



# OPHI WORKING PAPER NO. 83

## Multidimensional Poverty Measurement and Analysis: Chapter 2 – The Framework

Sabina Alkire\*, James E. Foster\*\*, Suman Seth\*\*\*, Maria  
Emma Santos\*\*\*, Jose M. Roche\*\*\*\* and Paola Ballon\*\*\*\*\*

January 2015

### Abstract

This working paper introduces the notation and basic concepts that are used throughout the *OPHI Working Papers* 82-91. The Paper has five sections. First we review unidimensional poverty measurement with particular attention to the well-known Foster-Greer-Thorbecke measures of income poverty as many methods presented in *OPHI Working Paper* 84 (Chapter 3 – Overview of Methods for Multidimensional Poverty Assessment) as well as the measure presented in *OPHI*

\* Sabina Alkire: Oxford Poverty & Human Development Initiative, Oxford Department of International Development, University of Oxford, 3 Mansfield Road, Oxford OX1 3TB, UK, +44-1865-271915, [sabina.alkire@qeh.ox.ac.uk](mailto:sabina.alkire@qeh.ox.ac.uk)

\*\* James E. Foster: Professor of Economics and International Affairs, Elliott School of International Affairs, 1957 E Street, NW, [fosterje@gwu.edu](mailto:fosterje@gwu.edu).

\*\*\* Suman Seth: Oxford Poverty & Human Development Initiative (OPHI), Queen Elizabeth House (QEH), Department of International Development, University of Oxford, UK, +44 1865 618643, [suman.seth@qeh.ox.ac.uk](mailto:suman.seth@qeh.ox.ac.uk).

\*\*\*\* Maria Emma Santos: Instituto de Investigaciones Económicas y Sociales del Sur (IIES), Departamento de Economía, Universidad Nacional del Sur (UNS) - Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), 12 de Octubre 1198, 7 Piso, 8000 Bahía Blanca, Argentina. Oxford Poverty and Human Development Initiative, University of Oxford. [msantos@uns.edu.ar](mailto:msantos@uns.edu.ar); [maria.santos@qeh.ox.ac.uk](mailto:maria.santos@qeh.ox.ac.uk).

\*\*\*\*\* Jose Manuel Roche: Save the Children UK, 1 St John's Lane, London EC1M 4AR, [j.roche@savethechildren.org.uk](mailto:j.roche@savethechildren.org.uk).

\*\*\*\*\* Paola Ballon: Assistant Professor of Econometrics, Department of Economics, Universidad del Pacifico, Lima, Peru; Research Associate, OPHI, Department of International Development, Oxford University, Oxford, U.K, [pm.ballonf@up.edu.pe](mailto:pm.ballonf@up.edu.pe).

This study has been prepared within the OPHI theme on multidimensional measurement and missing dimensions.

OPHI gratefully acknowledges support from the German Federal Ministry for Economic Cooperation and Development (BMZ), Praus, national offices of the United Nations Development Programme (UNDP), national governments, the International Food Policy Research Institute (IFPRI), and private benefactors. For their past support OPHI acknowledges the UK Economic and Social Research Council (ESRC)/(DFID) Joint Scheme, the Robertson Foundation, the John Fell Oxford University Press (OUP) Research Fund, the Human Development Report Office (HDRO/UNDP), the International Development Research Council (IDRC) of Canada, the Canadian International Development Agency (CIDA), the UK Department of International Development (DFID), and AusAID.

*Working Papers* 86-90 (Chapters 5-9) are based on these measures. The second section introduces the notation and basic concepts for multidimensional poverty measurement that are used in subsequent chapters. Third we define indicators' scales of measurement, and fourth, address issues of comparability across people and dimensions. The fifth section systematically explains different properties that have been proposed in axiomatic approaches to multidimensional poverty measurement, which enable the analyst to understand the ethical principles embodied in a measure and to be aware of the direction of change they will exhibit under certain transformations.

**Keywords:** Foster-Greer-Thorbecke measures, scales of measurement, ordinal data properties of poverty measurement, axiomatic methods, joint distribution, identification, aggregation.

