

# Measuring poverty in multidimensional contexts

Iñaki Permanyer<sup>1</sup>

**Abstract:** When measuring multidimensional poverty it is reasonable to expect that the trade-offs between variable pairs can differ depending on whether the concerned pairs are complements or substitutes. Yet, currently existing approaches based on deprivation count distributions unrealistically assume that all pairs of variables are related in the same way – an unfortunate circumstance that undermines the possibilities of identifying the poor, aggregating their poverty levels and modeling non-trivial interactions between variables in highly flexible ways. This paper, which aims at modeling non-trivial relational structures across variables both in the identification and aggregation steps, is a first contribution towards addressing these inadequacies. The approach has been axiomatically characterized to flesh out the normative foundations upon which it is based and has a vast potential for application.

**Keywords:** Multidimensional poverty measurement, Complementarity, Substitutability, Identification, Aggregation.

**JEL Classification:** I3; I32; D63; O1

---

<sup>1</sup> Corresponding Author. Centre d'Estudis Demogràfics, Carrer de Ca n'Altayó, Edifici E-2, Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra (Spain). Phone: (+34) 935813060, e-mail: inaki.permanyer@uab.es. **Acknowledgements:** Different versions of this paper have been presented at several conferences and seminars (SCW 2014, ECINEQ 2015, SCW 2016): I am grateful to its participants for their valuable comments and suggestions. In particular, I am indebted to Eugenio Peluso, Casilda Lasso de la Vega, Buhong Zheng, Salvador Barberà, Jordi Massó, Sabina Alkire, Suman Seth, Gastón Yalonetzky, Xavi Ramos, Coral del Rio, Olga Alonso-Villar, Carlos Gradín, Rafael Salas, Juan Gabriel Rodriguez and, more generally, to the members of EQUALITAS. The research leading to these results has received funding from the European Research Council (ERC-2014-StG-637768, EQUALIZE project), the CERCA Programme from Generalitat de Catalunya and from the Spanish Ministry of Economy and Competitiveness “Ramón y Cajal” Research Grant Program (RYC-2013-14196).

## 1. Introduction

Who is poor and who is not? How poor are the poor? These are the fundamental ‘identification’ and ‘aggregation’ questions suggested by Amartya Sen that must be addressed before any poverty eradication program can be implemented (Sen 1976). While the answer to these questions has been quite satisfactorily addressed when poverty is measured in the space of income distributions (after the seminal contribution by Sen in 1976 the literature on income poverty measurement is huge and is based on a very solid footing – see, for instance, Chakravarty 2009 for a recent survey on the topic), matters become more complicated when the poverty status and its levels are determined using several dimensions at the same time. After the influential writings of Sen (1985, 1987, 1992, 1993), it is nowadays acknowledged that poverty is a multidimensional phenomenon and many scholars have insisted on the necessity of defining poverty measures that go beyond the distribution of income or consumption expenditures alone (see, for instance, Anand and Sen 1997, Atkinson 2003, Bourguignon and Chakravarty 2003, Thorbecke 2007, Alkire and Foster 2011, Aaberge and Brandolini 2015).

While several contributions have identified different classes of multidimensional poverty measures (e.g. Chakravarty et al 1998, Tsui 2002, Bourguignon and Chakravarty 2003, Chakravarty and D’Ambrosio 2006, Alkire and Foster 2011, Silber and Yalonetzky 2014, Aaberge and Brandolini 2015, Aaberge et al 2015, Datt 2017, Pattanaik and Xu 2018) one of the most fundamental issues in this literature still needs to be addressed: the proper modeling of the relational structure across variables. When combining several variables of potentially different nature into a single measure it is reasonable to expect that the trade-offs between them could differ depending, for instance, on whether the concerned pairs are

complements or substitutes (see Ravallion 2011, 2012 for a conceptually related discussion). Yet, virtually all current approaches to multidimensional poverty measurement rely one way or another on the so-called ‘deprivation count distributions’ – an approach that takes the *amount* of variables in which individuals are deprived as its informational basis – implicitly assuming that all pairs of variables are related in the same way. This is unfortunate because the possibilities of identifying the poor, aggregating their poverty levels and modeling non-trivial interactions between variables in more realistic ways are severely undermined (see Ravallion 2011). Even if the trade-offs variability across alternative pairs of variables was readily identified as a central issue as soon as multidimensional poverty or welfare assessments were proposed (e.g. Atkinson and Bourguignon 1982, Bourguignon and Chakravarty 2003), as of now there are no measures that are able to capture such variability in a satisfactory way.

When poverty is assessed via pecuniary and non-pecuniary attributes simultaneously it is customary to partition the variables composing such measures in mutually exclusive *dimensions*, with several variables within each dimension. Such a partition – which is exogenously given – aims at imposing certain coherence and structure on the variables one is dealing with by clustering them in conceptually related areas (e.g: the dimensions of ‘Health’, ‘Education’ or ‘Standard of Living’). Here we argue that the partition of variables across dimensions lends itself to a natural relational structure – with the variables belonging to the same (resp. alternative) dimensions being more ‘similar’ (resp. ‘dissimilar’) among themselves – that has been systematically ignored in current approaches to multidimensional poverty measurement. Consider the following examples.

**Example 1.** Assume multidimensional poverty is assessed using the variables  $V_1$  = ‘Income’,  $V_2$  = ‘Years of Schooling’,  $V_3$  = ‘Self-assessed Health’ and  $V_4$  = ‘Health insurance’ (example

taken from Alkire and Foster 2011, p.483). These variables can be naturally partitioned in two dimensions: ‘Capacity to make a living’ (denoted as  $D_1$ , including  $V_1$  and  $V_2$ ) and ‘Health’ (denoted as  $D_2$ , including  $V_3$  and  $V_4$ ). When deciding how to identify the poor, one might argue that the lack of deprivation in one variable could eventually compensate for the deprivation experienced in the other variable within the same dimension (i.e. having a health insurance might somehow compensate a low health status in a ‘health’ dimension, or having a high-quality education could compensate temporary low income levels). Therefore, individuals could be labeled as ‘poor’ when they experience simultaneous deprivations at least in  $V_1$  and  $V_2$  (something which would severely hinder that individual’s capacity to make a decent living) or in  $V_3$  and  $V_4$  (an alarming circumstance for that individual’s health), but *not* when they experience deprivation in one variable within  $D_1$  and in one variable within  $D_2$ .

**Example 2.** In the same 4-variable 2-dimensional setting, one could alternatively argue that each variable is essential to enjoy a decent living in the corresponding dimension (so that there is no possibility of compensation within dimensions) but that only individuals that are deprived in both dimensions have good reasons to be identified as ‘poor’. Under this alternative specification, individuals experiencing deprivations in one dimension only would not be identified as ‘poor’.

In these examples, what counts to be identified as poor is not only the *quantity* of variables in which individuals are deprived but also the *qualitative* relationships that might exist between them. Presently, there is no multidimensional poverty methodology that is able to identify the poor in the ways described in examples 1 and 2 (see section 3.1). Existing identification methods basically *count* the number of existing deprivations irrespective of the dimensions they belong to (see Atkinson 2003, Alkire and Foster 2011, Silber and Yalonetzky

2014, Aaberge and Brandolini 2015), thus ignoring the non-trivial compensation patterns that might exist within and between dimensions.<sup>2</sup> Such disregard for the relational structure across variables also has crucial implications for the aggregation step. When measuring how poor are the poor, current approaches assume that *all* pairs of variables are either complements or substitutes and that the elasticity of substitution is constant across them – an unduly restrictive assumption that is very unlikely to be satisfied in practice. Using an axiomatically characterized approach, this paper is a first step towards addressing the aforementioned inadequacies both in the ‘identification’ and ‘aggregation’ steps.

The rest of the paper is organized as follows. After introducing some basic notations in section 2, in section 3 we present the axiomatic characterization of different families of ‘identification functions’. We start with the basic one-dimensional case (i.e. there is no partition of the underlying variables – this includes all approaches currently used in the literature<sup>3</sup>) and then proceed to the multidimensional case (i.e. the variables are partitioned across multiple dimensions). In section 4 we propose and axiomatically characterize new aggregation methods that allow introducing non-trivial relational structures across variables. Like in the previous section we start with the single dimensional case and then proceed to the multidimensional one. Section 5 concludes. The proofs are relegated to the appendix.

---

<sup>2</sup> Yalonetzky (2014) makes room for defining identification functions that can potentially take into account the relationship between variables when defining the poverty status of individuals. Yet, in the context of three or more variables (which is the one we are interested in this paper), the author only explores the extreme cases of the ‘union’ and ‘intersection’ approaches – see definitions in section 2.

<sup>3</sup> These are the well-known ‘union’, ‘intersection’ and, more generally, the ‘intermediate’ or ‘counting’ identification approaches which are massively used in empirical analysis – see Aaberge and Brandolini 2015 and Alkire et al 2015 and the definitions in section 2.

## 2. Notation and definitions

Let  $N$  be the set of individuals and  $D$  the set of variables<sup>4</sup> under consideration, with  $n := |N| \geq 1, d := |D| \geq 2$ . For any natural number  $d \geq 2$ , let  $X^d := \{0, 1\}^d$ . For  $\mathbf{x}, \mathbf{y} \in X^d$ , we write  $\mathbf{x} \geq \mathbf{y}$  whenever  $x_j \geq y_j$  for all  $j \in \{1, \dots, d\}$  and say that  $\mathbf{x}$  vector-dominates  $\mathbf{y}$ . Analogously, we write  $\mathbf{x} > \mathbf{y}$  whenever  $x_j \geq y_j$  for all  $j \in \{1, \dots, d\}$  with at least one strict inequality, and say that  $\mathbf{x}$  strictly vector-dominates  $\mathbf{y}$ .  $\mathbb{R}_+$  is the set of non-negative real numbers and  $\mathbb{N}_+$  the set of strictly positive natural numbers. Let  $\mathbf{a} = (a_1, \dots, a_d)$  be a vector of positive numbers summing up to 1, whose  $j^{\text{th}}$  coordinate  $a_j$  is interpreted as the normalized weight associated with variable  $j$ . Let  $\Delta_d = \{(a_1, \dots, a_d) \in \mathbb{R}_+^d \mid \sum_i a_i = 1\}$  be the  $d$ -dimensional simplex. Within this set, we denote by  $\mathbf{1}/d := (1/d, \dots, 1/d)$  the equal weights vector.

The achievement of individual  $i$  in attribute  $j$  is assumed to be measurable in a cardinal scale and will be denoted by  $y_{ij} \in \mathbb{R}_+$ . For each attribute  $j$  we consider a poverty threshold  $z_j \in \mathbb{R}_{++}$  indicating the minimum quantity necessary for a subsistence level— which in this paper we consider as exogenously given. Whenever  $y_{ij} \leq z_j$ , we say that individual  $i$  is *deprived* in attribute  $j$ . Very often, poverty measures are defined in the space of deprivations rather than achievements.<sup>5</sup> One common measure of the deprivation experienced by individual  $i$  in variable  $j$  is the following<sup>6</sup>

$$\gamma_{ij} := \left( \text{Max} \left\{ \frac{z_j - y_{ij}}{z_j}, 0 \right\} \right)^c, \quad (1)$$

<sup>4</sup> The terms ‘variable’, ‘indicator’ or ‘attribute’ will be used interchangeably in this paper.

<sup>5</sup> The alternative approach advocated by Ravallion (2011) of working in the space of attainments is not followed in this paper because (i) it might be possible for a poor person to be lifted out of poverty as a result of an increment in a nondeprived dimension, (ii) it does not allow keeping track of the dimension-specific deprivations simultaneously.

<sup>6</sup> There exist other definitions of deprivation gaps (see Table 1 in Permanyer (2014:4) for other examples). Since alternative definitions do not alter the findings of the paper, we have chosen the one that is more commonly used in the literature.

where  $c \geq 0$ . Whenever  $y_{ij}$  is measured in a cardinal scale  $\gamma_{ij}$  is well-defined for any  $c \geq 0$ ,<sup>7</sup> and  $\gamma_{ij} \in [0, 1]$ . In particular, when  $c = 1$ ,  $\gamma_{ij}$  is the so-called *normalized deprivation gap* (which measures in a  $[0, 1]$ -scale the distance between a given achievement  $y_{ij}$  and the corresponding poverty line  $z_j$ ) and when  $c = 2$ , we obtain the *squared deprivation gap*. For an individual  $i$ , we define the corresponding *vector of deprivations gaps* as  $\gamma_i := (\gamma_{i1}, \dots, \gamma_{id}) \in [0, 1]^d$  (when no confusion arises and its use is unnecessary, we might omit the individuals' label  $i$ ). A *deprivation matrix*  $\Gamma$  is a  $n \times d$  matrix with entries in  $[0, 1]$  containing the deprivation gap vectors of  $n$  individuals in the different rows. The set of all  $n \times d$  deprivation matrices is denoted as  $\mathcal{G}_{n \times d}$ , and we define  $\mathcal{G} := \bigcup_{n \in \mathbb{N}} \bigcup_{d \in \mathbb{N}} \mathcal{G}_{n \times d}$ . Following Bourguignon and Chakravarty (2003), an *identification function*  $\zeta : \mathcal{G}_{1 \times d} \rightarrow \{0, 1\}$  is a mapping from individual  $i$ 's deprivation gap vector  $\gamma_i$  to an indicator variable in such a way that  $\zeta(\gamma_i) = 1$  if person  $i$  is multidimensionally poor and  $\zeta(\gamma_i) = 0$  otherwise. For analytical clarity, we write the identification function  $\zeta$  as the composite  $\zeta = \rho \circ \omega$ , with

$$\omega : \mathcal{G}_{1 \times d} \rightarrow X^d \tag{2}$$

and

$$\rho : X^d \rightarrow \{0, 1\}. \tag{3}$$

The function  $\omega$  converts the deprivation gap vector  $\gamma_i$  into a vector of 0s and 1s of length  $d$  indicating whether individual  $i$  is deprived or not in the different variables taken into account (where 1 denotes deprivation and 0 non-deprivation). The set  $X^d$  contains all possible combinations of deprivations/non-deprivations across  $d$  variables. Its generic members – referred to as *deprivation profiles* – are denoted as  $\mathbf{x} = (x_1 \dots x_d)$ , with  $x_j \in \{0, 1\}$  indicating the deprivation status in variable  $j$ . The profile  $\mathbf{0} := (0 \dots 0)$  corresponds to someone who is

---

<sup>7</sup> The value  $c = 0$  is typically chosen when the variables that are used to assess multidimensional poverty are measured in an ordinal scale. However, in this paper we will focus on the cardinal case (see Remark 3 in section 4.2).

not deprived in any variable and  $\mathbf{1} := (1\dots 1)$  to someone who is deprived in all variables. For each  $j \in \{1, \dots, d\}$  we define  $\mathbf{e}_j := (0\dots 010\dots 0)$  as the deprivation profile where deprivation is only experienced in variable  $j$ . The following definitions are based on the relation of vector dominance among pairs of elements in  $X^d$ .

