprivation and a specific individual characteristic unconditional to the occurrence of the other characteristics. Starting from the gender, it becomes evident that the incidence of cumulative deprivation is higher among females. This trend is particularly pronounced in Italy, where the percentage of deprived females stands at 76%, compared to an average of 51% among those who are not deprived. The distribution of cumulative deprivation across age groups varies significantly among countries. Italy and France stand out as the only two nations where individuals aged 25 to 35 are more prevalent among the cumulatively deprived. Conversely, the remaining countries tend to have a larger proportion of people aged 35 to 55 among the poor. With regards to employment status, it is apparent that a substantial portion of cumulatively deprived individuals are either inactive or unemployed. Spain and Belgium exhibit a noteworthy fraction of Extra-EU citizens among cumulatively deprived individuals, being it more than twice as high compared to their presence among the non-deprived population. A high incidence of cumulative deprivation among households with children is observed in Spain, Italy, and Romania. Regarding the tenancy status, with the exception of Italy, a substantial proportion of cumulatively deprived people are tenants at standard market rates. A considerable number of those renting at reduced rates are as well falling in cumulative derivation. The percentage of deprived homeowners paying a mortgage is considerably low with respect to non-poor people.

The robust *probit* estimation outputs for each country are shown in the appendix while Figure 10 shows the marginal effect estimates of the dependent response to a variation in the given explanatory variable. Therefore, the marginal effects show the percentage variation in the probability to fall in cumulative deprivation due to a percentage variation in each characteristic. Due to the possible high correlation between gender with activity status or age with tenancy, the model including all regressors is presented together with estimates obtained progressively adding one single regressor starting with sex and age. The marginal effects are computed only on the final model. The p-values for the marginal effects are presented in Table B.12 in the Appendix. From each estimated model is possible to notice how the sign of the relation between gender and cumulative deprivation changes once introducing the activity status categories as regressors. Once controlling for activity status, sex appears less relevant and sometimes negatively related with cumulative deprivation probability, the most evident case is Italy where the sex coefficient is highly significant and positive while it turns negative and not significant once controlling for activity. This happens because the majority of inactive people are females and inactivity, together with unemployment highly influence probability to fall in cumulative deprivation. Age effects are highly varying across countries, primarily due to different age distributions, secondly due to the heterogeneous ageing risks and welfare supports guaranteed by the countries. The probability of being cumulatively deprived strongly rises at decreasing trends in Spain, Romania and Czech Republic. This relation is achieved as well in Belgium once controlling for activity, and in Germany once controlling for tenancy. Migration status appears to be relevant only in Germany. Home ownership is negatively related with cumulative deprivation in all countries. However this relation seems significant only in Belgium, Czech Republic and Sweden. The access to ownership Through mortgages is as well negatively related with the probability to be deprived in Continental, and northern countries. Highly significant ad positive is the effect of living in a poor region in France and Italy. The zeroes in the estimation outcomes indicate too low number of observations for a given category in a country.

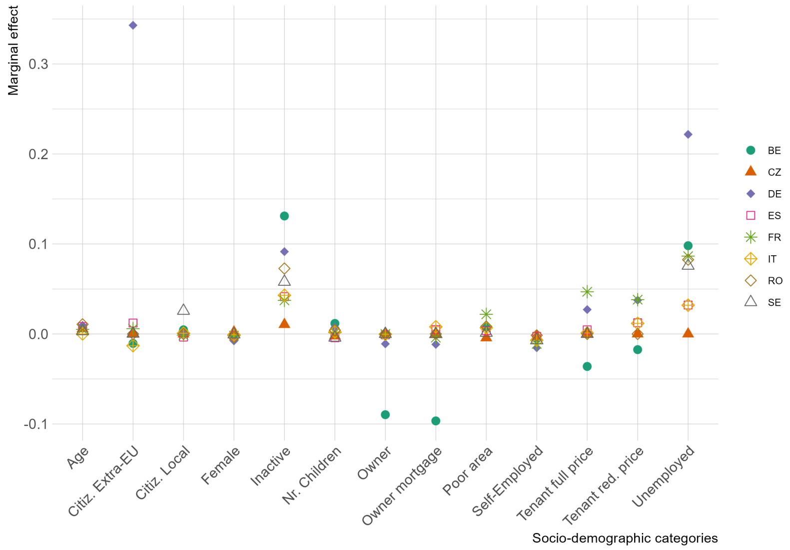


Figure 10: Marginal effects from *probit* regressions

5 Conclusions

This paper contributes to the field of multidimensional poverty and deprivation studies. While poverty and deprivation studies often focus on aggregating multiple dimensions, this research emphasises the need to investigate the interrelations among dimensions of poverty. This paper offers an empirical application of a copula-based tool to study multidimensional dependence within cumulative deprivation conditions. The study has been carried out with EU-SILC data on Belgium, France, Czech Republic, Germany, Italy, Romania, Spain and Sweden between 2007 and 2019. The dimensions selected are income, educational attainment, working conditions, health and housing conditions. The novel technique employed with this study quantifies the non-randomness in the occurrence of cumulative deprivation, revealing the associations among income, health, housing conditions, education and job condition across different European countries and years. This approach, in the broader context of multidimensional poverty measurement, can complement the research by exploring interactions among socioeconomic deprivations and supporting the phases of dimension selection, weights assignment, and phenomenon definition. The study of multidimensional dependence when cumulative deprivation occurs can support the understanding of welfare systems abilities to address the unequal access to social rights and public services.