**JEL classification:** D63, I32.

### Acknowledgements

We received very helpful comments, corrections, improvements, and suggestions from many across the years. We are also grateful for direct comments on this working paper from Tony Atkinson, Gordon Anderson, Indranil Dutta, Kaushik Basu, Satya Chakravarty, Natalie Quinn, Prasanta Pattanaik, Sudhir Anand, Francois Bourguignon, Gaston Yalonetzky and Koen Decancq.

**Citation:** Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M., and Ballon, P. (2015). *Multidimensional Poverty Measurement and Analysis*, Oxford: Oxford University Press, ch. 2.

*The Oxford Poverty and Human Development Initiative (OPHI) is a research centre within the Oxford Department of International Development, Queen Elizabeth House, at the University of Oxford. Led by Sabina Alkire, OPHI aspires to build and advance a more systematic methodological and economic framework for reducing multidimensional poverty, grounded in people's experiences and values.*

The copyright holder of this publication is Oxford Poverty and Human Development Initiative (OPHI). This publication will be published on OPHI website and will be archived in Oxford University Research Archive (ORA) as a Green Open Access publication. The author may submit this paper to other journals.

This publication is copyright, however it may be reproduced without fee for teaching or non-profit purposes, but not for resale. Formal permission is required for all such uses, and will normally be granted immediately. For copying in any other circumstances, or for re-use in other publications, or for translation or adaptation, prior written permission must be obtained from OPHI and may be subject to a fee.

Oxford Poverty & Human Development Initiative (OPHI)  
 Oxford Department of International Development  
 Queen Elizabeth House (QEH), University of Oxford  
 3 Mansfield Road, Oxford OX1 3TB, UK  
 Tel. +44 (0)1865 271915 Fax +44 (0)1865 281801  
 ophi@qeh.ox.ac.uk http://www.ophi.org.uk

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by OPHI or the University of Oxford, nor by the sponsors, of any of the views expressed.

## 2.0 Introduction

This chapter introduces the notation and basic concepts that will be used throughout the book. The chapter is organized into four sections. Section 2.1 starts with a review of unidimensional poverty measurement with particular attention to the well-known FGT measures (Foster, Greer, and Thorbecke 1984) because many methods presented in Chapter 3, as well as the Alkire and Foster (2007, 2011a) measures presented in Chapter 5, are based on FGT indices. Section 2.2 introduces the notation and basic concepts for multidimensional poverty measurement that will be used in subsequent chapters. Section 2.3 delves into the issue of indicators' scales of measurement, an aspect often overlooked when discussing methods for multidimensional analysis and which is central to this book. Section 2.4 addresses comparability across people and dimensions. Finally Section 2.5 presents in a detailed form the different properties that have been proposed in axiomatic approaches to multidimensional poverty measurement. Such properties enable the analyst to understand the ethical principles embodied in a measure and to be aware of the direction of change they will exhibit under certain transformations.

### 2.1 Review of Unidimensional Measurement and FGT Measures

The measurement of multidimensional poverty builds upon a long tradition of unidimensional poverty measurement. Because both approaches are technically closely linked, the measurement of poverty in a unidimensional way can be seen as a special case of multidimensional poverty measurement. This section introduces the basic concepts of unidimensional poverty measurement using the lens of the multidimensional framework, so serves as a springboard for the later work.

The measurement of poverty requires a reference population, such as all people in a country. We refer to the reference population under study as a *society*. We assume that any society consists of at least one observation or unit of analysis. This unit varies depending on the measurement exercise. For example, the unit of analysis is a child if one is measuring child poverty, it is an elderly person if one is measuring poverty among the elderly, and it is a person or—sometimes due to data constraints—the household for measures covering the whole population. For simplicity, unless otherwise indicated, we refer to the unit of analysis within a society as a *person* (Chapter 6 and

Chapter 7). We denote the number of person(s) within a society by  $n$ , such that  $n$  is in  $\mathbb{N}$  or  $n \in \mathbb{N}$ , where  $\mathbb{N}$  is the set of positive integers. Note that unless otherwise specified,  $n$  refers to the total population of a society and not a sample of it.