**Definition 1:** Let  $Z$  be any non-empty subset of  $X^d$ . The *up-set* of  $Z$  is defined as  $Z^\uparrow := \{\mathbf{x} \in X^d \mid \exists \mathbf{z} \in Z \text{ s.t. } \mathbf{z} \leq \mathbf{x}\}$  and the *set of least deprived profiles in  $Z$*  is defined as  $\mathcal{L}(Z) := \{\mathbf{x} \in Z \mid \nexists \mathbf{y} \in Z \setminus \{\mathbf{x}\} \text{ s.t. } \mathbf{y} \leq \mathbf{x}\}$ .

If  $Z$  is a subset of  $X^d$ ,  $Z^\uparrow$  is the set of deprivation profiles vector-dominating at least one member of  $Z$ . On the other hand,  $\mathcal{L}(Z)$  is the set of elements in  $Z$  that do not vector-dominate any other element in  $Z$ .

Since  $\omega$  is simply an indicator function specifying whether individuals are deprived or not in the different variables, here we focus our attention on different ways in which  $\rho$  can be defined. With a slight abuse of notation, the functions  $\rho$  will also be referred to as ‘*identification functions*’. Let  $\Omega_d := \{\rho : X^d \rightarrow \{0, 1\}\}$  be the set of all possible identification functions for  $d$  variables. Since each identification function  $\rho$  is uniquely characterized by the set of elements  $\rho^{-1}(1) \subseteq X^d$  (or its complement  $\rho^{-1}(0)$ ) and  $X^d$  has  $2^d$  elements, it follows that  $\Omega_d$  has  $2^{2^d}$  elements. To simplify notation we will write  $P_\rho := \rho^{-1}(1)$  and  $R_\rho := \rho^{-1}(0)$  (i.e.  $P_\rho$  and  $R_\rho$  are the set of deprivation profiles that  $\rho$  identifies as ‘poor’ and ‘non-poor’ respectively). The set  $\Omega := \{\Omega_d\}_{d \in \mathbb{N}_+ \setminus \{1\}}$  contains all identification functions for all possible  $d \geq 2$ . In section 3 we impose several conditions on the elements of  $\Omega_d$  to pin down several classes of identification functions  $S_d \subseteq \Omega_d$ . For clarification purposes, it is sometimes useful to graph the Hasse diagram corresponding to the set  $X^d$  (whose elements are the nodes of the diagram) and the partial order generated by vector dominance  $\leq$  (represented by the



edges between nodes) to represent identification functions  $\rho \in \Omega_d$ . In Figure 1 we show two examples of identification functions  $(\rho_1, \rho_2)$  for the case  $d = 4$  that will be useful to illustrate other sections of the paper.

Let  $Q_\rho := \{i \in N \mid \rho(\omega(\gamma_i)) = 1\}$  be the set of individuals considered to be poor according to the identification function  $\rho$ . After completing the identification step, Sen (1976) (and all the ensuing literature on poverty measurement after him) suggests to proceed to the aggregation step, i.e. summarize the information on the extent of poverty among the poor into a single number. While the identification functions analyzed in section 3 are based on the vectors of 0s and 1s that obtain after applying the  $\omega$  function to individuals' deprivation gap vectors, the aggregation step takes  $\rho$  as given and associates with a deprivation matrix  $\Gamma$  an overall level of multidimensional poverty. The chosen aggregation method will be denoted as  $\mathcal{A}$  (section 4 presents and axiomatically characterizes several of such methods). Borrowing notation from Alkire and Foster (2011:477), we define a *multidimensional poverty methodology* as the tuple  $(\rho, \mathcal{A})$ .

## 2.1 The counting approach

The 'counting approach identification functions' can be written as the composite  $\iota_k \circ c_{\mathbf{a}}$ , with

$$c_{\mathbf{a}} : X^d \rightarrow [0, 1] \tag{4}$$

and

$$\iota_k : [0, 1] \rightarrow \{0, 1\}. \tag{5}$$

For any  $\mathbf{x} \in X^d$  and any  $\mathbf{a} \in \Delta_d$ , the function  $c_{\mathbf{a}}$  is defined as  $c_{\mathbf{a}}(\mathbf{x}) = \sum_{j=1}^{j=d} a_j x_j$ , that is:  $c_{\mathbf{a}}$  computes the weighted proportion of deprivations experienced by someone with deprivation

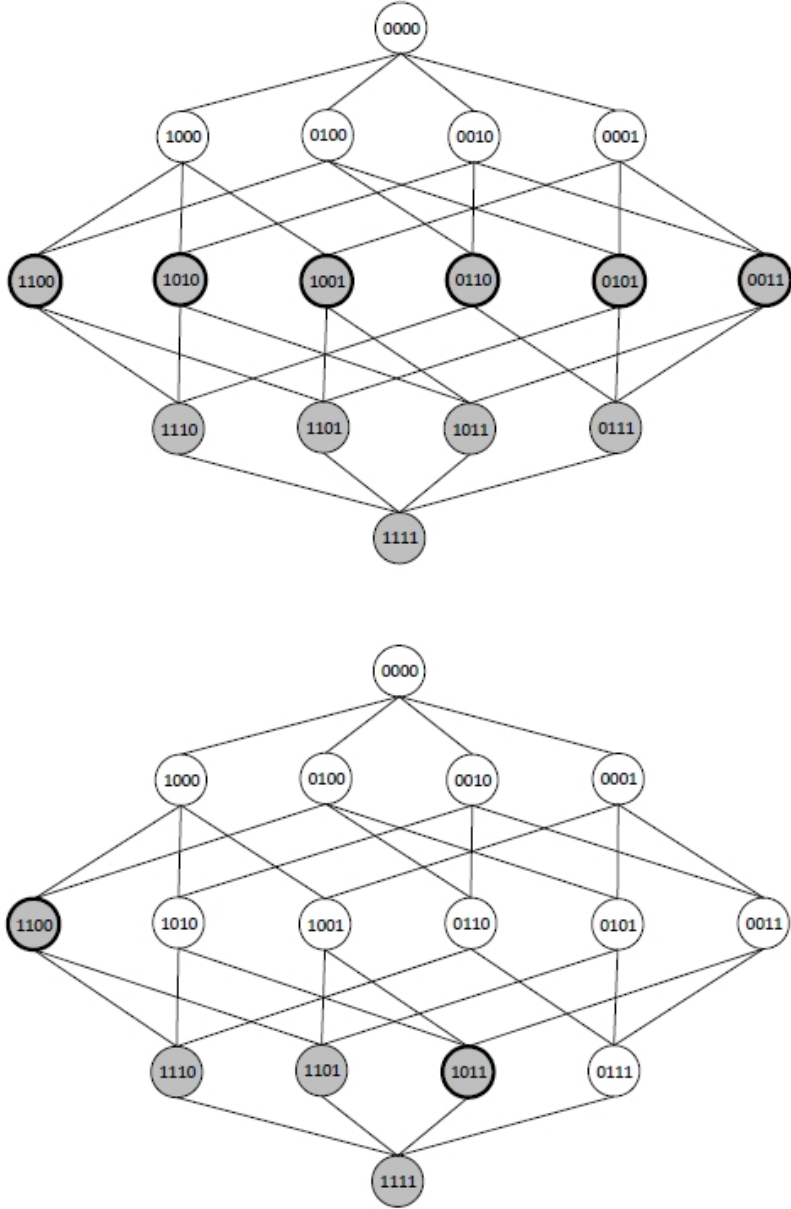


Figure 1: Figure 1a (top), 1b (bottom). Two examples of identification functions for the partially ordered set  $(X^4, \leq)$ . The shaded circles in the top and bottom panels are the members of  $P_{\rho_1}$  and  $P_{\rho_2}$  respectively. The first identification function comes from the ‘unweighted counting approach’ using  $k = 1/2$  as deprivation threshold (see equation (8)). The second one comes from the ‘weighted counting approach’, using  $a_1 = 1/2, a_2 = 1/4, a_3 = 1/8, a_4 = 1/8$  as weights and  $k = 3/4$  as deprivation threshold (see equation (7)). The least deprived profiles for  $\rho_1$  and  $\rho_2$  (i.e.  $\mathcal{L}(P_{\rho_1})$  and  $\mathcal{L}(P_{\rho_2})$ ) are highlighted in bold.

profile  $\mathbf{x}$ . Lastly, for any  $s \in [0, 1]$  and for any  $k \in (0, 1]$ ,  $\iota_k$  is defined as

$$\iota_k(s) = \begin{cases} 1 & \text{if } s \geq k \\ 0 & \text{if } s < k \end{cases}. \quad (6)$$

The  $\iota_k \circ c_{\mathbf{a}}$  function takes a value of 1 whenever the weighted proportion of deprivations attains a certain deprivation threshold  $k$  (which is exogenously given) and a value of 0 otherwise. With this notation, if one fixes any  $k \in (0, 1]$  we can define the following class of identification functions:

$$\mathcal{W}_d(k) := \{\rho \in \Omega_d \mid \rho(\mathbf{x}) = \iota_k(c_{\mathbf{a}}(\mathbf{x})) \text{ for some } \mathbf{a} \in \Delta_d\}. \quad (7)$$

This is the set of identification functions belonging to the so-called ‘weighted counting approach’ with deprivation threshold  $k$ . The higher the value of  $k$ , the more difficult it is that an individual ends up being classified as poor. When  $k \leq \min_j a_j$ , a function  $\rho \in \mathcal{W}_d(k)$  corresponds to the so-called ‘union approach’, and when  $k = 1$ ,  $\rho \in \mathcal{W}_d(1)$  is equivalent to the ‘intersection approach’. The Hasse diagrams shown in Figures 1a and 1b illustrate examples of identification functions  $\rho \in \mathcal{W}_d(k)$  for certain combinations of  $\mathbf{a}$  and  $k$  when  $d = 4$ . In Figure 1a, we have chosen  $a_1 = a_2 = a_3 = a_4 = 1/4$  and  $k = 1/2$  (representing  $\rho_1$ ) and in Figure 1b,  $a_1 = 1/2, a_2 = 1/4, a_3 = 1/8, a_4 = 1/8$  and  $k = 3/4$  (representing  $\rho_2$ ). When the weighting vector  $\mathbf{a}$  turns out to weight all variables equally we obtain the class

$$\mathcal{C}_d(k) := \{\rho \in \Omega_d \mid \rho(\mathbf{x}) = \iota_k(c_{\mathbf{1}/d}(\mathbf{x}))\}, \quad (8)$$

which will be referred to as ‘unweighted counting approach’ with deprivation threshold  $k$ . Since both approaches use deprivation thresholds within variables ( $z_j$ ) and an overall threshold  $k$  across them, they are generally known as the ‘dual cutoff’ identification method – or,

simply, the ‘counting approach’ (see Alkire and Foster 2011). The sets

$$\mathcal{W}_d : = \bigcup_{k \in (0,1]} \mathcal{W}_d(k) \quad (9)$$

$$\mathcal{C}_d : = \bigcup_{k \in (0,1]} \mathcal{C}_d(k) \quad (10)$$

contain all weighted and unweighted counting identification methods for  $d$  variables respectively. Lastly, the sets  $\mathcal{W} := \{\mathcal{W}_d\}_{d \in \mathbb{N}_+ \setminus \{1\}}$  and  $\mathcal{C} := \{\mathcal{C}_d\}_{d \in \mathbb{N}_+ \setminus \{1\}}$  are the collection of all weighted and unweighted counting identification methods for all possible variables.

### 3. Identification of the poor

In this section we present different families of identification functions obtained after imposing increasingly demanding axioms on the elements of  $\Omega_d$ . We start assuming that all variables belong to the same dimension (the following subsection deals with the more general case where variables are partitioned across several dimensions). Let  $S_d \subseteq \Omega_d$  be a set of identification functions.

*Non-triviality* (NTR):  $\rho$  is a non-constant function for all  $\rho \in S_d$ .

*Monotonicity* (MON): Let  $\mathbf{x}, \mathbf{y} \in X^d$ . If  $\mathbf{x} \leq \mathbf{y}$ , then  $\rho(\mathbf{x}) \leq \rho(\mathbf{y})$  for all  $\rho \in S_d$ .

NTR prevents the identification function being constant across all deprivation profiles. MON ensures that if individual  $A$  experiences deprivations at least in the same variables as those where another individual  $B$  experiences deprivations, and possibly in others, then  $A$  qualifies at least as much as  $B$  to be identified as multidimensionally poor. Because of their uncontroversial nature, we posit that the set of identification functions satisfying these two axioms should be the universe of reference from which identification functions should be drawn; it will be denoted as  $\mathcal{I}_d$  and referred to as the set of *consistent* identification

functions.<sup>8</sup> The set  $\mathcal{I} := \{\mathcal{I}_d\}_{d \in \mathbb{N}_+ \setminus \{1\}}$  is the collection of consistent identification functions for all possible sets of variables. The following result uniquely characterizes the elements of  $\mathcal{I}$ .

**Proposition 1.** For all  $d \in \mathbb{N}_+ \setminus \{1\}, \rho \in \mathcal{I}_d \Leftrightarrow (\mathcal{L}(P_\rho))^\dagger = P_\rho$ .

**Proof:** See the appendix.

According to Proposition 1, the sets of ‘poor profiles’  $P_\rho$  derived from consistent identification functions are *uniquely* characterized and represented by the corresponding subsets of ‘least deprived elements’  $\mathcal{L}(P_\rho)$ . When choosing a sensible set of poor profiles  $P_\rho$ , the subsets  $\mathcal{L}(P_\rho)$  are particularly important because their elements determine the least deprived conditions that individuals should experience in order to be considered as poor. Indeed, the sets  $\mathcal{L}(P_\rho)$  can be thought as a generalization of the concept of a poverty line to the multidimensional context (i.e. they determine the boundary separating the poor from the non-poor: when  $\mathbf{x} \in \mathcal{L}(P_\rho)$  and  $\mathbf{y}, \mathbf{z} \in X^d$  are such that  $\mathbf{y} < \mathbf{x} \leq \mathbf{z}$ , then  $\mathbf{y} \in R_\rho$  and  $\mathbf{z} \in P_\rho$ ).

The minimal structure imposed on consistent identification functions makes ample room to incorporate different criteria – many of which can be qualitative in nature – when deciding who should be considered as multidimensionally poor. The downside of such flexibility is that the set of consistent identification functions can perhaps be too unwieldy for certain practical purposes. Indeed, current approaches to multidimensional poverty measurement have considerably reduced the class of admissible identification functions by imposing a set of axioms that, as we will now see, are quite restrictive.