Results indicate that dependence within poverty dimensions ranges between 5 and 15% in the considered countries. Dependence has been rising since 2011 strongly for Belgium, France, Romania and Spain. Moreover, it emerges an increasing trend in the incidence of cumulative deprivation (the condition experienced by those who fall in the bottom 40% for all dimensions) across all countries, particularly in Belgium, Italy, and France. The highest dependence observed through a pair-wise dominance comparison, shows that post-financial crisis years are dominating the others and in particular for Germany and Belgium. The marginal contribution to overall diagonal dependence is provided by the income and job condition dimensions. The paper describes as well the general conditions of the cumulatively deprived people, showing that the highest income gap between deprived and non-deprived people is taking place in continental and Scandinavian countries. Working conditions are the dimension that most similarly across countries and consistently over time differs between cumulatively deprived and non-cumulatively deprived people. Unsurprisingly, the weaker actors in the society (females, migrants, unemployed people), turned out to be more frequently counted among the cumulatively deprived.

References

Aaberge, R., Atkinson, A. B., and Königs, S. (2018). From classes to copulas: Wages, capital, and top incomes. *The Journal of Economic Inequality*, 16(2):295–320.

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

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

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.

Bossert, W., Chakravarty, S. R., and D'Ambrosio, C. (2013). Multidimensional poverty and material deprivation with discrete data. *Review of Income and Wealth*, 59(1):29–43.

Bourguignon, F. and Chakravarty, S. R. (1999). A family of multidimensional poverty measures. In *Advances in econometrics, income distribution and scientific methodology*, pages 331–344. Springer.

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

Bourguignon, F. and Chakravarty, S. R. (2019). The measurement of multidimensional poverty. In *Poverty, Social Exclusion and Stochastic Dominance*, pages 83–107. Springer.

Chakravarty, S. R. and D'Ambrosio, C. (2006). The measurement of social exclusion. *Review of Income and wealth*, 52(3):377–398.

Deaton, A. (2016). Measuring and understanding behavior, welfare, and poverty. *American Economic Review*, 106(6):1221–1243.

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.

Decancq, K. and Lugo, M. A. (2013). Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews*, 32(1):7–34.

Deutsch, J. and Silber, J. (2005). Measuring multidimensional poverty: An empirical comparison of various approaches. *Review of Income and Wealth*, 51(1):145–174.

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., Sahn, D. E., and Younger, S. D. (2006). Robust multidimensional poverty comparisons. *The economic journal*, 116(514):943–968.

Espinoza-Delgado, J. and Silber, J. (2021). A new inequality-sensitive multidimensional deprivation index (mdi) for dichotomous variables. Technical report, Discussion Papers.

García-Gómez, C., Pérez, A., and Prieto-Alaiz, M. (2020). Copula-based analysis of multivariate dependence patterns between dimensions of poverty in europe. *Review of Income and Wealth*.

Hofert, M., Kojadinovic, I., Mächler, M., and Yan, J. (2019). *Elements of Copula Modeling with R*. Springer.

Idler, E. L. and Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior*, pages 21–37.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC press.

Lemmi, A. A. and Betti, G. (2006). *Fuzzy set approach to multidimensional poverty measurement*, volume 3. Springer Science & Business Media.

Levy, H. and Robinson, M. (2006). *Stochastic dominance: Investment decision making under uncertainty*, volume 34. Springer.

Nelsen, R. B. (2007). *An Introduction to Copulas*. Springer Science & Business Media.

Nussbaum, M. and Sen, A. (1993). *The quality of life*. Clarendon Press.

Pérez, A. (2015). Measuring The Dependence Between Dimensions of Welfare: a Study Based on Spearman's Footrule and Gini's Gamma. page 20.

Pérez, A. and Prieto-Alaiz, M. (2016). 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.

Quinn, C. (2007). Using copulas to measure association between ordinal measures of health and income. Technical report, HEDG, c/o Department of Economics, University of York.

Sen, A. K. (1985). *Commodities and Capabilities*. Elsevier Science Publishers.

Sklar, A. (1996). Random variables, distribution functions, and copulas: a personal look backward and forward. *Lecture Notes-Monograph Series*, pages 1–14.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Univ. de Paris*, 8:229–231.