Assume that poverty is to be assessed using  $d$  number of dimensions, such that  $d \in \mathbb{N}$ . We refer to the performance of a person in a dimension as an achievement in a very general way, and we assume that achievements in each dimension can be represented by a non-negative real valued indicator. We denote the achievement of person  $i$  in dimension  $j$  by  $x_{ij} \in \mathbb{R}_+$  for all  $i = 1, \dots, n$  and  $j = 1, \dots, d$ , where  $\mathbb{R}_+$  is the set of non-negative real numbers, which is a proper subset of the set of real numbers  $\mathbb{R}$ .<sup>1</sup> Subsequently, we denote the set of strictly positive real numbers by  $\mathbb{R}_{++}$ .

Throughout this book, we allow the population size of a society to vary, which allows comparisons of societies with different populations. When we seek to permit comparability of poverty estimates across different populations, we assume  $d$  to denote a fixed set (and number) of dimensions.<sup>2</sup> The achievements of all persons within a society are denoted by an  $n \times d$ -dimensional achievement matrix  $X$  which looks as follows:

$$X = \begin{array}{c} \text{Dimensions} \\ \left[ \begin{array}{ccc} x_{11} & \cdots & x_{1d} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nd} \end{array} \right] \\ \text{People} \end{array}$$

We denote the set of all possible matrices of size  $n \times d$  by  $\mathcal{X}_n \in \mathbb{R}_+^{n \times d}$  and the set of all possible achievement matrices by  $\mathcal{X}$ , such that  $\mathcal{X} = \bigcup_n \mathcal{X}_n$ . If  $X \in \mathcal{X}_n$ , then matrix  $X$  contains achievements for  $n$  persons in  $d$  dimensions. Unless specified otherwise, whenever we refer to matrix  $X$ , we assume  $X \in \mathcal{X}$ . The achievements of any person  $i$  in all  $d$  dimensions, which is row  $i$  of matrix  $X$ , are represented by the  $d$ -dimensional vector  $x_i$  for all  $i = 1, \dots, n$ . The achievements in any dimension  $j$  for all  $n$  persons, which is column  $j$  of matrix  $X$ , are represented by the  $n$ -dimensional vector  $x_{\cdot j}$  for all  $j = 1, \dots, d$ .

---

<sup>1</sup> Empirical applications may encounter negative or zero income values, which require special treatment for certain poverty measures to be implemented.

<sup>2</sup> In practical implementations of the unidimensional method, a fixed set and number of dimensions is rarely obtained. Survey-based consumption items or income sources often differ in number and content.

In the unidimensional context, the  $d$  dimensions considered in matrix  $X \in \mathcal{X}$ —which are typically assumed to be cardinal—can be meaningfully combined into a well-defined overall achievement or resource variable for each person  $i$ , which is denoted by  $x_i$ . One possibility, from a welfarist approach, would be to construct each person’s welfare from her vector of achievements using a utility function  $x_i = u(x_{i1}, \dots, x_{id})$ .<sup>3</sup> Another possibility is that each dimension  $j$  refers to a different source of income (labour income, rents, family allowances, etc.). Then, one can construct the total income level for each person  $i$  as the sum of the income level obtained from each source, that is  $x_i = \sum_{j=1}^d x_{ij}$ . Alternatively, each dimension  $j$  can be measured in the quantity of a good or service that can be acquired in a market. Then, one can construct the total consumption expenditure level for each person  $i$  as the sum of the quantities acquired at market price, that is  $x_i = \sum_{j=1}^d p_j x_{ij}$ , where  $p_j$ , the price of commodity  $j$ , is used as its weight. In any of these three cases, the achievement matrix  $X$  is reduced to a vector  $x = (x_1, \dots, x_n)$  containing the welfare level or the resource variables of all  $n$  persons. In other words, the distinctive feature of the unidimensional approach is not that it necessarily considers only one dimension, but rather that it maps *multiple* dimensions of poverty assessment into a *single* dimension using a common unit of account.<sup>4</sup>

### 1.1.1. 2.1.1. Identification of the Income Poor

Since Sen (1976), the measurement of poverty has been conceptualized as following two main steps: **identification** of who the poor are and **aggregation** of the information about poverty across society. In unidimensional space, the identification of who is poor is relatively straightforward: the