---

<sup>8</sup> The naming ‘consistent identification functions’ is reminiscent of the so-called ‘poverty consistency’ property introduced by Lasso de la Vega (2010) in the context of the counting approach (see section 2.1). In that paper an identification function is ‘poverty consistent’ if, when identifying a person with a deprivation score equal to  $s$  as poor, it also considers as poor anybody whose deprivation score count is at least as high as  $s$ . Clearly, the consistency condition proposed here is more general than the ‘poverty consistency’ property in that the former applies to *any* identification function, not just those relying on scores.

*Variable Anonymity* (VAN): For any  $i, j \in \{1, \dots, d\}$ ,  $\rho(\mathbf{e}_i) = \rho(\mathbf{e}_j)$  for all  $\rho \in S_d$ .

*Independence* (IND): Let  $\mathbf{x}, \mathbf{y} \in X^d$  be two deprivation profiles such that for some variable  $i \in \{1, \dots, d\}$ ,  $x_i = y_i$ . Let  $\mathbf{x}', \mathbf{y}' \in X^d$  be two other deprivation profiles such that  $x_j = x'_j$  and  $y_j = y'_j$  for  $j \neq i$  and  $x'_i = y'_i$ . Then  $\rho(\mathbf{x}) \leq \rho(\mathbf{y})$  implies  $\rho(\mathbf{x}') \leq \rho(\mathbf{y}')$  for all  $\rho \in S_d$ .

VAN requires all variables to be treated symmetrically. IND is a classical separability assumption ensuring that the removal or addition of the same deprivation from two deprivation profiles should preserve the weak ordering among them. With these two additional axioms we can present the following result.

**Theorem 1:** Let  $S_d \subseteq \Omega_d$ . The identification functions  $\rho \in S_d$  satisfy MON, NTR, VAN and IND if and only if  $S_d = \mathcal{C}_d$ .

**Proof:** See the appendix.

Theorem 1 is inspired by the seminal work of Pattanaik and Xu (1990) in the field of freedom of choice measurement. Whenever one is willing to accept the four aforementioned axioms simultaneously, the unweighted counting approach obtains. When it comes to characterize the weighted counting approach  $\mathcal{W}_d$  one would be tempted to simply drop the VAN axiom from the list. Yet, it turns out that IND is not powerful enough, so it needs to be strengthened. For our next axiom, we need the following definition.

**Definition 2:** Consider two hypothetical societies, each with  $m > 1$  individuals, with deprivation profiles  $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ ,  $(\mathbf{y}_1, \dots, \mathbf{y}_m)$ . We say that these two societies are *equivalent* if for each variable  $j \in \{1, \dots, d\}$  the number of individuals that are deprived in that variable is the same in both societies, that is:  $\sum_{i=1}^{i=m} x_{ij} = \sum_{i=1}^{i=m} y_{ij}$  for all  $j \in \{1, \dots, d\}$ .

*Compensation* (COM): Consider two equivalent societies with deprivation profiles  $(\mathbf{x}_1, \dots, \mathbf{x}_m)$  and  $(\mathbf{y}_1, \dots, \mathbf{y}_m)$ . Assume that  $\rho(\mathbf{x}_1) \geq \rho(\mathbf{y}_1), \dots, \rho(\mathbf{x}_{m-1}) \geq \rho(\mathbf{y}_{m-1})$  for all  $\rho \in S_d$ . Then, one must have that  $\rho(\mathbf{x}_m) \leq \rho(\mathbf{y}_m)$  for all  $\rho \in S_d$ .

COM states that in two equivalent societies, it is not possible that *all* individuals in one of them qualify at least as much as the others to be identified as multidimensionally poor. That is: it is not possible to match the individuals of these equivalent societies in such a way that each individual of the former is ranked at least as high as the corresponding individual of the latter with some of these rankings being strict. It is easy to show that COM imposes separability across variables – indeed, COM can be seen as a stronger version of classical independence axioms usually employed in welfare analysis (see, for instance, Blackorby, Primont and Russell 1978): if COM holds for a given  $S_d \subseteq \Omega_d$ , then IND holds as well for all  $\rho \in S_d$ ; however, the opposite is not necessarily true (the proof of this statement is shown in the Appendix before the proof of Theorem 2).<sup>9</sup>

**Theorem 2:** Let  $S_d \subseteq \Omega_d$ . The identification functions  $\rho \in S_d$  satisfy MON, COM and NTR if and only if  $S_d = \mathcal{W}_d$ .

**Proof:** See the appendix.

With the last axioms the class of consistent identification functions has been narrowed down considerably to obtain  $\mathcal{W}_d$  and  $\mathcal{C}_d$ , the state-of-the art methodologies that are massively used in practice to identify the multidimensionally poor. However, such simplification comes at a cost: while NTR and MON are indisputable, VAN, IND and COM are contentious. VAN is not very meaningful when some variables are much more relevant than

<sup>9</sup> In the context of finite sets it turns out that Independence axioms are not strong enough to guarantee the additive representations we are looking for – as opposed to what happens in continuous settings with richer structures (see, Fishburn 1970 and Blackorby, Primont and Russell 1978).

others, while IND and COM impose full separability across all variables and treat them as if they were, so-to-speak, mutually orthogonal. These axioms are responsible for the exceedingly uniform way in which all pairs of variables are treated, irrespective of whether they might be complements or substitutes. Such quantitatively-driven approach can be highly unsatisfactory as it might fail to identify the poor in those contexts where the different variables we are dealing with are partitioned across several dimensions and where there might be non-trivial compensation patterns between and within them (see examples 1 and 2 in the introduction). In the following section we address these issues by proposing another subclass of consistent identification functions that does not satisfy the overly restrictive IND and COM axioms.

### 3.1 Identification in multiple dimensions

Suppose now that the set of variables  $D$  is partitioned in  $G$  dimensions ( $G$  being a natural number strictly greater than one). Such exogenously given partition in thematic areas is naturally derived from the index's design and is unlikely to generate much controversy. Let  $\Psi_G$  denote the set of partitions of  $D$  into  $G$  dimensions  $D_1, \dots, D_G$  where at least two dimensions contain at least two variables<sup>10</sup> and let  $d_g := |D_g|$ . Clearly,  $d = \sum_g d_g$ . Given any such partition there is a one-to-one correspondence between  $X^d$  and  $X^{d_1} \times \dots \times X^{d_G}$ , so any deprivation profile  $\mathbf{x} = (x_1, \dots, x_d) \in X^d$  can be uniquely associated with  $(\mathbf{x}_1, \dots, \mathbf{x}_g, \dots, \mathbf{x}_G)$ , where  $\mathbf{x}_g = (x_{g1}, \dots, x_{gd_g}) \in X^{d_g}$  and  $x_{gv} \in \{0, 1\}$  indicates the deprivation status in variable  $v$  within dimension  $g$ .

**Definition 3:** Let  $S_d \subseteq \Omega_d$  be a set of identification functions belonging to a collection  $\{S_j\}_{j \in \mathbb{N}_+ \setminus \{1\}}$  and let  $\psi \in \Psi_G$ . We define the *two-stage identification functions* associated

---

<sup>10</sup>This excludes trivial partitions in which each dimension is composed of one variable only and forces  $d \geq 4$ .



with the pair  $(S_d, \{S_j\}_{j \in \mathbb{N}_+ \setminus \{1\}})$  as the set

$$S_d^\psi := \{\rho \in \Omega_d \mid \rho(\mathbf{x}) = \rho^b(\rho_1^w(\mathbf{x}_1), \dots, \rho_G^w(\mathbf{x}_G)), \text{ with } \rho^b \in S_G, \rho_g^w \in S_{d_g} \forall g \in \{1, \dots, G\}\}. \quad (11)$$

The members of  $S_d^\psi$  are identification functions constructed in two steps. Initially, the functions  $\rho_g^w : X^{d_g} \rightarrow \{0, 1\}$  decide about the deprivation status in each dimension and secondly, overall deprivation across dimensions is assessed via  $\rho^b : X^G \rightarrow \{0, 1\}$ . These functions – all of which belong to the collection  $\{S_j\}_{j \in \mathbb{N}_+ \setminus \{1\}}$  – are referred to as *within-* and *across-dimension* identification functions respectively. The superscript  $\psi$  is used to indicate the dependence of the two-stage identification functions on the choice of the partition of  $D$  into  $G$  dimensions  $D_1, \dots, D_G$ . Hence,  $\mathcal{C}_d^\psi$  (resp.  $\mathcal{W}_d^\psi, \mathcal{I}_d^\psi$ ) is the set of two-stage identification functions for  $d$  variables whereby both the within- and across-dimension identification functions belong to the unweighted counting approach  $\mathcal{C}$  (resp. the weighted counting approach  $\mathcal{W}$ , consistent identification functions  $\mathcal{I}$ ).

The identification functions verbally described in Examples 1 and 2 (see Introduction) are members of  $\mathcal{C}_d^\psi$ . In those examples, we have four variables partitioned in two dimensions (i.e.  $d = 4, G = 2$ ) and  $\psi \in \Psi_2$  is the partition that groups the first two variables ( $V_1, V_2$ ) in the first dimension ( $D_1$ ) and the last two ( $V_3, V_4$ ) in the second one ( $D_2$ ). In Example 1, both within-dimension identification functions are based on the intersection approach, that is: individuals have to be deprived in both variables within the same dimension to be considered as deprived in that dimension. Afterwards, the across-dimension identification function is based on the union approach, i.e. whenever individuals are deprived in any of the two dimensions, they are considered to be multidimensionally poor. In Example 2, both within-dimension identification functions are based on the union approach and the across-dimension identification function is based on the intersection approach. In general, if there

are good reasons to believe that the relationships between pairs of variables differ within and across dimensions, two-stage identification functions can a priori be more appropriate than counting approaches. The following result highlights the key difference between them.

**Proposition 2:** Let  $\psi \in \Psi_G$ . One has that  $\mathcal{C}_d^\psi, \mathcal{W}_d^\psi$  and  $\mathcal{I}_d^\psi$  satisfy NTR and MON, but none of them satisfy COM.

**Proof:** See the appendix.

The first part of proposition 2 ensures that, whenever  $S_d \in \{\mathcal{C}_d, \mathcal{W}_d, \mathcal{I}_d\}$ , the two-stage identification functions in  $S_d^\psi$  satisfy the minimal consistency requirements (i.e. they satisfy NTR and MON, so  $S_d^\psi \subset \mathcal{I}_d$ ). The second one fundamentally differentiates  $S_d^\psi$  from the weighted counting approach  $\mathcal{W}$ . Since any set of identification functions satisfying NTR, MON and COM *must* correspond to the weighted counting approach (see Theorem 2), the failure of  $\mathcal{C}_d^\psi, \mathcal{W}_d^\psi$  and  $\mathcal{I}_d^\psi$  to satisfy COM ensures that none of those sets of two-stage identification functions coincides with  $\mathcal{W}$ . Stated otherwise: no matter what weighting scheme and deprivation threshold we choose, the weighted counting approach is unable to generate the sets of two-stage identification functions  $\mathcal{C}_d^\psi, \mathcal{W}_d^\psi$  and  $\mathcal{I}_d^\psi$ . Hence, no matter what weighting scheme and deprivation threshold we choose, the weighted counting approach will not be able to generate the two-stage identification functions shown in Examples 1 and 2.

Since  $\mathcal{I}_d \supset \mathcal{W}_d \supset \mathcal{C}_d$ , it is straightforward to check that  $\mathcal{I}_d^\psi \supset \mathcal{W}_d^\psi \supset \mathcal{C}_d^\psi$  for any  $\psi \in \Psi_G$ . In addition, when the partition  $\psi$  is trivial ( $G = 1$ ) the multiple dimensions approach reduces to the classical single-dimensional case. This way, we have generated a new set of consistent identification functions that does not comply with the COM axiom and makes room to model non-trivial compensation patterns that might exist between and within dimensions –

thus considerably enlarging the toolkit available to those practitioners aiming at measuring multidimensional poverty.

#### 4. Aggregation of the poor

So far we have been discussing how the partition of variables in different dimensions affects the identification of the poor. We are now going to explore its implications for the ‘aggregation step’. As is standard in the literature, we take a certain identification function  $\rho$  as exogenously given and propose different ways to summarize the extent of poverty among the poor with a single real number (i.e. we suggest different aggregation methods  $\mathcal{A}$ ). For that purpose we introduce the following notation. Let  $\mathcal{G}_{n \times d}^S$  denote the set of  $n \times d$  deprivation matrices whose rows are the same. For a given deprivation gap vector  $\gamma \in \mathcal{G}_{1 \times d}$ , let  $[\gamma] \in \mathcal{G}_{n \times d}^S$  denote the  $n \times d$  deprivation matrix whose rows are equal to  $\gamma$ . A *family of multivariate poverty indices*  $\mathbf{f} : \mathcal{G} \rightarrow \mathbb{R}$  is a set of non trivial functions  $\{f_d\}_{d \in \mathbb{N}}$  where each function  $f_d$  converts an element from the space of deprivation matrices  $\Gamma \in \mathcal{G}_{n \times d}$  into a real number  $f_d(\Gamma)$  indicating the extent of poverty in the corresponding distribution. In this section we present some basic properties one might want to impose on a family of multivariate poverty indices  $\mathbf{f} : \mathcal{G} \rightarrow \mathbb{R}$  to characterize it axiomatically. We begin with the one-dimensional case (i.e. all variables belong to the same dimension) and then proceed to the multidimensional case in the next subsection. Our axioms and results are designed for the cardinal case (i.e. all variables are assumed to be measurable in a cardinal scale).

*Subgroup Decomposability* (SDC): Let  $\Gamma_1 \in \mathcal{G}_{n_1 \times d}, \dots, \Gamma_m \in \mathcal{G}_{n_m \times d}$  be a list of deprivation matrices for  $m$  disjoint subpopulations. Then, for all  $d \in \mathbb{N}$ , poverty on the overall population can be written as  $f_d(\Gamma_1 \cup \dots \cup \Gamma_m) = \sum_l \frac{n_l}{n} f_d(\Gamma_l)$ , where  $n = \sum_l n_l$ .

*Monotonicity in aggregation* (MOA): For all  $A = (a_{ij}), B = (b_{ij}) \in \mathcal{G}_{n \times d}$ , if  $A \neq B$  and

$a_{ij} \geq b_{ij}$  for all  $i \in \{1, \dots, n\}$  and all  $j \in \{1, \dots, d\}$ , then  $f_d(A) \geq f_d(B)$  for all  $d \in \mathbb{N}$ .