Stiglitz, J. E., Sen, A., Fitoussi, J.-P., et al. (2009). Report by the commission on the measurement of economic performance and social progress.

Taylor, M. (2016). Multivariate measures of concordance for copulas and their marginals. *Dependence Modeling*, 4(1).

Terzi, S. and Moroni, L. (2022). Local concordance and some applications. *Social Indicators Research*, 161(2-3):457–470.

Tkach, K. and Gigliarano, C. (2020). Multidimensional poverty index with dependence-based weights. *Social Indicators Research*, pages 1–30.

Tsui, K.-y. (2002). Multidimensional poverty indices. *Social choice and welfare*, 19(1):69–93.

Van Doorslaer, E. and Jones, A. M. (2003). Inequalities in self-reported health: Validation of a new approach to measurement. *Journal of Health Economics*, 22(1):61–87.

Yitzhaki, S. and Schechtman, E. (2013). *The Gini methodology: A primer on a statistical methodology*. Springer.

A Appendix: Figures

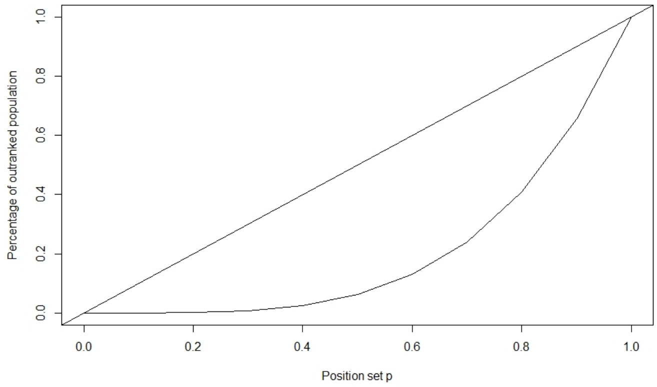


Figure 11: Downward Diagonal Dependence Curve

B Appendix: Tables

Table B.1: DDDI and Standard Errors from 100 repetitions

Country	Year	DDDI	Standard Error	Country	Year	DDDI	Standard Error
BE	2009	0.106	0.00004	DE*	2009	0.113	0.00003
BE	2010	0.103	0.00004	DE*	2010	0.111	0.00003
BE	2011	0.121	0.00004	DE*	2011	0.114	0.00005
BE	2012	0.116	0.00004	DE*	2012	0.128	0.00003
BE	2013	0.142	0.00003	DE*	2013	0.127	0.00003
BE	2014	0.165	0.00004	DE*	2014	0.141	0.00003
BE	2015	0.165	0.00004	DE*	2015	0.134	0.00004
BE	2016	0.166	0.00005	DE*	2016	0.121	0.00003
BE	2017	0.170	0.00004	DE*	2017	0.131	0.00005
BE	2018	0.152	0.00004	DE*	2018	0.123	0.00004
BE	2019	0.172	0.00005	DE*	2019	0.111	0.00005
IT	2007	0.081	0.00003	FR	2007	0.096	0.00003
IT	2008	0.076	0.00002	FR	2008	0.095	0.00003
IT	2009	0.086	0.00003	FR	2009	0.103	0.00003
IT	2010	0.082	0.00002	FR	2010	0.097	0.00004
IT	2011	0.097	0.00002	FR	2011	0.102	0.00003
IT	2012	0.100	0.00002	FR	2012	0.094	0.00003
IT	2013	0.107	0.00002	FR	2013	0.119	0.00003
IT	2014	0.119	0.00002	FR	2014	0.133	0.00003
IT	2015	0.112	0.00002	FR	2015	0.129	0.00003
IT	2016	0.098	0.00002	FR	2016	0.129	0.00004
IT	2017	0.097	0.00003	FR	2017	0.133	0.00004
IT	2018	0.092	0.00003	FR	2018	0.152	0.00004
IT	2019	0.096	0.00003	FR	2019	0.153	0.00005
RO	2007	0.022	0.00004	SE	2008	0.026	0.00004
RO	2008	0.047	0.00009	SE	2009	0.031	0.00004

Continued on next page

Table B.1 Continued from previous page

Country	Year	DDDI	Standard Error	Country	Year	DDDI	Standard Error
RO	2009	0.063	0.00014	SE	2010	0.034	0.00004
RO	2010	0.055	0.00007	SE	2011	0.026	0.00005
RO	2011	0.034	0.00013	SE	2012	0.046	0.00004
RO	2012	0.045	0.00010	SE	2013	0.043	0.00004
RO	2013	0.059	0.00015	SE	2014	0.054	0.00005
RO	2014	0.075	0.00018	SE	2015	0.051	0.00006
RO	2015	0.081	0.00012	SE	2016	0.049	0.00005
RO	2016	0.093	0.00018	SE	2017	0.047	0.00006
RO	2017	0.087	0.00009	SE	2018	0.042	0.00006
RO	2018	0.089	0.00013	SE	2019	0.039	0.00006
RO	2019	0.085	0.00009	CZ	2007	0.039	0.00004
ES	2007	0.087	0.00002	CZ	2008	0.036	0.00004
ES	2008	0.088	0.00002	CZ	2009	0.035	0.00004
ES	2009	0.107	0.00002	CZ	2010	0.030	0.00005
ES	2010	0.096	0.00002	CZ	2011	0.029	0.00005
ES	2011	0.102	0.00002	CZ	2012	0.028	0.00005
ES	2012	0.106	0.00003	CZ	2013	0.041	0.00005
ES	2013	0.125	0.00002	CZ	2014	0.049	0.00007
ES	2014	0.146	0.00002	CZ	2015	0.050	0.00005
ES	2015	0.137	0.00002	CZ	2016	0.045	0.00006
ES	2016	0.139	0.00002	CZ	2017	0.040	0.00006
ES	2017	0.130	0.00002	CZ	2018	0.040	0.00009
ES	2018	0.130	0.00002	CZ	2019	0.037	0.00008
ES	2019	0.129	0.00003				

Note: DDDI values for Germany are referred to four dimensions: Income, health, work, education

Table B.2: P-values for t-test on mean difference between $DDDI(d)$ on all dimensions and $DDDI(d-1)$ computed removing one dimension at a time.

country / year	BE	CZ	DE	ES	FR	IT	RO	SE
2007					no inc.: 0.011		no edu.: 0.145	
2008								no hou.: 0.025
2009		no hea: 0.076						
2010							no wrk.: 0.995	
2011								
2012							no wrk.: 0.337	
2013								
2014		no hea: 0.996						
2015					no hou.: 0.052			
2016								
2017		no hou, hlt, wrk.: 0.995						
2018		no hou, hea, wrk, edu.: 0.995						
2019		no hou, hea, wrk, edu.: 0.995			no wrk.: 0.001			

Note: all blank cells indicate that all t-test p-values are strictly below 0.001

Table B.3: Deascriptive statistics of maximal score position - Year 2019

Country	min	max	median	mean	Q 0.01	Q 0.05	Q 0.1	Q 0.3	Q 0.5
BE	0.0752	0.9999	0.8244	0.7816	0.2400	0.4016	0.4935	0.7151	0.8244
CZ	0.1532	0.9999	0.8799	0.8396	0.4268	0.5775	0.6594	0.7851	0.8799
DE	0.0596	0.9999	0.8310	0.7667	0.2411	0.3754	0.4718	0.7070	0.8310
ES	0.0724	0.9999	0.8508	0.8182	0.3500	0.5108	0.5946	0.7835	0.8508
FR	0.0563	0.9999	0.8255	0.7872	0.2782	0.4434	0.5231	0.7409	0.8255
IT	0.1048	0.9999	0.8650	0.8145	0.3325	0.4957	0.5788	0.7738	0.8650
RO	0.0637	0.9999	0.8523	0.8050	0.3377	0.5062	0.5715	0.7447	0.8523
SE	0.1310	0.9998	0.8707	0.8197	0.3595	0.4998	0.5849	0.7693	0.8707

Table B.4: Probit regression - Belgium 2019

Dependent variable: Cumulative deprivation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.1437*	0.1414*	0.1499*	0.1472*	-0.0108	-0.0023	-0.0029
	(2.47)	(2.41)	(2.55)	(2.51)	(-0.16)	(-0.03)	(-0.04)
Age		0.0336	0.0360	0.0312	0.0761**	0.0846**	0.0826**
		(1.33)	(1.43)	(1.16)	(2.68)	(2.86)	(2.80)
Age^2		-0.0002	-0.0003	-0.0002	-0.0008*	-0.0008*	-0.0008*
		(-0.84)	(-0.89)	(-0.62)	(-2.38)	(-2.44)	(-2.37)
Citizenship							
Local			-0.1602	-0.1600	-0.0532	0.0456	0.0582
			(-1.60)	(-1.60)	(-0.46)	(0.39)	(0.48)
Extra-EU			0.2263	0.2251	-0.0275	-0.1605	-0.1519
			(1.63)	(1.63)	(-0.17)	(-0.99)	(-0.94)
Nr. of children				0.0242	0.0972**	0.1388***	0.1364***
				(0.68)	(2.62)	(3.68)	(3.58)
Activity status							
Self-Employed					-0.1460	-0.1930	-0.1872
					(-0.75)	(-1.03)	(-1.00)
Unemployed					1.2386***	0.9885***	0.9749***
					(11.69)	(9.13)	(8.90)
Inactive					1.3018***	1.1505***	1.1406***
					(16.35)	(14.20)	(13.92)
Tenancy							
Owner						-0.8339**	-0.8217**
						(-3.08)	(-3.02)
Owner with Mortgage						-0.9535***	-0.9375***
						(-3.57)	(-3.49)
Tenant full price						-0.2314	-0.2280
						(-0.88)	(-0.86)
tenant red. price						-0.1052	-0.0929
						(-0.39)	(-0.34)
Poor area							0.0941
							(1.30)
Constant	-1.7080***	-2.6820***	-2.6467***	-2.5864***	-3.9669***	-3.7066***	-3.7293***
	(-38.55)	(-5.08)	(-5.00)	(-4.71)	(-6.65)	(-5.47)	(-5.45)
Observations	7114	7114	7084	7084	7084	7084	7084