*Normalization* (NRM): Consider  $\Gamma = (\gamma_{ij}) \in \mathcal{G}_{n \times d}$ . If  $\gamma_{ij} = 0$  for all  $i \in \{1, \dots, n\}$  and all  $j \in \{1, \dots, d\}$ , then  $f_d(\Gamma) = 0$  for all  $d \in \mathbb{N}$ . If  $\gamma_{ij} = 1$  for all  $i \in \{1, \dots, n\}$  and all  $j \in \{1, \dots, d\}$ , then  $f_d(\Gamma) = 1$  for all  $d \in \mathbb{N}$ .

*Separability* (SEP): Let  $\boldsymbol{\gamma}, \boldsymbol{\delta} \in [0, 1]^d$  be two deprivation gap vectors such that for some variable  $j \in \{1, \dots, d\}$ ,  $\gamma_j = \delta_j$ . Let  $\boldsymbol{\gamma}', \boldsymbol{\delta}' \in [0, 1]^d$  be two other deprivation gap vectors such that  $\gamma_l = \gamma'_l$  and  $\delta_l = \delta'_l$  for  $l \neq j$  and  $\gamma'_j = \delta'_j$ . Then  $f_d([\boldsymbol{\gamma}]) \leq f_d([\boldsymbol{\delta}])$  implies  $f_d([\boldsymbol{\gamma}']) \leq f_d([\boldsymbol{\delta}'])$  for all  $d \in \mathbb{N}$ .

*Homogeneity* (HMG): For any  $\Gamma \in \mathcal{G}_{n \times d}$  and any  $\lambda \in (0, 1]$  one has that  $f_d(\lambda\Gamma) = \lambda f_d(\Gamma)$  for all  $d \in \mathbb{N}$ , where  $\lambda\Gamma$  is the  $\lambda$ -scaling of  $\Gamma$ .

*Homotheticity* (HMT): For any  $\Gamma_1, \Gamma_2 \in \mathcal{G}_{n \times d}$  and any  $\lambda \in (0, 1]$  one has that  $f_d(\Gamma_1) \geq f_d(\Gamma_2) \Leftrightarrow f_d(\lambda\Gamma_1) \geq f_d(\lambda\Gamma_2)$  for all  $d \in \mathbb{N}$ , where  $\lambda\Gamma_1, \lambda\Gamma_2$  are the deprivation matrices  $\Gamma_1, \Gamma_2$  with all their elements scaled by  $\lambda$ .

*Continuity* (CON): For all  $d \in \mathbb{N}$ ,  $f_d$  is a continuous function in its arguments.

SDC allows identifying subgroups where poverty is particularly high and evaluating their contribution to overall poverty levels. Indeed, it is such an intuitive and useful property that it has been imposed on *all* multivariate poverty indices presented in the literature so far. One of the consequences of SDC is that overall poverty can be written as the arithmetic mean of individual poverty levels, thus ensuring that our aggregation method is consistent with the principle of anonymity (i.e. overall poverty does not depend on the identity of the individuals experiencing deprivations). MOA ensures that if the deprivation felt by any individual in any attribute increases, assuming the rest of deprivations do not decrease,

then overall deprivation should increase. Formally, this axiom is the equivalent of MON adapted to the ‘aggregation framework’. NRM establishes that when nobody is deprived in any variable, then poverty should be equal to zero. In addition, if all individuals are fully deprived in all variables, NRM stipulates that poverty should reach its maximal value of one. According to SEP the poverty ranking between two deprivation gap vectors only depends on the set of variables where their values do not coincide, irrespective of what happens in the ones where they coincide. Following Blackorby et al (1978), SEP stipulates that each variable is separable from its complement. HMG ensures that when all deprivation gaps are scaled by a proportionality factor, overall deprivation is scaled by the same factor. HMT is a weaker version of HMG and ensures that the weak poverty ordering between two societies does not change when all deprivations are scaled by the same proportionality factor. Clearly, Homogeneity implies Homotheticity, but not the other way around. Since both HMG and HMT are very standard in the literature, we have used them both to show the difference it makes to impose the one or the other to our poverty measures. Lastly, CON requires that small changes in the deprivations of individuals produce small changes in the corresponding poverty measures (i.e. poverty levels do not change abruptly when individuals’ deprivations are slightly altered). This property ensures that poverty levels will not be dramatically affected by small measurement errors in the data.

**Theorem 3:** Assume we identify the set of poor individuals ( $Q_\rho$ ) via the identification function  $\rho \in \mathcal{I}_d$ . A family of multivariate poverty indices  $\mathbf{f} = \{f_d\}_{d \in \mathbb{N}}$  satisfies SDC, MOA, NRM, SEP, HMG and CON if and only if, for each  $\Gamma \in \mathcal{G}_{n \times d}$ ,  $f_d(\Gamma)$  can be written as

$$\Pi_\theta := \frac{1}{n} \sum_{i \in Q_\rho} \left( \sum_{j=1}^d a_j (\gamma_{ij})^\theta \right)^{1/\theta}, \quad (12)$$

for all  $d \in \mathbb{N}$ , where  $\theta > 0$ ,  $a_j > 0 \forall j \in \{1, \dots, d\}$  and  $\sum_j a_j = 1$ .

**Proof:** See the appendix.

The family of multivariate poverty indices shown in (12) measures individuals' poverty levels averaging the corresponding deprivation gaps vector  $\gamma_i$  using a weighted generalized mean of order  $\theta$ .<sup>11</sup> Clearly, when  $\theta = 1$ ,  $\Pi_\theta$  corresponds to the class of poverty measures  $M_\alpha$  suggested by Alkire and Foster (2011) when  $\alpha = c$ . In addition, (12) can also be seen as a member of some of the multidimensional poverty indices proposed by Bourguignon and Chakravarty (2003). While the previous two measures were originally defined under the assumption that the poor are identified via the counting and the union approach, respectively, the new measure shown in (12) broadens the class of admissible poverty measures incorporating the more general identification functions embodied in  $\rho \in \mathcal{I}_d$ . The choice of  $\theta$  allows modelling different elasticities of substitution between pairs of deprivation gaps: when  $\theta = 1$  there is perfect substitutability and when  $\theta \rightarrow \infty$  there is perfect complementarity. As highlighted by Bourguignon and Chakravarty (2003:40), such elasticity of substitution is the same across *all* pairs of deprivations, a restriction that often might not be very realistic. Inspecting the axiomatic characterization shown in Theorem 3, it is clear that SEP – which treats all dimensions as if they were mutually orthogonal – is responsible for this state of affairs. In the following subsection, we introduce other axioms allowing more flexibility in this regard.

When homogeneity (HMG) is substituted by homotheticity (HMT), it is easy to show that the functional form of the multivariate poverty indices characterized in Theorem 3 can

---

<sup>11</sup>The class of weighted generalized means is well-known and has been widely used in welfare analysis. Higher values of  $\theta$  give more importance to the upper tails of the distribution and vice versa. In the limit, as  $\theta \rightarrow \infty$  (resp.  $\theta \rightarrow -\infty$ ) the generalized mean converges towards the maximum (resp. minimum) of the distribution.

be written as

$$\frac{1}{n} \sum_{i \in Q_\rho} \left( \sum_{j=1}^d a_j (\gamma_{ij})^\theta \right)^{\beta/\theta}, \quad (13)$$

where  $\beta > 0$  (result not shown here but available upon request). This functional form coincides with some of the measures proposed by Bourguignon and Chakravarty (2003). In addition, when  $\theta = 1$  and  $\beta > 1$ , the aggregation function shown in (13) coincides with the one used in the “distribution-sensitive” measures proposed by Datt (2017) and by Pattanaik and Xu (2018).<sup>12</sup> As discussed in those papers, equation (13) can be sensitive to the extent of inequality in the distribution of deprivations depending on the values of  $\beta$ . The use of HMT enlarges the class of admissible indices characterized by Theorem 3 at the cost of introducing some extra complexity in the interpretation of our poverty measures.

## 4.1 Aggregation in multiple dimensions

We now assume that the set of variables  $D$  is partitioned in  $G \geq 1$  dimensions (i.e.  $D = \bigsqcup_g D_g$ ), with each dimension  $D_g$  containing  $d_g$  variables ( $\sum_g d_g = d$ ). We relabel individual’s  $i$  deprivation gaps vector  $\gamma_i = (\gamma_{i1}, \dots, \gamma_{id})$  to identify the specific dimensions where the different deprivations belong to. Given the one-to-one correspondence between  $[0, 1]^d$  and  $[0, 1]^{d_1} \times \dots \times [0, 1]^{d_G}$  we can rewrite  $\gamma_i$  as  $(\gamma_{i1}, \dots, \gamma_{ig}, \dots, \gamma_{iG})$ , where  $\gamma_{ig} = (\gamma_{ig1}, \dots, \gamma_{igd_g})$  is the vector of deprivation gaps in dimension  $D_g$  for each  $g \in \{1, \dots, G\}$ . Hence  $\gamma_{igv}$  is individual’s  $i$  deprivation gap in variable  $v$  belonging to dimension  $g$ . When no confusion arises, we might omit the label  $i$  when it is not necessary. For any  $g \in \{1, \dots, G\}$ , let

$$[0, 1]_g^d := \{(\gamma_1, \dots, \gamma_G) \in [0, 1]^{d_1} \times \dots \times [0, 1]^{d_G} \mid \gamma_j = \mathbf{0} \forall j \neq g\}. \quad (14)$$

<sup>12</sup>Yet, these two measures and the one shown in (13) differ in the identification function they use to select the members of  $Q_\rho$ . While Datt (2017) uses the union approach to identify the poor, Pattanaik and Xu (2018) propose a novel identification function based on individuals’ weighted sum of deprivation gaps  $\sum_j a_j \gamma_{ij}$  (rather than the traditional weighted proportion of deprivations  $\sum_j a_j \omega(\gamma_{ij})$  used in the counting approach).

This is the set of deprivation gap vectors for those individuals who only experience deprivations within dimension  $g$ . Given any partition  $(D_1, \dots, D_G)$  of  $D$ , let

$$[0, 1]_E^d := \{(\gamma_1, \dots, \gamma_G) \in [0, 1]^{d_1} \times \dots \times [0, 1]^{d_G} \mid \gamma_{gu} = \gamma_{gv} \forall u \neq v, \forall g\}. \quad (15)$$

This is the set of deprivation gap vectors whose components are constant within each dimension. Lastly, observe that any deprivation gap matrix  $\Gamma \in \mathcal{G}_{n \times d}$  can be obtained after appending different  $\Gamma_g \in \mathcal{G}_{n \times d_g}$  (one per dimension), so that  $\Gamma = (\Gamma_1 | \dots | \Gamma_G)$ . In order to generalize the family of multivariate poverty indices shown in Theorem 3 to the multidimensional context, we introduce the following axioms.

*Within Dimension Separability* (WDS): Let  $(D_1, \dots, D_G)$  be any partition of  $D$  in  $G \geq 1$  dimensions. Consider any dimension  $g \in \{1, \dots, G\}$  and let  $\gamma, \delta \in [0, 1]_g^d$  be two deprivation gap vectors such that for some variable  $v \in \{1, \dots, d_g\}$ ,  $\gamma_{gv} = \delta_{gv}$ . Let  $\gamma', \delta' \in [0, 1]_g^d$  be two other deprivation gap vectors such that  $\gamma_{gu} = \gamma'_{gu}$  and  $\delta_{gu} = \delta'_{gu}$  for  $u \neq v$  and  $\gamma'_{gv} = \delta'_{gv}$ . Then  $f_d([\gamma]) \leq f_d([\delta])$  implies  $f_d([\gamma']) \leq f_d([\delta'])$  for all  $d \in \mathbb{N}$ .

*Between Dimension Separability* (BDS): Let  $(D_1, \dots, D_G)$  be any partition of  $D$  in  $G \geq 1$  dimensions. Let  $\mathbf{r}, \mathbf{s} \in [0, 1]_E^d$  be two deprivation gap vectors such that  $\mathbf{r}_g = \mathbf{s}_g$  for some dimension  $g \in \{1, \dots, G\}$ . Let  $\mathbf{r}', \mathbf{s}' \in [0, 1]_E^d$  be two other deprivation gap vectors such that  $\mathbf{r}'_j = \mathbf{r}_j$  and  $\mathbf{s}'_j = \mathbf{s}_j$  for  $j \neq g$  and  $\mathbf{r}'_g = \mathbf{s}'_g$ . Then  $f_d([\mathbf{r}]) \leq f_d([\mathbf{s}])$  implies  $f_d([\mathbf{r}']) \leq f_d([\mathbf{s}'])$  for all  $d \in \mathbb{N}$ .

WDS stipulates that the poverty ranking between two individuals who are only deprived in dimension  $g$  just depends on the set of variables where their values differ, irrespective of what happens with the ones where they coincide. In other words, it requires that any variable should be separable from its complement *within the same dimension only*. In this respect,



WDS is much weaker than SEP (the latter requiring any variable to be separable from all other variables irrespective of the dimension they belong to). Indeed, SEP can be seen as a particular case of WDS when  $G = 1$ . Likewise, BDS requires the different dimensions to be separable from each other. Interestingly, SEP can also be seen as a particular case of BDS when  $G = d$  (each variable constitutes one dimension).

**Theorem 4:** Assume we identify the set of poor individuals ( $Q_\rho$ ) via the identification function  $\rho \in \mathcal{I}_d$ . Let  $(D_1, \dots, D_G)$  be any partition of  $D$  in  $G \geq 1$  dimensions. A family of multidimensional poverty indices  $\mathbf{f} = \{f_d\}_{d \in \mathbb{N}}$  satisfies SDC, MOA, NRM, FOC, HMG, CON, WDS and BDS if and only if, for each  $\Gamma \in \mathcal{G}_{n \times d}$ ,  $f_d(\Gamma)$  can be written as

$$\Pi_{\boldsymbol{\theta}}^G := \frac{1}{n} \sum_{i \in Q_\rho} \left( \sum_{g=1}^G a_g \cdot \left[ \sum_{v=1}^{d_g} w_{gv} (\gamma_{igv})^{\theta_g} \right]^{\theta/\theta_g} \right)^{1/\theta}, \quad (16)$$

for all  $d \in \mathbb{N}$ , where  $\boldsymbol{\theta} = (\theta, \theta_1, \dots, \theta_G) \in \mathbb{R}_{++}^{G+1}$ ,  $a_{gv} > 0 \forall g \in \{1, \dots, G\}, \forall v \in \{1, \dots, d_g\}$ ,  $a_g := \sum_{v=1}^{d_g} a_{gv}$ ,  $w_{gv} := a_{gv}/a_g$  and  $\sum_g \sum_v a_{gv} = 1$ .

**Proof:** See the appendix.

## 4.2 Implications and remarks

**Remark 1.** The family of multidimensional poverty indices characterized in Theorem 4 satisfies the following identity:  $f_d(\Gamma) = f_G(f_{d_1}(\Gamma_1), \dots, f_{d_G}(\Gamma_G))$ , where each  $\Gamma_g \in \mathcal{G}_{n \times d_g}$  is the restriction of  $\Gamma$  to dimension  $g \in \{1, \dots, G\}$ . Stated otherwise:  $f_d(\Gamma)$  is a mean of means, where deprivations are averaged within dimensions first and then averaged across dimensions. This way of aggregating deprivation gap vectors can be seen as a particular example of the following – more general – aggregation method. Assume there exist functions  $\Phi : [0, 1]^G \rightarrow [0, 1]$  and  $\varphi_g : [0, 1]^{d_g} \rightarrow [0, 1]$  for each  $g \in \{1, \dots, G\}$  so that individual  $i$ 's poverty level can be measured as  $\Phi(\varphi_1(\gamma_{i11}, \dots, \gamma_{i1d_1}), \dots, \varphi_G(\gamma_{iG1}, \dots, \gamma_{iGd_G}))$ . For obvious

reasons, we call this the *dimension-first two-stage aggregation* method, and leave its more complete exploration for future research.

**Remark 2:** When homogeneity (HMG) is substituted by homotheticity (HMT), it is easy to show that the functional form of the multivariate poverty indices characterized in Theorem 4 can be written as

$$\frac{1}{n} \sum_{i \in Q(P_\rho)} \left( \sum_{g=1}^G a_g \left[ \sum_{v=1}^{d_g} w_{gv} (\gamma_{igv})^{\theta_g} \right]^{\beta_1 \theta / \theta_g} \right)^{\beta_2 / \theta}, \quad (17)$$

where  $\beta_1, \beta_2 > 0$  (result not shown here but available upon request). Hence, relaxing the somewhat restrictive HMG by HMT, we enlarge the class of admissible dimension-first two-stage aggregation methods we can use to measure multidimensional poverty. Yet, such increased flexibility is gained at the cost of complicating the interpretation of these measures. Since the more complicated functional forms shown in (17) are not strictly necessary for the purposes of this paper, in the rest of this section we stick to their simpler version shown in (16).

**Remark 3:** What happens in the ordinal setting where the deprivation gaps are dichotomous (i.e.  $c = 0$  in equation (1))? In that case, CON, HMG and HMT are not well-defined. In addition, the proofs of Theorems 3 and 4 rely on the fact that the deprivation gaps  $\gamma_{ij}$  can take any real value in  $[0, 1]$  (see Appendix). Hence, these two theorems do not apply and their ordinal scale versions should await further research.<sup>13</sup>

**Remark 4:** An important characteristic of multidimensional poverty measures is their sensitivity to correlation increasing switches<sup>14</sup>. As discussed in Bourguignon and Chakravarty

<sup>13</sup>Yet, the lack of axiomatic characterization should not necessarily preclude the use of  $\Pi_\theta$  or  $\Pi_\theta^G$  in empirical applications where the variables one is dealing with are ordinal. Indeed, the  $M_0$  index suggested by Alkire and Foster (2011) for the ordinal setting is not axiomatically characterized but is massively employed in practice.

<sup>14</sup>

(2003:35), these measures might increase or decrease after correlation increasing switches depending on whether the attributes we are taking into consideration are complements or substitutes.<sup>15</sup> Unfortunately, currently existing measures assume that all pairs of variables are either complements or substitutes – an assumption that might not necessarily hold. The poverty measure  $\Pi_{\theta}^G$  shown in (16) makes room for the possibility of having pairs of variables that are complements or substitutes depending on whether they belong to the same or alternative dimensions (see proposition 3 below). Clearly, when  $G = 1$ , equation (16) reduces to equation (12). The new poverty measure depends on parameter  $\theta$  (governing the complementarity or substitutability across dimensions) and the different  $\theta_g$  (governing the complementarity or substitutability between variables within dimension  $D_g$ ). As is clear, whenever  $\theta = \theta_1 = \dots = \theta_G$  the ‘ $G$ –dimensional measure’  $\Pi_{\theta}^G$  is equivalent to the ‘1–dimensional measure’ shown in equation (12), so all pairs of deprivations have the same elasticity of substitution. However, when one departs from that trivial case the levels of complementarity / substitutability between deprivations in the poverty measure  $\Pi_{\theta}^G$  can vary across dimensions. Following Bourguignon and Chakravarty (2003:35), it is trivial to check that when using  $\Pi_{\theta}^G$ , poverty does not decrease (resp. increase) after a correlation increasing switch whenever the concerned variables satisfy the conditions of substitutability (resp. complementarity) stated in the following proposition.

**Proposition 3:** Consider the multidimensional poverty measure  $\Pi_{\theta}^G$ . (i) For any dimension  $D_g$  ( $g \in \{1, \dots, G\}$ ), two variables  $u, v$  belonging to that dimension (i.e.  $u, v \in D_g$ )

---

Consider a scenario with two individuals  $a, b$  and two variables  $u, v$ , where  $a$  has more of  $u$  but less of  $v$  than  $b$ . A ‘correlation increasing switch’ occurs when we interchange the amounts of attribute  $v$  between the two persons. After such change, the marginal distributions are unchanged but  $a$  has more of  $u$  and  $v$  than  $b$ .

<sup>15</sup>In this paper we use the standard ALEP definition of complementarity / substitutability (see Kannai (1980)). That is: when the cross partial derivative of the individual poverty function is positive (resp. negative), the attributes are considered complements (resp. substitutes).

are complements whenever  $\theta_g < \min\{1, \theta\}$ . On the other hand, the same two variables are substitutes whenever  $\theta_g > \max\{1, \theta\}$ . (ii) Assume now the two variables  $u, v$  belong to different dimensions  $D_g, D_h$  ( $g, h \in \{1, \dots, G\}$ ). Then  $u, v$  are complements whenever  $\theta < 1$  and substitutes when  $\theta > 1$ .

**Proof:** See the Appendix.

**Remark 5:** Following Sen (1976) and the ensuing literature on poverty measurement, the multidimensional poverty methodologies proposed in this paper have two separate parts: the identification step ( $\rho$ ) and the aggregation step ( $\mathcal{A}$ ). It is only after identifying who is poor and who is not that one proceeds to summarize the extent of poverty among the poor into a single number. While this sequential approach does not generate any problem for ‘traditional’ income poverty measures, in the multidimensional context it can potentially create different kinds of inconsistencies. One of them might arise because of the lack of coherence between the identification and aggregation steps – which in this paper have been characterized *independently*. While this independence gives ample room to generate identification and aggregation functions in highly flexible ways, it could eventually lead to some kind of mismatch between them.<sup>16</sup>

The other undesirable consequence of defining  $\rho$  and  $\mathcal{A}$  separately is that whenever the identification of the poor  $\rho$  is *not* performed via the union approach, the multidimensional poverty methodology  $(\rho, \mathcal{A})$  will fail to be continuous – even if the aggregation methods  $\mathcal{A}$  proposed in (12), (13), (16) and (17) use continuous functions. Technically, such discontinuities arise because when  $\rho$  does not correspond to the union approach, the deprivations of

---

<sup>16</sup>Yet, it is currently difficult to ascertain whether identification and aggregation functions treat the complementarity / substitutability among variables in a truly coherent way. While it is clear that the ALEP criterion can be used for aggregation functions, we are not aware of any analogous method that can be implemented in the discrete setting of identification functions – an important topic that should be explored in future research.

the non-poor are censored – an issue that currently affects virtually all the multidimensional poverty methodologies using the dual cutoff method proposed by Alkire and Foster (2011) and the consistent identification functions proposed in this paper. This problem and its implications have been discussed at length by Datt (2017) and Pattanaik and Xu (2018). While Datt (2017) suggests to use the union approach to avoid such discontinuities and other distributive-related problems, Pattanaik and Xu (2018) suggest a novel method in which the identification of the poor is based on the weighted sum of deprivation *gaps* (rather than the weighted count of dichotomous deprivations that is customary in the counting approach – see footnote #12). This identification method – which seems to dilute the hitherto crisp boundaries between the identification and aggregation steps – offers an interesting avenue of research that should be explored in the near future.

## 5. Discussion and conclusion

This paper addresses one of the most fundamental challenges that are still pending in the measurement of multidimensional poverty: the modeling of non-trivial relational structures across variables. As opposed to currently existing methods – which assume that *all* pairs of variables are either complements or substitutes and that the elasticity of substitution is constant across them – our approach offers the possibility to flexibly model the tradeoff complexities and subtleties involved in poverty measurement. The fact that some pairs of poverty indicators are complements while others are substitutes can have crucial implications both when identifying the poor and when aggregating their poverty levels. The techniques presented in this paper allow, for the first time, to take these considerations into account both in the identification and aggregation steps. This is accomplished by generalizing and going beyond the ‘deprivation count distributions’ or ‘counting methods’ – whereby individuals’

multidimensional poverty status and depth of poverty are assessed on the basis of the *amount* of variables in which these individuals are deprived – that pervade current approaches to multidimensional poverty measurement (see Atkinson 2003, Bourguignon and Chakravarty 2003, Alkire and Foster 2011, Silber and Yalonetzky 2014, Aaberge and Brandolini 2015). The new identification and aggregation methods proposed here have been axiomatically characterized to flesh out the normative foundations upon which they are based.

Our approach is very general and includes most of the currently existing multidimensional poverty methodologies proposed in the literature as particular cases. Inter alia, it includes the counting approach suggested by Alkire and Foster (2011) that is massively used in empirical applications (like, for instance, in the United Nations’ Multidimensional Poverty Index). Among the multidimensional poverty methodologies that are not covered by the techniques proposed in this paper, it is worth highlighting the ‘dual approach’ proposed by Aaberge et al (2015). Using a social evaluation function that applies a certain distortion to the CDF of the deprivation count across the population, the dual approach is a flexible method that explicitly takes into consideration the complementarity / substitutability among attributes in the aggregation step. Yet, like all other aggregation methodologies proposed in the literature so far, the dual approach assumes that *all* pairs of attributes are either complements or substitutes.

All modeling exercises face inescapable trade-offs between parsimony and realism. In this regard, the two-stage identification and aggregation methods introduced in this paper lie between two extremes: (i) extreme parsimony, in which the relational structure between variables is simply ignored, and (ii) extreme realism, in which one attempts to model the association between *all* variable pairs ( $d(d-1)/2$ ). While the former approach (which represents the current state of affairs) is excessively rough, the latter quickly becomes statistically

intractable as the number of variables increases. Alternatively, the two-stage approach suggested here is flexible enough to capture important aspects of the relational structure between variables without falling prey of statistical over-sophistication.

The tools proposed here have a vast potential for empirical applications in a wide variety of settings, ranging from microeconomic theory to development economics. Among others, they allow analysts to model multidimensional poverty more realistically but face them with questions that are more difficult to answer (e.g. ¿How to choose the identification functions  $\rho \in \mathcal{I}_d$ ?, ¿How to determine the degree of complementarity / substitutability across and within dimensions?). The answer to these questions is highly context-specific and is likely to require econometric callibrating models to estimate the values of the parameters governing the trade-offs within and across dimensions.

To the extent that the success of micro level anti-poverty programs depends on targeting the right individuals and properly assessing their deprivation levels, and that current international cooperation, development and aid programs are guided by the macro level results derived from the corresponding measures, the issues analyzed in this paper have practical and financial implications for the design of effective poverty eradication strategies. Having recently reached the Millennium Development Goals (MDGs) target year, many scholars and policy-makers are currently engaged in an intense debate about what kind of headline poverty indicator should be the most appropriate to guide poverty eradication strategies in the post-2015 global development agenda. Like its predecessor, the first of the so-called Sustainable Development Goals (the SDGs) aims to ‘End Poverty in all its forms everywhere’. This is a good moment to take stock and reflect before uncritically extending use of the counting approach. Other procedures, such as the ones suggested here, exist to identify recipients and assess their poverty levels under one of the greatest international endeavours

of our time to eradicate poverty.

## 6. Appendix

**Proof of Proposition 1:** We start with the ‘if’ part of the proof. Assume that  $\rho \in \mathcal{I}_d$ . We have to prove that  $(\mathcal{L}(P_\rho))^\uparrow = P_\rho$ .

1) We start proving  $(\mathcal{L}(P_\rho))^\uparrow \subset P_\rho$ . Take  $\mathbf{x} \in (\mathcal{L}(P_\rho))^\uparrow$ . Then, there exists some  $\mathbf{z} \in \mathcal{L}(P_\rho)$  such that  $\mathbf{z} \leq \mathbf{x}$  (if  $\mathbf{x} \in \mathcal{L}(P_\rho)$ , then  $\mathbf{z} = \mathbf{x}$ ). Since  $\mathcal{L}(P_\rho) \subset P_\rho$ ,  $\mathbf{z} \in P_\rho$ . In addition, since  $\mathbf{x} \in \mathbf{z}^\uparrow$  and  $\rho \in \mathcal{I}_d$ , one can conclude that  $\mathbf{x} \in P_\rho$ .

2) We now prove  $(\mathcal{L}(P_\rho))^\uparrow \supset P_\rho$ . Take  $\mathbf{x} \in P_\rho$ . If it turns out that  $\mathbf{x} \in \mathcal{L}(P_\rho)$  then we are done. If  $\mathbf{x} \notin \mathcal{L}(P_\rho)$  then there must exist some  $\mathbf{y} \in P_\rho \setminus \{\mathbf{x}\}$  such that  $\mathbf{y} \leq \mathbf{x}$ . Now, if  $\mathbf{y} \in \mathcal{L}(P_\rho) \subset P_\rho$  then  $\mathbf{x} \in \mathbf{y}^\uparrow$ . Since  $\rho \in \mathcal{I}_d$ , one can conclude that  $\mathbf{x} \in (\mathcal{L}(P_\rho))^\uparrow$ . Otherwise, if  $\mathbf{y} \notin \mathcal{L}(P_\rho)$  then we can proceed iteratively until reaching an element belonging to  $\mathcal{L}(P_\rho)$ . That is: since  $X^d$  is finite ( $|X^d| = 2^d$ ) there must exist a finite sequence of vector dominations  $\mathbf{z}_i \leq \mathbf{z}_{i+1}$  from some element  $\mathbf{z}_1 \in \mathcal{L}(P_\rho)$  up to  $\mathbf{x}$  (i.e.:  $\mathbf{z}_1 \leq \mathbf{z}_2 \dots \leq \mathbf{z}_n \leq \mathbf{x}$ ), so that  $\mathbf{x} \in \mathbf{z}_1^\uparrow$ . Since  $\rho \in \mathcal{I}_d$ , one can conclude that  $\mathbf{x} \in (\mathcal{L}(P_\rho))^\uparrow$ .