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.5: Probit regression - Czech Republic 2019

Dependent variable: Cumulative deprivation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.1614	0.1765	0.1763	0.1678	0.1153	0.1132	0.1082
	(1.57)	(1.57)	(1.57)	(1.49)	(1.03)	(0.99)	(0.95)
Age		0.2420***	0.2386***	0.2301***	0.2836***	0.2797***	0.2835***
		(4.01)	(4.00)	(4.12)	(4.11)	(4.22)	(4.18)
Age^2		-0.0021***	-0.0021***	-0.0020***	-0.0026***	-0.0026***	-0.0026***
		(-3.32)	(-3.30)	(-3.39)	(-3.50)	(-3.55)	(-3.54)
Citizenship							
Local			-0.5952	-0.5879	-0.5716	-0.3902	-0.2639
			(-1.83)	(-1.81)	(-1.81)	(-1.42)	(-0.93)
Extra-EU			0.0000	0.0000	0.0000	0.0000	0.0000
			(.)	(.)	(.)	(.)	(.)
Nr. of children				-0.0838	-0.0814	-0.0960	-0.0969
				(-1.13)	(-1.08)	(-1.35)	(-1.36)
Activity status							
Self-Employed					-0.1191	-0.1328	-0.1440
					(-0.65)	(-0.71)	(-0.76)
Unemployed					0.1852	0.1739	0.2017
					(0.61)	(0.55)	(0.63)
Inactive					0.4270**	0.4027**	0.3981**
					(3.01)	(2.72)	(2.71)
Tenancy							
Owner						-0.7228**	-0.7194**
						(-2.96)	(-2.94)
Owner with Mortgage						-0.4851	-0.4792
						(-1.86)	(-1.82)
Tenant full price						-0.6106*	-0.6351*
						(-2.39)	(-2.48)
Tenant red. price						0.0000	0.0000
						(.)	(.)
Poor area							-0.2025
							(-1.38)
Constant	-2.4956***	-9.0554***	-8.3923***	-8.0251***	-9.2727***	-8.7636***	-8.8004***
	(-31.69)	(-6.50)	(-6.13)	(-6.36)	(-5.94)	(-5.60)	(-5.53)
Observations	8542	8542	8408	8408	8408	8157	8157

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.6: Probit regression - Germany 2019

Dependent variable: Cumulative deprivation	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.0458	0.0530	0.0521	0.0522	-0.0844	-0.0864
	(1.12)	(1.25)	(1.23)	(1.23)	(-1.73)	(-1.76)
Age		0.0377	0.0393	0.0428	0.0988***	0.1072***
		(1.73)	(1.81)	(1.93)	(3.74)	(3.99)
Age^2		-0.0001	-0.0002	-0.0002	-0.0008**	-0.0008**
		(-0.55)	(-0.65)	(-0.82)	(-2.81)	(-2.89)
Citizenship						
Extra-EU			0.3753**	0.3712**	0.4847**	0.5404**
			(2.82)	(2.79)	(3.01)	(3.27)
Nr. of children				-0.0234	0.0307	0.0824*
				(-0.78)	(0.94)	(2.50)
Activity status						
Self-Employed					-0.4136**	-0.3938*
					(-2.63)	(-2.49)
Unemployed					1.5027***	1.3539***
					(19.47)	(16.99)
Inactive					1.1202***	1.0871***
					(19.59)	(19.03)
Tenancy						
Owner						-0.1556
						(-0.85)
Owner with Mortgage						-0.1657
						(-0.90)
Tenant full price						0.2831
						(1.58)
Tenant red. price						0.3696
						(1.92)
Constant	-1.6241***	-3.0634***	-3.0755***	-3.1169***	-4.6972***	-5.1180***
	(-52.78)	(-6.34)	(-6.38)	(-6.44)	(-7.91)	(-8.26)
Observations	11383	11383	11383	11383	11379	11379

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.7: Probit regression - Spain 2019