This proves the ‘if’ part of the proposition. The ‘only if’ part of the proof goes as follows. Assume  $P_\rho$  is a subset of  $X^d$  such that  $(\mathcal{L}(P_\rho))^\uparrow = P_\rho$ . We have to prove that  $\rho \in \mathcal{I}_d$ . Take any  $\mathbf{x} \in P_\rho$ . Since  $(\mathcal{L}(P_\rho))^\uparrow = P_\rho$  we can say that  $\mathbf{x} \in \mathbf{z}^\uparrow$  for some  $\mathbf{z} \in \mathcal{L}(P_\rho)$ . Consider now any  $\mathbf{y} \in \mathbf{x}^\uparrow$ . By the transitivity of  $\leq$  one has that  $\mathbf{y} \in \mathbf{z}^\uparrow$ . Since  $(\mathcal{L}(P_\rho))^\uparrow = P_\rho$ , we can conclude that  $\mathbf{y} \in P_\rho$ .

Q.E.D.

**Proof of Theorem 1:** It is easy to verify that when  $S_d = \mathcal{C}_d$ , then any  $\rho \in S$  satisfies MON, IND, VAN and NTR. Therefore, we will prove that when a group of identification



functions  $S_d \subseteq \Omega_d$  satisfies these four axioms then it must be equal to  $\mathcal{C}_d$ .

We will first prove the following auxiliary lemma:

**Auxiliary Lemma 1:** Whenever IND and VAN hold for some  $\rho \in S_d \subseteq \Omega_d$ , then, one has that  $|\mathbf{x}| = |\mathbf{y}| \Rightarrow \rho(\mathbf{x}) = \rho(\mathbf{y})$  for all  $\mathbf{x}, \mathbf{y} \in X^d$  and all  $\rho \in S_d$ .

Let  $\rho \in S_d$  and let  $m > 1$  be an integer for which  $|\mathbf{x}| = |\mathbf{y}|$  entails  $\rho(\mathbf{x}) = \rho(\mathbf{y})$  for all  $\mathbf{x}, \mathbf{y} \in X^d$  such that  $|\mathbf{x}| = |\mathbf{y}| < m$ . We will now prove the result also holds true whenever  $|\mathbf{x}| = |\mathbf{y}| = m$ . Let  $\mathbf{x}, \mathbf{y} \in X^d$  be two vectors with  $|\mathbf{x}| = |\mathbf{y}| = m$ . There are now two mutually exclusive cases.

*Case 1.* Assume there exists a variable  $j \in \{1, \dots, d\}$  such that  $x_j = y_j = 1$ . Consider the vectors  $\mathbf{x} - \mathbf{e}_j, \mathbf{y} - \mathbf{e}_j$ . Since  $|\mathbf{x} - \mathbf{e}_j| = |\mathbf{y} - \mathbf{e}_j| = m - 1 < m$ , the induction hypothesis holds, so  $\rho(\mathbf{x} - \mathbf{e}_j) = \rho(\mathbf{y} - \mathbf{e}_j)$ . Now, by IND one has that  $\rho(\mathbf{x}) = \rho(\mathbf{y})$ , as desired.

*Case 2.* Assume there does not exist any variable  $j \in \{1, \dots, d\}$  such that  $x_j = y_j = 1$ . Consider two different variables  $j, l \in \{1, \dots, d\}$  such that  $x_j = 0, y_j = 1$  and  $x_l = 1, y_l = 0$ . Since  $|\mathbf{x} - \mathbf{e}_l| = |\mathbf{y} - \mathbf{e}_j| = m - 1 < m$ , the induction hypothesis holds, so  $\rho(\mathbf{x} - \mathbf{e}_l) = \rho(\mathbf{y} - \mathbf{e}_j)$ . Now, by IND one has that

$$\rho(\mathbf{x}) = \rho(\mathbf{y} - \mathbf{e}_j + \mathbf{e}_l). \quad (\text{A1})$$

Observe that

$$\mathbf{e}_j = \mathbf{y} - \bigcup_{i \in \delta(\mathbf{y}) \setminus \{j\}} \mathbf{e}_i, \quad (\text{A2})$$

$$\mathbf{e}_l = \mathbf{y} - \bigcup_{i \in \delta(\mathbf{y}) \setminus \{j\}} \mathbf{e}_i - \mathbf{e}_j + \mathbf{e}_l \quad (\text{A3})$$

where  $\delta(\mathbf{y}) := \{i \in \{1, \dots, d\} | y_i = 1\}$  is the subset of variables in profile  $\mathbf{y}$  where  $y_i = 1$ .

By VAN, one has that  $\rho(\mathbf{e}_j) = \rho(\mathbf{e}_l)$ . Applying now IND several times (once per variable

$i \in \delta(\mathbf{y}) \setminus \{j\}$ ), one can conclude that

$$\rho(\mathbf{e}_j + \bigcup_{i \in \delta(\mathbf{y}) \setminus \{j\}} \mathbf{e}_i) = \rho(\mathbf{e}_l + \bigcup_{i \in \delta(\mathbf{y}) \setminus \{j\}} \mathbf{e}_i) \quad (\text{A4})$$

By (A2) and (A3), (A4) can be rewritten as

$$\rho(\mathbf{y}) = \rho(\mathbf{y} - \mathbf{e}_j + \mathbf{e}_l). \quad (\text{A5})$$

Comparing (A1) with (A5), we can conclude that  $\rho(\mathbf{x}) = \rho(\mathbf{y})$ , as desired. This proves the auxiliary lemma 1.

Q.E.D.

Take now any two vectors  $\mathbf{x}, \mathbf{y} \in X^d$  such that  $|\mathbf{x}| \geq |\mathbf{y}|$ . Define now  $\mathbf{w} \in X^d$  in such a way that  $|\mathbf{w}| = |\mathbf{y}|$  and  $\delta(\mathbf{w}) \subseteq \delta(\mathbf{x})$ . By auxiliary lemma 1, one has  $\rho(\mathbf{w}) = \rho(\mathbf{y})$ . On the other hand, by MON  $\rho(\mathbf{x}) \geq \rho(\mathbf{w})$ , so one can conclude that  $\rho(\mathbf{x}) \geq \rho(\mathbf{y})$ . This ensures that  $\rho$  is a counting measure for all  $\rho \in S_d$ .

Observe that NTR and MON imply that  $\rho(\mathbf{0}) = 0$  and  $\rho(\mathbf{1}) = 1$  (if  $\rho(\mathbf{0}) = 1$ , then  $\rho(\mathbf{x}) = 1$  for all  $\mathbf{x} \in X^d$  and if  $\rho^b(\mathbf{1}) = 0$ , then  $\rho^b(\mathbf{x}) = 0$  for all  $\mathbf{x} \in X^d$  – in both cases contradicting NTR). Lastly, by MON and NTR there must exist a  $k \in \{1, \dots, d\}$  such that  $\rho(\mathbf{x}) = 0$  whenever  $|\mathbf{x}| < k$  and  $\rho(\mathbf{x}) = 1$  whenever  $|\mathbf{x}| \geq k$ . Therefore,  $\rho \in \mathcal{C}_d$ , as desired. This proves Theorem 1.

Q.E.D.

**Statement:** If COM holds for a given  $S_d \subseteq \Omega_d$ , then IND holds as well for all  $\rho \in S_d$ ; however, the opposite is not necessarily true.

**Proof of the statement:** To verify this claim let's start assuming that COM applies for a certain set of identification functions  $S_d \subseteq \Omega_d$ . Consider now the deprivation vectors

$\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in X^d$  as in the statement of IND. Then, it is trivial to verify that  $(\mathbf{x}, \mathbf{y}')$  and  $(\mathbf{y}, \mathbf{x}')$  are equivalent societies (with  $m = 2$ ). Imposing COM, one has that  $\rho(\mathbf{x}) \geq \rho(\mathbf{y})$  implies  $\rho(\mathbf{x}') \geq \rho(\mathbf{y}')$  for all  $\rho \in S_d$ : this is precisely what IND states. On the other hand, one can find infinitely many examples of sets of identification functions satisfying IND but failing to satisfy COM. A very simple example for the case  $d = 3$  can be  $S_d = \{\rho_0\}$ , where

$$\rho_0(\mathbf{x}) = \begin{cases} 0 & \text{if } \sum_i x_i < 2 \\ 1 & \text{if } \sum_i x_i \geq 2 \end{cases}. \quad (\text{A6})$$

It is trivial to check that  $S_d = \{\rho_0\}$  satisfies IND. However, it does not satisfy COM. To verify this, consider the following pair of three-person societies  $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$  and  $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$  with  $\mathbf{x}_1 = (111), \mathbf{x}_2 = (101), \mathbf{x}_3 = (001), \mathbf{y}_1 = (101), \mathbf{y}_2 = (011), \mathbf{y}_3 = (101)$ . Clearly, both societies are equivalent (the number of individuals experiencing deprivation in each variable coincides). However, we have that  $\rho_0(\mathbf{x}_1) \leq \rho_0(\mathbf{y}_1)$  and  $\rho_0(\mathbf{x}_2) \leq \rho_0(\mathbf{y}_2)$  for all  $\rho \in S_d$  and yet  $\rho_0(\mathbf{x}_3) < \rho_0(\mathbf{y}_3)$ , thus contradicting COM.

Q.E.D.

**Proof of Theorem 2:** It is easy to verify that when  $S_d = \mathcal{W}_d$ , then any  $\rho \in S_d$  satisfies MON, COM and NTR. Therefore, we will prove that when a group of identification functions  $S_d \subseteq \Omega_d$  satisfies these three axioms then it must be equal to  $\mathcal{W}_d(k)$  for some  $k \in (0, 1]$ . Since (i)  $X^d$  is finite, (ii) each  $\rho \in S_d$  induces a complete ordering in  $X^d \times X^d$ , and (iii) COM holds for all  $\rho \in S_d$ , the hypotheses of Theorem 4.1.B in Fishburn (1970) are satisfied. Therefore, for all  $\mathbf{x}, \mathbf{y} \in X^d$  and for all  $\rho \in S_d$  one has that

$$\rho(\mathbf{x}) \leq \rho(\mathbf{y}) \Leftrightarrow \sum_{j=1}^d u_j(x_j) \leq \sum_{j=1}^d u_j(y_j) \quad (\text{A7})$$

for some real-valued functions  $u_1, \dots, u_d$  on  $\{0, 1\}$ . One can rewrite the last expression

as follows

$$\rho(\mathbf{x}) \leq \rho(\mathbf{y}) \Leftrightarrow \sum_{j=1}^d (u_j(x_j) - u_j(0)) \leq \sum_{j=1}^d (u_j(y_j) - u_j(0)) \quad (\text{A8})$$

In turn, this expression can be rewritten as

$$\rho(\mathbf{x}) \leq \rho(\mathbf{y}) \Leftrightarrow \sum_{j=1}^d \tilde{u}_j(x_j) \leq \sum_{j=1}^d \tilde{u}_j(y_j) \quad (\text{A9})$$

where  $\tilde{u}_j(x_j) := u_j(x_j) - u_j(0)$ . Clearly,  $\tilde{u}_j(0) = 0$ . If we define  $w_j := \tilde{u}_j(1)$ , (A9) can be written as

$$\rho(\mathbf{x}) \leq \rho(\mathbf{y}) \Leftrightarrow \sum_{j=1}^d w_j x_j \leq \sum_{j=1}^d w_j y_j \quad (\text{A10})$$

By MON, one must have that  $w_j \geq 0 \forall j$ . This ensures that  $\rho$  is a counting measure for all  $\rho \in S$ . By MON and NTR,  $\rho(\mathbf{0}) = 0$  and  $\rho(\mathbf{1}) = 1$ .<sup>17</sup> Lastly, by MON and NTR there must exist a real number  $q \in (0, \sum_i w_i]$  such that  $\rho(\mathbf{x}) = 0$  whenever  $\sum_{j=1}^d w_j x_j < q$  and  $\rho(\mathbf{x}) = 1$  whenever  $\sum_{j=1}^d w_j x_j \geq q$ . Defining  $a_i := w_i / \sum_i w_i$  and  $k := q / \sum_i w_i$  we have found a vector of weights  $\mathbf{a} \in \Delta_d$  and a deprivation threshold  $k \in (0, 1]$  such that  $\rho \in \mathcal{W}_d$ , as desired. This proves Theorem 2.

Q.E.D.

**Proof of Proposition 2:** It is straightforward to prove that if  $S_d \in \{\mathcal{I}_d, \mathcal{W}_d, \mathcal{C}_d\}$  then  $S_d^\psi$  satisfies NTR and MON. We will only show that  $S_d^\psi$  does not satisfy COM. For that purpose we need to prove an auxiliary lemma (see below). According to (11)  $\mathcal{C}_d^\psi$  is the set of identification functions for  $d$  variables where the unweighted counting approach is used within- and between-dimensions (given  $\psi \in \Psi_G$ ). To allow proper labelling, this set will be rewritten as  $\mathcal{C}_d^\psi(k_1, \dots, k_G; k^b)$ , where  $k_1, \dots, k_G$  denote the domain specific deprivation

<sup>17</sup>If one had that  $\rho(\mathbf{0}) = 1$ , then MON would imply that  $\rho(\mathbf{x}) = 1$  for all  $\mathbf{x} \in X^d$ , and if  $\rho(\mathbf{1}) = 0$ , then MON would imply  $\rho(\mathbf{x}) = 0$  for all  $\mathbf{x} \in X^d$  – in both cases contradicting NTR.

thresholds and  $k^b$  the between domain deprivation threshold. Within this set, define

$$\widehat{\mathcal{C}}_d^\psi := \left\{ \rho \in \mathcal{C}_d^\psi(k_1, \dots, k_G; k^b) \mid k^b < 1 \text{ and } k_{g_j} \geq 2/d_{g_j} \text{ for at least two } g_j \in \{1, \dots, G\} \right\}, \quad (\text{A11})$$

$$\widetilde{\mathcal{C}}_d^\psi := \left\{ \rho \in \mathcal{C}_d^\psi(k_1, \dots, k_G; k^b) \mid k^b = 1 \text{ and } k_{g_j} < 1 \text{ for at least two } g_j \in \{1, \dots, G\} \right\}. \quad (\text{A12})$$

The set  $\widehat{\mathcal{C}}_d^\psi$  contains identification functions where poor individuals do not have to experience deprivation in all dimensions simultaneously and in some of them they must be deprived in at least two variables. The set  $\widetilde{\mathcal{C}}_d^\psi$  contains identification functions where poor individuals have to experience deprivation in all dimensions simultaneously but where deprivation needs not to be universal within at least two of these dimensions. The sets  $\widehat{\mathcal{C}}_d^\psi$  and  $\widetilde{\mathcal{C}}_d^\psi$  are generalizations of Examples 1 and 2 to the multiple dimension context.

**Auxiliary Lemma 2.**  $\widehat{\mathcal{C}}_d^\psi \cap \mathcal{W}_d = \emptyset$  and  $\widetilde{\mathcal{C}}_d^\psi \cap \mathcal{W}_d = \emptyset$ .

**Proof of Auxiliary Lemma 2:** In both cases we follow the same strategy: if  $\rho \in \widehat{\mathcal{C}}_d^\psi$  or  $\rho \in \widetilde{\mathcal{C}}_d^\psi$  we start assuming that there *is* a weighting scheme  $\mathbf{a} \in \Delta_d$  and a deprivation threshold  $k$  such that  $\rho \in \mathcal{W}_d$  to arrive at a contradiction. Given the partition of  $D$  in  $G$  dimensions  $(D_1, \dots, D_G) \in \Psi_G$ , we will denote the elements of the weighting vector  $\mathbf{a}$  as  $a_{gv}$ , where  $g \in \{1, \dots, G\}$  indexes the member of the partition  $D_g$  to which the weight belongs and  $v \in \{1, \dots, d_g\}$  indexes the members within domain  $D_g$ . We can assume without loss of generality that *within* each domain  $D_g$  the weights are sorted in a non-ascending order, i.e.:  $a_{gv} \geq a_{g(v+1)}$  for all  $g \in \{1, \dots, G\}$  and all  $v \in \{1, \dots, d_g - 1\}$ .

We start with  $\widehat{\mathcal{C}}_d^\psi$ . Without loss of generality, we can assume that the two dimensions  $g_1, g_2 \in \{1, \dots, G\}$  with  $k_{g_1} \geq 2/d_{g_1}$  and  $k_{g_2} \geq 2/d_{g_2}$  are  $g_1 = 1$  and  $g_2 = 2$ . Since  $\rho \in \widehat{\mathcal{C}}_d^\psi$ , there exist  $\rho^b \in \mathcal{C}_G(k^b)$  and  $\rho_g^w \in \mathcal{C}_{d_g}(k_g)$  such that  $\rho(\mathbf{x}) = \rho^b(\rho_1^w(\mathbf{x}_1), \dots, \rho_G^w(\mathbf{x}_G))$ . Let  $\mathcal{L}^b \subset X^G$  be the set of *least* deprived profiles in  $(\rho^b)^{-1}(1)$  (i.e. if  $\mathbf{x} \in \mathcal{L}^b$  and  $\mathbf{y} \in X^G$  is

such that  $\mathbf{y} < \mathbf{x}$ , then  $\mathbf{y} \in (\rho^b)^{-1}(0)$ . Consider  $\mathbf{u}, \mathbf{u}' \in \mathcal{L}^b$ . Without loss of generality we will write them as  $\mathbf{u} = (1 \ 0 \ u_3 \ \dots \ u_G)$ ,  $\mathbf{u}' = (0 \ 1 \ u_3 \ \dots \ u_G)$ . By definition, the following inequalities must hold:

$$a_{11} + a_{12} + \sum_{v=3}^{v=m_1} a_{1v} + \sum_{v=3}^{v=m_2} a_{2v} + \sum_{g \in \delta(u_3 \dots u_G)} \sum_{v=1}^{v=m_g} a_{gv} \geq k \quad (\text{A13})$$

$$a_{21} + a_{22} + \sum_{v=3}^{v=m_1} a_{1v} + \sum_{v=3}^{v=m_2} a_{2v} + \sum_{g \in \delta(u_3 \dots u_G)} \sum_{v=1}^{v=m_g} a_{gv} \geq k \quad (\text{A14})$$

where  $m_g := \lfloor k_g d_g \rfloor$  and  $\delta(u_3 \dots u_G) := \{g | u_g = 1\}$  is the subset of elements in vector  $(u_3 \dots u_G)$  where  $u_g = 1$ . Consider now a third vector  $\mathbf{u}'' = (0 \ 0 \ u_3 \ \dots \ u_G)$ . Since  $\mathbf{u}, \mathbf{u}' \in \mathcal{L}^b$ , one has that  $\mathbf{u}'' \in (\rho^b)^{-1}(0)$ . This implies that the following inequalities must hold.

$$a_{11} + \sum_{v=3}^{v=m_1} a_{1v} + a_{21} + \sum_{v=3}^{v=m_2} a_{2v} + \sum_{g \in \delta(u_3 \dots u_G)} \sum_{v=1}^{v=m_g} a_{gv} < k \quad (\text{A15})$$

$$a_{11} + \sum_{v=3}^{v=m_1} a_{1v} + a_{22} + \sum_{v=3}^{v=m_2} a_{2v} + \sum_{g \in \delta(u_3 \dots u_G)} \sum_{v=1}^{v=m_g} a_{gv} < k \quad (\text{A16})$$

$$a_{12} + \sum_{v=3}^{v=m_1} a_{1v} + a_{21} + \sum_{v=3}^{v=m_2} a_{2v} + \sum_{g \in \delta(u_3 \dots u_G)} \sum_{v=1}^{v=m_g} a_{gv} < k \quad (\text{A17})$$

$$a_{12} + \sum_{v=3}^{v=m_1} a_{1v} + a_{22} + \sum_{v=3}^{v=m_2} a_{2v} + \sum_{g \in \delta(u_3 \dots u_G)} \sum_{v=1}^{v=m_g} a_{gv} < k \quad (\text{A18})$$

If one defines

$$k' := k - \left( \sum_{v=3}^{v=m_1} a_{1v} + \sum_{v=3}^{v=m_2} a_{2v} + \sum_{g \in \delta(u_3 \dots u_G)} \sum_{v=1}^{v=m_g} a_{gv} \right) \quad (\text{A19})$$

the inequalities (A13)-(A18) can be rewritten as

$$\left\{ \begin{array}{l} a_{11} + a_{12} \geq k', \quad a_{11} + a_{21} < k', \quad a_{12} + a_{21} < k', \\ a_{21} + a_{22} \geq k', \quad a_{11} + a_{22} < k', \quad a_{12} + a_{22} < k'. \end{array} \right\} \quad (\text{A20})$$

It is trivial to show that the inequalities system shown in (A20) does not have feasible solutions. In the first inequality of the system, either  $a_{11}$  or  $a_{12}$  must be greater or equal

than  $k'/2$ . The same goes for  $a_{21}, a_{22}$  in the second inequality of the system: at least one of them must be greater or equal than  $k'/2$ . Picking the largest elements between  $a_{11}, a_{12}$  and  $a_{21}, a_{22}$  and adding them up results in a number that is greater or equal than  $k'$ , therefore contradicting at least one of the four last inequalities of the system. We have reached the contradiction we were looking for.

Let us now consider case  $\tilde{\mathcal{C}}_d^\psi$ . Without loss of generality, we can assume that the two dimensions  $g_1, g_2 \in \{1, \dots, G\}$  with  $k_{g_1} < 1$  and  $k_{g_2} < 1$  are  $g_1 = 1$  and  $g_2 = 2$ . Since  $\rho \in \tilde{\mathcal{C}}_d^\psi$ , there exist  $\rho^b \in \mathcal{C}_G(1)$  and  $\rho_g^w \in \mathcal{C}_{d_g}(k_g)$  such that  $\rho(\mathbf{x}) = \rho^b(\rho_1^w(\mathbf{x}_1), \dots, \rho_G^w(\mathbf{x}_G))$ . By definition, the following inequalities must hold:

$$a_{11} + \sum_{v=3}^{v=m_1+1} a_{1v} + a_{21} + \sum_{v=3}^{v=m_2+1} a_{2v} + \sum_{g=3}^{g=G} \sum_{v=1}^{v=m_g} a_{gv} \geq k \quad (\text{A21})$$

$$a_{11} + \sum_{v=3}^{v=m_1+1} a_{1v} + a_{22} + \sum_{v=3}^{v=m_2+1} a_{2v} + \sum_{g=3}^{g=G} \sum_{v=1}^{v=m_g} a_{gv} \geq k \quad (\text{A22})$$

$$a_{12} + \sum_{v=3}^{v=m_1+1} a_{1v} + a_{21} + \sum_{v=3}^{v=m_2+1} a_{2v} + \sum_{g=3}^{g=G} \sum_{v=1}^{v=m_g} a_{gv} \geq k \quad (\text{A23})$$

$$a_{12} + \sum_{v=3}^{v=m_1+1} a_{1v} + a_{22} + \sum_{v=3}^{v=m_2+1} a_{2v} + \sum_{g=3}^{g=G} \sum_{v=1}^{v=m_g} a_{gv} \geq k \quad (\text{A24})$$

where  $m_g := \lfloor k_g d_g \rfloor$ . Consider now the following  $G$ -dimensional binary vectors:  $\mathbf{v} = (1 \ 0 \ 1 \dots 1)$ ,  $\mathbf{v}' = (0 \ 1 \ 1 \dots 1)$ . Since  $\mathbf{v}, \mathbf{v}' \in (\rho^b)^{-1}(0)$ , the following inequalities must hold:

$$a_{11} + a_{12} + \sum_{v=3}^{v=m_1+1} a_{1v} + \sum_{v=3}^{v=m_2+1} a_{2v} + \sum_{g=3}^{g=G} \sum_{v=1}^{v=m_g} a_{gv} < k \quad (\text{A25})$$

$$a_{21} + a_{22} + \sum_{v=3}^{v=m_1+1} a_{1v} + \sum_{v=3}^{v=m_2+1} a_{2v} + \sum_{g=3}^{g=G} \sum_{v=1}^{v=m_g} a_{gv} < k \quad (\text{A26})$$

Defining

$$k' := k - \left( \sum_{v=3}^{v=m_1+1} a_{1v} + \sum_{v=3}^{v=m_2+1} a_{2v} + \sum_{g=3}^{g=G} \sum_{v=1}^{v=m_g} a_{gv} \right) \quad (\text{A27})$$

the inequalities (A21)-(A26) can be rewritten as

$$\left\{ \begin{array}{l} a_{11} + a_{21} \geq k', \quad a_{12} + a_{21} \geq k', \quad a_{11} + a_{12} < k', \\ a_{11} + a_{22} \geq k', \quad a_{12} + a_{22} \geq k', \quad a_{21} + a_{22} < k'. \end{array} \right\} \quad (\text{A28})$$

Again, it is trivial to prove that the inequalities system shown in (A28) does not have feasible solutions. In the second to last inequality of the system, either  $a_{11}$  or  $a_{12}$  must be smaller than  $k'/2$ . The same goes for  $a_{21}, a_{22}$  in the last inequality of the system: at least one of them must be smaller than  $k'/2$ . Picking the smallest elements between  $a_{11}, a_{12}$  and  $a_{21}, a_{22}$  and adding them up results in a number that is smaller than  $k'$ , therefore contradicting at least one of the four first inequalities of the system. We have reached the contradiction we were looking for. This proves auxiliary lemma 2.

Q.E.D.

The proof of proposition 1 is now almost immediate. According to auxiliary lemma 2,  $\mathcal{C}_d^\psi \setminus \mathcal{W}_d \neq \emptyset$  (essentially, it is only when the union or intersection approaches are used both within and across dimensions –i.e. either  $k_g = 1/d_g \forall g, k^b = 1/G$  or  $k_g = 1 \forall g, k^b = 1$  – that the counting approach  $\mathcal{W}_d$  is able to generate some identification functions included in  $\mathcal{C}_d^\psi$ ). Since COM uniquely characterizes  $\mathcal{W}_d$  (see Theorem 2),  $\mathcal{C}_d^\psi$  cannot satisfy that axiom as well. Finally, since  $\mathcal{I}_d^\psi \supset \mathcal{W}_d^\psi \supset \mathcal{C}_d^\psi$  and  $\mathcal{C}_d^\psi$  does not satisfy COM, neither  $\mathcal{W}_d^\psi$  nor  $\mathcal{I}_d^\psi$  can satisfy COM. This proves Proposition 2.

Q.E.D.

**Proof of Theorem 3:** It is clear that the multivariate poverty index shown in equation (12) satisfies SDC, MOA, NRM, FOC, SEP, HMG and CON. We are going to prove the opposite implication. Since  $\mathbf{f} = \{f_d\}_{d \in \mathbb{N}}$  satisfies SDC, it can be written as

$$f_d(\Gamma) = \frac{1}{n} \sum_{i=1}^n p(\gamma_i) \quad (\text{A29})$$



for some function  $p : [0, 1]^d \rightarrow \mathbb{R}$ . Clearly, for any  $\gamma \in [0, 1]^d$ ,  $p(\gamma) = f_d([\gamma])$  for some deprivation matrix  $[\gamma] \in \mathcal{G}_{n \times d}^S$  where all rows are equal to  $\gamma$ . Therefore, since  $f_d$  satisfies SDC, CON, HMG, SEP and MOA,  $p$  will satisfy them too. It can be shown that MOA implies minimal increasingness and strict essentiality (see Blackorby and Donaldson 1982:251). Moreover, the domain of  $p$  is  $[0, 1]^d$ , which is connected and topologically separable. In an analogous way to Blackorby and Donaldson (1982: 252), based on Gorman (1968:369) and Blackorby, Primont and Russel (1978:127) it can be shown that  $p$  is additively separable and can be written as

$$p(\gamma) = p^* \left( \sum_{j=1}^d p_j(\gamma_j) \right) \quad (\text{A30})$$

where  $p^*$  and  $p_j, j \in \{1, \dots, d\}$  are continuous real-valued functions and  $p^*$  is increasing. By HMG, one has that

$$p(\lambda\gamma) = \lambda p(\gamma) \quad (\text{A31})$$

for any  $\lambda \in (0, 1]$ . Using equation (A30), for each  $l \in \{1, \dots, d\}$  we can define the functions

$$h_l(\gamma_l) := p(0, \dots, 0, \gamma_l, 0, \dots, 0) = p^* \left( p_l(\gamma_l) + \sum_{j \neq l} p_j(0) \right). \quad (\text{A32})$$

Since  $p$  is linearly homogeneous on  $\gamma$ , so are the functions  $h_l(\gamma_l)$ . Therefore

$$h_l(\lambda\gamma_l) = \lambda h_l(\gamma_l) \quad (\text{A33})$$

for any  $\lambda \in (0, 1]$  and for all  $l \in \{1, \dots, d\}$ . As a consequence, there exist constants  $c_l$  such that

$$h_l(\gamma_l) = \gamma_l h_l(1) = c_l \gamma_l \quad (\text{A34})$$

Plugging equations (A32) and (A34) we have that

$$c_l \gamma_l = p^* \left( p_l(\gamma_l) + \sum_{j \neq l} p_j(0) \right). \quad (\text{A35})$$