Dependent variable Cumulative deprivation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.1238*	0.1299*	0.1256*	0.1297*	-0.0368	-0.0408	-0.0380
	(2.02)	(2.05)	(2.01)	(2.07)	(-0.49)	(-0.54)	(-0.50)
Age		0.1621***	0.1704***	0.1941***	0.2172***	0.2149***	0.2153***
		(5.12)	(5.57)	(6.32)	(6.62)	(6.55)	(6.59)
Age^2		-0.0017***	-0.0017***	-0.0020***	-0.0023***	-0.0023***	-0.0023***
		(-4.94)	(-5.24)	(-5.98)	(-6.40)	(-6.29)	(-6.31)
Citizenship							
Local			-0.1296	-0.1202	-0.1414	-0.0957	-0.0885
			(-0.87)	(-0.80)	(-0.81)	(-0.49)	(-0.46)
Extra-EU			0.2885	0.3138	0.2252	0.2232	0.2359
			(1.46)	(1.59)	(0.97)	(0.96)	(1.00)
n_child				-0.1277***	-0.0985*	-0.1088*	-0.1100*
				(-3.32)	(-2.22)	(-2.46)	(-2.48)
Activity status							
Self-Employed					-0.1716	-0.1653	-0.1711
					(-1.41)	(-1.35)	(-1.40)
Unemployed					0.7142***	0.7275***	0.7110***
					(7.13)	(7.20)	(6.84)
Inactive					0.8098***	0.8229***	0.8148***
					(8.62)	(8.56)	(8.36)
Tenancy							
Owner						-0.0381	-0.0294
						(-0.25)	(-0.19)
Owner with Mortgage						0.1153	0.1261
						(0.74)	(0.81)
Tenant full price						0.1016	0.1189
						(0.60)	(0.70)
Tenant red. price						0.2573	0.2767
						(1.15)	(1.24)
Poor area							0.0944
							(1.27)
Constant	-2.1844***	-5.9601***	-6.1135***	-6.4867***	-7.0975***	-7.1752***	-7.2556***
	(-51.92)	(-8.11)	(-8.68)	(-9.30)	(-9.35)	(-9.05)	(-9.30)
Observations	19931	19931	19815	19815	19758	19758	19758

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.8: Probit regression - France 2019

Dependent variable Cumulative deprivation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.0374	0.0417	0.0365	0.0411	-0.0291	-0.0405	-0.0288
	(0.63)	(0.70)	(0.61)	(0.69)	(-0.45)	(-0.62)	(-0.44)
Age		0.0315	0.0329	0.0464	0.0676	0.0794*	0.0785*
		(1.01)	(1.04)	(1.37)	(1.91)	(2.31)	(2.26)
Age^2		-0.0004	-0.0004	-0.0006	-0.0008*	-0.0009*	-0.0009*
		(-1.09)	(-1.09)	(-1.45)	(-2.03)	(-2.31)	(-2.23)
Citizenship							
Local			-0.2325	-0.2349	-0.1499	-0.0122	-0.0219
			(-1.27)	(-1.28)	(-0.89)	(-0.07)	(-0.12)
Extra-EU			0.1442	0.1565	0.0320	0.0123	0.0842
			(0.68)	(0.73)	(0.15)	(0.06)	(0.40)
Nr. of children				-0.0543	-0.0193	0.0185	0.0211
				(-1.70)	(-0.60)	(0.56)	(0.62)
Activity status							
Self-Employed					-0.3381*	-0.2907*	-0.3041*
					(-2.39)	(-1.97)	(-2.06)
Unemployed					0.9963***	0.8547***	0.8382***
					(11.77)	(9.89)	(9.65)
Inactive					0.6207***	0.5073***	0.4893***
					(7.33)	(5.94)	(5.66)
Owner						0.0323	-0.0009
						(0.14)	(-0.00)
Tenancy							
Owner with Mortgage						-0.0861	-0.1149
						(-0.38)	(-0.49)
Tenant full price						0.6491**	0.6321**
						(2.86)	(2.73)
Tenant red. price						0.5505*	0.5534*
						(2.42)	(2.39)
Poor area							0.3388**
							(3.07)
Constant	-1.8535***	-2.4478***	-2.2904***	-2.4812***	-3.1693***	-3.9056***	-4.1606***
	(-42.20)	(-3.75)	(-3.43)	(-3.60)	(-4.26)	(-5.40)	(-5.70)
Observations	11327	11327	11189	11189	11186	11186	11182

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.9: Probit regression - Italy 2019