Hence

$$p_l(\gamma_l) = p^{*-1}(c_l \gamma_l) - \sum_{j \neq l} p_j(0) \quad (\text{A36})$$

Substituting equation (A36) in equation (A30), one has that

$$p(\boldsymbol{\gamma}) = p^* \left( \sum_{j=1}^d \left[ p^{*-1}(c_j \gamma_j) - \sum_{l \neq j} p_l(0) \right] \right) = p^* \left( \sum_{j=1}^d p^{*-1}(c_j \gamma_j) + \varsigma \right) \quad (\text{A37})$$

for some constants  $c_j, \varsigma$  and a continuous increasing function  $p^*$ . Equation (A37) is essentially the same as equation (34) in Blackorby and Donaldson (1982:260). Therefore, following those authors—who in turn draw from Eichhorn (1978:32-34)—it can be proven that  $p^{*-1} =: f$  must satisfy the following functional equation

$$f(\lambda u) = \alpha(\lambda) f(u) + b(\lambda) \quad (\text{A38})$$

Without the domain restrictions on  $\lambda$  and  $u$ , the solutions to equation (A38) are well-known (Aczel et al 1986). It is straightforward to show that the solution for equation (A38) on the present restricted domain is

$$f(u) = \left\{ \begin{array}{l} au^\theta + b \\ c \ln(u) + d \end{array} \right\} \quad (\text{A39})$$

for some parameters  $a, b, c, d, \theta$  (with  $\theta \neq 0$ ). Since continuity of  $f$  at 0 precludes the logarithmic solution, the general solution of equation (A37) can be written as

$$p(\boldsymbol{\gamma}) = p^* \left( \sum_{j=1}^d a (c_j \gamma_j)^\theta + db + \varsigma \right) = \left( \sum_{j=1}^d (c_j \gamma_j)^\theta + \frac{b(d-1) + \varsigma}{a} \right)^{1/\theta}. \quad (\text{A40})$$

Since  $p$  is linearly homogeneous on  $\boldsymbol{\gamma}$ , one must have that  $(b(d-1) + \varsigma)/a = 0$ . Therefore, equation (A30) can be rewritten as

$$p(\boldsymbol{\gamma}) = \psi \left( \left( \sum_{j=1}^d (c_j \gamma_j)^\theta \right)^{1/\theta} \right) \quad (\text{A41})$$

for some continuous increasing function  $\psi$ . Since  $p$  satisfies HMG one has

$$\lambda\psi\left(\left(\sum_{j=1}^d(c_j\gamma_j)^\theta\right)^{1/\theta}\right)=\lambda p(\boldsymbol{\gamma})=p(\lambda\boldsymbol{\gamma})=\psi\left(\left(\sum_{j=1}^d(c_j\lambda\gamma_j)^\theta\right)^{1/\theta}\right)=\psi\left(\lambda\left(\sum_{j=1}^d(c_j\gamma_j)^\theta\right)^{1/\theta}\right). \quad (\text{A42})$$

Equation (A42) implies that

$$\lambda\psi(x)=\psi(\lambda x) \quad (\text{A43})$$

for all  $x \geq 0, \lambda \in (0, 1]$ . As a consequence, there exists a constant  $q$  such that

$$\psi(x)=x\psi(1)=qx \quad (\text{A44})$$

Imposing NRM and MOA and rewriting accordingly, one obtains the desired functional form.

Q.E.D.

**Proof of Theorem 4:** Given its similarity with Theorem 3, we will simply present a brief sketch of the proof (the complete proof is available upon request). The sufficiency part of the theorem is clear, so we focus on the reverse implication. Axioms SDC, FOC, MOA and NRM imply that  $f_d(\Gamma)$  can be written as

$$\frac{1}{n}\sum_{i \in Q(P_\rho)}\Psi(\gamma_{i1}, \dots, \gamma_{id}) \quad (\text{A45})$$

for some increasing function  $\Psi : [0, 1]^d \rightarrow [0, 1]$  with  $\Psi(\mathbf{1}) = 1$  and  $\Psi(\mathbf{0}) = 0$ . Imposing CON, HMG, MOA and WDS, it turns out that  $\Psi(\gamma_{i1}, \dots, \gamma_{id}) = \Phi(P_1(\boldsymbol{\gamma}_1), \dots, P_G(\boldsymbol{\gamma}_G))$ , where

$$P_g(\boldsymbol{\gamma}_g)=\left[\sum_{v=1}^{d_g}w_{gv}(\gamma_{gv})^{\theta_g}\right]^{1/\theta_g} \quad (\text{A46})$$

for some  $\theta_g > 0, w_{gv} > 0$ . Lastly, CON, HMG, MOA and BDS imply that

$$\Phi(P_1, \dots, P_G)=\left(\sum_{g=1}^{G}a_g.P_g^\theta\right)^{1/\theta} \quad (\text{A47})$$

for some  $\theta > 0, a_{g.} > 0$ , so we obtain the desired functional form.

Q.E.D.

**Proof of Proposition 3:** Let

$$\pi_{\theta} := \left( \sum_{g=1}^{g=G} a_{g.} \left[ \sum_{v=1}^{d_g} w_{gv}(\gamma_{gv})^{\theta_g} \right]^{\theta/\theta_g} \right)^{1/\theta} \quad (\text{A48})$$

be the individual level poverty function corresponding to (16). Therefore, one has that

$$\frac{\partial \pi_{\theta}}{\partial \gamma_{gv}} = \left( \sum_{g=1}^{g=G} a_{g.} \left[ \sum_{v=1}^{d_g} w_{gv}(\gamma_{gv})^{\theta_g} \right]^{\theta/\theta_g} \right)^{\frac{1}{\theta}-1} \left[ \sum_{v=1}^{d_g} w_{gv}(\gamma_{gv})^{\theta_g} \right]^{\frac{\theta}{\theta_g}-1} w_{gv}(\gamma_{gv})^{\theta_g-1} \quad (\text{A49})$$

After several algebraic manipulations it is easy to show that

$$\frac{\partial^2 \pi_{\theta}}{\partial \gamma_{gv} \partial \gamma_{gu}} \equiv \left[ (1 - \theta) \left( \sum_{v=1}^{d_g} w_{gv}(\gamma_{gv})^{\theta_g} \right)^{\theta/\theta_g} + (\theta - \theta_g) \left( \sum_{g=1}^{g=G} \left[ \sum_{v=1}^{d_g} w_{gv}(\gamma_{gv})^{\theta_g} \right]^{\theta/\theta_g} \right) \right] \quad (\text{A50})$$

The last expression can be rearranged and written as follows:

$$\frac{\partial^2 \pi_{\theta}}{\partial \gamma_{gv} \partial \gamma_{gu}} \equiv \left[ (1 - \theta_g) \left( \sum_{v=1}^{d_g} w_{gv}(\gamma_{gv})^{\theta_g} \right)^{\theta/\theta_g} + (\theta - \theta_g) \left( \sum_{h=1, h \neq g}^{h=G} \left[ \sum_{v=1}^{d_h} w_{hv}(\gamma_{hv})^{\theta_h} \right]^{\theta/\theta_h} \right) \right] \quad (\text{A51})$$

Therefore, one can basically say that

$$\frac{\partial^2 \pi_{\theta}}{\partial \gamma_{gv} \partial \gamma_{gu}} \equiv A(1 - \theta_g) + B(\theta - \theta_g) \quad (\text{A52})$$

for some real constants  $A, B > 0$ . Hence, whenever  $\theta_g < \min\{1, \theta\}$ ,  $(\partial^2 \pi_{\theta}) / (\partial \gamma_{gv} \partial \gamma_{gu}) > 0$ , so the attributes  $u, v$  belonging to the same dimension are complements. On the other hand, whenever  $\theta_g > \max\{1, \theta\}$ ,  $(\partial^2 \pi_{\theta}) / (\partial \gamma_{gv} \partial \gamma_{gu}) < 0$ , so the attributes  $u, v$  belonging to the same dimension are substitutes. This proves part (i). For part (ii), we need to compute  $(\partial^2 \pi_{\theta}) / (\partial \gamma_{gv} \partial \gamma_{hu})$ . After algebraic manipulations it can be shown that

$$\frac{\partial^2 \pi_{\theta}}{\partial \gamma_{gv} \partial \gamma_{hu}} \equiv (1 - \theta) \left( \sum_{g=1}^{g=G} a_{g.} \left[ \sum_{v=1}^{d_g} w_{gv}(\gamma_{gv})^{\theta_g} \right]^{\theta/\theta_g} \right)^{\frac{1}{\theta}-2} \left[ \sum_{u=1}^{d_h} w_{hu}(\gamma_{hu})^{\theta_g} \right]^{\frac{\theta}{\theta_h}} w_{hu}(\gamma_{hu})^{\theta_h-1} \quad (\text{A53})$$

From the previous equation we can say that

$$\frac{\partial^2 \pi_{\theta}}{\partial \gamma_{gv} \partial \gamma_{hu}} \equiv C(1 - \theta) \quad (\text{A54})$$

for some real constant  $C > 0$ . Therefore, whenever  $\theta < 1$ ,  $(\partial^2 \pi_{\theta}) / (\partial \gamma_{gv} \partial \gamma_{hu}) > 0$ , so the attributes  $u, v$  belonging to different dimensions are complements. Analogously, when  $\theta > 1$ ,  $(\partial^2 \pi_{\theta}) / (\partial \gamma_{gv} \partial \gamma_{hu}) < 0$ , so the attributes  $u, v$  belonging to different dimensions are substitutes. This proves part (ii).

Q.E.D.

## REFERENCES

- Aaberge, R. and Brandolini, A. (2015), “Multidimensional poverty and inequality”, in Atkinson, A. and Bourguignon, F. (Eds.) *Handbook of Income Distribution*, Volume 2A, Chapter 3, North Holland.
- Aaberge, R., Peluso, E. and Sigstad, H. (2015), “The dual approach for measuring multidimensional deprivation and poverty”, Discussion Papers 820, Statistics Norway, Research Department.
- Aczél, J., Roberts, F. and Rosenbaum, Z. (1986), “On scientific laws without dimensional constants”, *Journal of Mathematical Analysis and Applications* 119: 389-416.
- Alkire, S. and Foster, J. (2011), “Counting and multidimensional poverty measurement”, *Journal of Public Economics* 95 (7-8): 476-487.
- Alkire, S. Foster, J., Seth, S., Santos, M., Roche, J. and Ballón, P. (2015). *Multidimensional poverty measurement and analysis*. Oxford: Oxford University Press.
- Anand, S. and Sen, A. (1997), “Concepts of human development and poverty: A multidimensional perspective”, UNDP, New York.
- Atkinson, A. (2003), “Multidimensional deprivation: contrasting social welfare and counting approaches”, *Journal of Economic Inequality* 1:51-65.
- Atkinson, A. and Bourguignon, F. (1982), “The comparison of multi-dimensioned distributions of economic status”, *Review of Economic Studies* 49:183-201.
- Blackorby, C. and Donaldson, D. (1982), “Ratio-scale and translation-scale full interpersonal comparability without domain restrictions: admissible social-evaluation functions”, *International Economic Review* 23(2):249-268.
- Blackorby, C., Primont, D. and Russell, R. (1978), *Duality, separability, and functional structure: Theory and economic applications*, North-Holland, New York / Amsterdam.
- Bourguignon, F. and Chakravarty, S. (2003), “The measurement of multidimensional poverty”, *Journal of Economic Inequality* 1:25-49.
- Chakravarty, S., Mukherjee, D., Renade, R. (1998), “On the family of subgroup and factor decomposable measures of multidimensional poverty”, *Research on Economic Inequality* 8:175–194.
- Chakravarty, S. and D’Ambrosio, C. (2006), “The measurement of social exclusion”, *Review of Income and Wealth* 52(3):377-398.
- Chakravarty, S. (2009), *Inequality, Polarization and Poverty. Advances in Distributional Analysis*. Ramat Gan: Springer.

- Datt, G. (2017), “Distribution-sensitive multidimensional poverty measures”, *The World Bank Economic Review*, lhx017, <https://doi.org/10.1093/wber/lhx017>.
- Fishburn, P. (1970), *Utility theory for decision making*. Wiley.
- Gorman, W. (1968), “The structure of utility functions”, *Review of Economic Studies* 32:369–390.
- Kannai, Y. (1980), “The ALEP definition of complementarity and least concave utility functions”, *Journal of Economic Theory* 22: 115-117.
- Lasso de la Vega, C. (2010), “Counting poverty orderings and deprivation curves”, *Research on Economic Inequality* 18:153-172.
- Pattanaik, P. and Xu, Y. (1990), “On ranking opportunity sets in terms of freedom of choice”, *Recherches Économiques de Louvain / Louvain Economic Review* 56:383-390.
- Pattanaik, P. and Xu, Y. (2018), “On measuring multidimensional deprivation”, *Journal of Economic Literature*, 56(2):657-672.
- Permanyer, I. (2014), “Assessing individuals’ deprivation in a multidimensional framework”, *Journal of Development Economics* 109:1-16.
- Ravallion, M. (2011), “On multidimensional indices of poverty ”, *Journal of Economic Inequality* 9:235-248.
- Ravallion, M. (2012), “Troubling trade-offs in the Human Development Index”, *Journal of Development Economics* 99:201-209.
- Sen, A. (1976), “Poverty: An ordinal approach to measurement”, *Econometrica* 42 (2): 219-231.
- Sen, A. (1985), *Commodities and Capabilities*. Elsevier, Amsterdam; New York. North-Holland.
- Sen, A. (1987), “The Standard of Living”, in Hawthorn, G. (Ed.), *The Standard of Living*. Cambridge University Press, Cambridge: 1–38.
- Sen, A. (1992), *Inequality Re-examined*. Clarendon Press, Oxford.
- Sen, A. (1993), “Capability and Well-Being”, in Nussbaum, M., Sen, A. (Eds.), *Quality of Life*. Clarendon Press, Oxford: 30-53.
- Silber, J. and Yalonetzky, G. (2014), “Measuring multidimensional deprivation with dichotomized and ordinal variables”, in Betti, G. and Lemmi, A. (Eds.) *Poverty and Social Exclusion. New methods of analysis*, Chapter 2, Routledge.

Thorbecke, E. (2007), “Multidimensional poverty: conceptual and measurement issues”, in Kakwani, N. and Silber, J. (Eds.), *The many dimensions of poverty*, Chapter 1, Palgrave Macmillan.

Tsui, K. (2002), “Multidimensional poverty indices”, *Social Choice and Welfare* 19:69-93.

Yalonetzky, G. (2014), “Conditions for the most robust multidimensional poverty comparisons using counting measures and ordinal variables”, *Social Choice and Welfare* 43:773-807.