Dependent variable Cumulative deprivation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.2451***	0.2419***	0.2413***	0.2378***	-0.0277	-0.0273	-0.0167
	(4.36)	(4.27)	(4.26)	(4.19)	(-0.41)	(-0.41)	(-0.25)
Age		-0.0242	-0.0254	-0.0349	0.0005	0.0002	0.0023
		(-1.20)	(-1.26)	(-1.58)	(0.02)	(0.01)	(0.09)
Age^2		0.0003	0.0003	0.0004	0.0000	0.0000	-0.0000
		(1.31)	(1.32)	(1.66)	(0.02)	(0.05)	(-0.03)
Citizenship							
Local			0.0501	0.0507	0.0451	0.0561	0.0166
			(0.36)	(0.36)	(0.31)	(0.38)	(0.11)
Extra-EU			-0.3210	-0.3304	-0.3664	-0.3687	-0.3653
			(-1.68)	(-1.73)	(-1.87)	(-1.88)	(-1.86)
Nr. of children				0.0473	0.0584	0.0514	0.0456
				(1.41)	(1.59)	(1.46)	(1.31)
Activity status							
Self-Employed					-0.3399*	-0.3352*	-0.3348*
					(-2.32)	(-2.30)	(-2.31)
Unemployed					0.5828***	0.5858***	0.5628***
					(7.35)	(7.40)	(6.89)
Inactive					0.6927***	0.7017***	0.6720***
					(9.61)	(9.79)	(9.09)
Tenancy							
Owner						-0.0103	-0.0075
						(-0.09)	(-0.06)
Owner with Mortgage						0.1299	0.1474
						(0.97)	(1.09)
Tenant full price						0.0240	0.0220
						(0.19)	(0.17)
Tenant red. price						0.2084	0.2059
						(0.94)	(0.94)
Poor area							0.1387*
							(2.29)
Constant	-2.1476***	-1.7072***	-1.6899***	-1.5475**	-2.3621***	-2.3985***	-2.4629***
	(-46.95)	(-4.00)	(-3.68)	(-3.20)	(-4.51)	(-4.50)	(-4.64)
Observations	20491	20491	20487	20487	20425	20425	20425

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.10: Probit regression - Romania 2019

Dependent variable Cumulative deprivation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.2575**	0.2560**	0.2560**	0.2506**	-0.0838	-0.0815	-0.0759
	(2.72)	(2.69)	(2.69)	(2.64)	(-0.74)	(-0.72)	(-0.67)
Age		0.1573***	0.1573***	0.1342***	0.2232***	0.2238***	0.2267***
		(4.08)	(4.09)	(3.73)	(5.65)	(5.60)	(5.75)
AGe^2		-0.0019***	-0.0019***	-0.0015***	-0.0027***	-0.0027***	-0.0027***
		(-4.26)	(-4.26)	(-3.87)	(-5.97)	(-5.93)	(-6.06)
Citizenship							
Local			0.0000	0.0000	0.0000	0.0000	0.0000
			(.)	(.)	(.)	(.)	(.)
Extra-EU			0.0000	0.0000	0.0000	0.0000	0.0000
			(.)	(.)	(.)	(.)	(.)
Nr. of children				0.1544**	0.1267*	0.1336*	0.1323*
				(2.65)	(2.10)	(2.19)	(2.13)
Activity status							
Self-Employed					-0.0439	-0.0330	-0.0654
					(-0.27)	(-0.20)	(-0.38)
Unemployed					1.0424***	1.0498***	1.0496***
					(3.49)	(3.49)	(3.54)
Inactive					0.9962***	0.9997***	0.9852***
					(8.08)	(8.05)	(7.84)
Tenancy							
Owner						0.1630	0.1655
						(0.37)	(0.37)
Owner with Mortgage						0.0000	0.0000
						(.)	(.)
Tenant full price						0.2292	0.2503
						(0.43)	(0.46)
Tenant red. price						0.0000	0.0000
						(.)	(.)
Poor area							0.1754
							(1.68)
Constant	-2.1540***	-5.3278***	-5.3288***	-5.0357***	-6.8216***	-6.9883***	-7.1364***
	(-29.10)	(-6.35)	(-6.35)	(-6.18)	(-7.74)	(-7.28)	(-7.56)
Observations	8055	8055	8052	8052	8052	7864	7864

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.11: Probit regression - Sweden 2019

Dependent variable Cumulative deprivation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.0606	0.0573	0.0465	0.0592	0.0109	-0.0088	-0.0089
	(0.67)	(0.63)	(0.51)	(0.66)	(0.12)	(-0.09)	(-0.09)
Age		0.0002	0.0038	0.0347	0.0458	0.0509	0.0509
		(0.01)	(0.10)	(0.86)	(1.08)	(1.17)	(1.17)
Age^2		0.0000	-0.0000	-0.0004	-0.0005	-0.0006	-0.0006
		(0.11)	(-0.03)	(-0.89)	(-1.08)	(-1.12)	(-1.12)
Citizenship							
Local			0.3023	0.2958	0.3373	0.6228*	0.6203*
			(1.29)	(1.25)	(1.30)	(2.03)	(2.02)
Extra-EU			0.1894	0.2333	0.0690	0.0030	0.0001
			(0.69)	(0.83)	(0.22)	(0.01)	(0.00)
Nr. of children				-0.1445**	-0.1090*	-0.0699	-0.0700
				(-2.92)	(-2.27)	(-1.44)	(-1.44)
Activity status							
Self-Employed					-0.2577	-0.1760	-0.1748
					(-1.61)	(-1.04)	(-1.04)
Unemployed					0.9255***	0.7768***	0.7774***
					(5.00)	(3.96)	(3.97)
Inactive					0.7551***	0.6590***	0.6591***
					(5.63)	(4.77)	(4.77)
Tenancy							
Owner						-0.5442**	-0.5473**
						(-2.74)	(-2.74)
Owner with Mortgage						-0.5907***	-0.5918***
						(-5.44)	(-5.45)
Tenant full price						0.0000	0.0000
						(.)	(.)
Poor area							0.0216
							(0.17)
Constant	-1.9251***	-2.0194*	-2.3488**	-2.8054***	-3.2413***	-3.3510***	-3.3527***
	(-28.93)	(-2.54)	(-2.88)	(-3.32)	(-3.58)	(-3.57)	(-3.56)
Observations	6093	6093	6066	6066	6066	5926	5926

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.12: Marginal effects of probit regressions

	BE	CZ	DE	ES	FR	IT	RO	SE	Pooled
Female	0.0011	0.0020	-0.0057	-0.0052	0.0017	0.0012	-0.0061	0.0004	-0.0009
	(0.20)	(0.84)	(-1.36)	(-2.57)	(0.49)	(0.57)	(-1.65)	(0.10)	(-0.84)
Age	0.0088***	0.0046***	0.0093***	0.0069***	0.0058**	0.0001	0.0103***	0.0014	0.0040***
	(4.17)	(3.60)	(4.17)	(6.51)	(3.19)	(0.12)	(7.19)	(0.80)	(7.76)
Citizenship									
Local	0.0082	0	0	-0.0136	-0.0119	0.0003	0	0.0104	0
	(1.16)	(.)	(.)	(-1.63)	(-1.01)	(0.05)	(.)	(1.75)	(0.00)
Extra-EU	-0.0099	0	0.3430***	-0.0026	-0.0066	-0.0070	0	-0.0037	-0.0035
	(-1.16)	(.)	(-3.63)	(-0.30)	(-0.51)	(-0.98)	(.)	(-0.53)	(-1.03)
Nr. children	0.0075**	-0.0012	0.0076**	-0.0032**	0.0043*	0.0027**	0.0017	-0.0014	0.0007
	(2.74)	(-1.04)	(2.61)	(-2.79)	(2.30)	(2.71)	(1.00)	(-0.65)	(1.17)
Activity status									
Self-employed	-0.0063	-0.0025	-0.0120*	-0.0012	-0.0111**	-0.0062***	0.0013	0.0024	-0.0038***
	(-1.30)	(-1.02)	(-2.23)	(-0.66)	(-2.68)	(-5.30)	(0.41)	(0.41)	(-3.95)
Unemployed	0.0990***	0.0041	0.2398***	0.0326***	0.0834***	0.0425***	0.0677*	0.0570*	0.0503***
	(7.30)	(0.60)	(10.97)	(8.12)	(8.16)	(7.11)	(2.04)	(2.20)	(16.94)
Inactive	0.1220***	0.0107*	0.1693***	0.0483***	0.0414***	0.0443***	0.0703***	0.0331**	0.0539***
	(11.40)	(2.14)	(15.41)	(10.38)	(6.10)	(9.34)	(6.84)	(3.14)	(20.59)
Tenancy									
Owner	-0.0701*	0	-0.0136	-0.0016	-0.0052	0.0032	0	0	0.0004
	(-2.00)	(.)	(-0.94)	(-0.42)	(-0.43)	(0.92)	(.)	(.)	(0.16)
Owner mtg	-0.0768*	0	-0.0128	0.0057	-0.0102	0.0096	0	0	-0.0004
	(-2.19)	(.)	(-0.85)	(1.40)	(-0.88)	(1.95)	(.)	(.)	(-0.16)
Tenant mkt.	-0.0156	0	0.0291	0.0092	0.0382**	0.0025	0	0	0.0182***
	(-0.43)	(.)	(1.85)	(1.80)	(2.91)	(0.68)	(.)	(.)	(6.90)
Tenant red.	-0.0041	0	0.0446*	0.0029	0.0345*	0.0069	0		0.0256***
	(-0.12)	(.)	(2.42)	(0.46)	(2.55)	(0.86)	(.)		(6.74)
Poor region	0.0105*	-0.0045		0.0035	0.0228***	0.0065***	0.0075*	0.0035	0.0071***
	(2.07)	(-1.86)		(1.77)	(4.80)	(3.45)	(2.28)	(0.76)	(6.07)
N	7084	8157	11379	19758	11182	20425	7864	5926	81124

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$