---

<sup>3</sup> A utility function is a (mathematical) instrument that intends to measure the level of satisfaction of a person with all possible sets of achievements (usually consumption baskets). Utility functions represent consumer preferences. The use of the utility framework for distributional analysis faces two well-known problems. First, in principle, utility functions are merely ordinal, that is, they indicate that a certain consumption basket (or achievement vector) is preferred to some other, without providing the magnitudes of the difference between two utility values. Second, in principle, the utility framework does not allow interpersonal comparability, in the sense that one cannot decide whether some utility loss of a given person (say a rich one) is less important than some utility gain of another person (say a poor one). As Sen observed, ‘... the attempt to handle social choice without using interpersonal comparability or cardinality had the natural consequence of the social welfare function being defined on the set of individual orderings. And this is precisely what makes this framework so unsuited to the analysis of distributional questions’ (Sen 1973: 12–13). In order to make this framework applicable to distributional analysis, one needs to broaden individual preferences to include interpersonally comparable cardinal welfare functions (Sen 1973: 15). One particular way in which this has been implemented is through the so-called utilitarianism approach, which defines the measure of social welfare as the sum of individual utilities; moreover, it is frequently assumed—as in the framework described above—that everyone has the same utility function.

<sup>4</sup> Alkire and Foster (2011b) provide further discussion on uni- vs multidimensional approaches.

poor are those whose overall achievement or resource variable falls below the poverty line  $z_U$ , where the subscript  $U$  simply signals that this is a poverty line used in the unidimensional space. Analogous to the construction of the resource variable, the poverty line can be obtained aggregating the minimum quantities or achievements  $z_j$  considered necessary in each dimension. It is assumed that such quantities or levels are positive values, that is  $z_j \in \mathbb{R}_{++}^d$ .<sup>5</sup> These minimum levels are collected in the  $d$ -dimensional vector  $z = (z_1, \dots, z_d)$ .

If the overall achievement is the level of utility, a utility poverty line needs to be set as  $z_U = u(z_1, \dots, z_d)$ .<sup>6</sup> On the other hand, when the overall achievement is total income or total consumption expenditure, the poverty line is given by the estimated cost of the basic consumption basket  $z_U = \sum_{j=1}^d p_j z_j$ —or some increment thereof.<sup>7</sup>

Then, given the person's overall resource value or utility value and the poverty line, we can define the identification function as follows:  $\rho_U(x_i; z_U) = 1$  identifies person  $i$  as poor if  $x_i < z_U$ , that is, whenever the resource or utility variable is below the poverty line, and  $\rho_U(x_i; z) = 0$  identifies person  $i$  as non-poor if  $x_i \geq z_U$ . We denote the number of unidimensionally poor persons in a society by  $q_U$  and the set of poor persons in a society by  $Z_U$ , such that  $Z_U = \{i | \rho_U(x_i; z_U) = 1\}$ .

### 1.1.2. 2.1.2. Aggregation of the Income Poor

In terms of aggregation, a variety of indices have been proposed.<sup>8</sup> Among them, the Foster, Greer, and Thorbecke or FGT (1984) family of indices has been the most widely used measures of poverty

---

<sup>5</sup> The concept of the poverty line dates to the late 1800s. Booth (1894, 1903), Rowntree (1901), and Bowley and Burnett-Hurst (1915) wrote seminal studies based on surveys in some UK cities. As Rowntree write, the poverty line represented the 'minimum necessities for the maintenance of merely physical efficiency' (i.e. nutritional requirements) in monetary terms, plus certain minimum sums for clothing, fuel, and household sundries according to the family size (Townsend 1954: 131).

<sup>6</sup> Axiomatic measures described in section 3.6.2 takes this approach.

<sup>7</sup> The interpretation of the variable is different if total income or total expenditure is used, with the former reflecting 'what could be' and the latter reflecting 'what is' (Atkinson 1989 cited in Alkire and Foster 2011b: 292).

<sup>8</sup> See Foster and Sen (1997), Zheng (1997), and Foster (2006) for a review of unidimensional poverty indices and Foster et al. (2013) for pedagogic coverage of poverty and other unidimensional measures, with tools for practical implementation.

by international organizations such as the World Bank and UN agencies, national governments, researchers, and practitioners.<sup>9</sup>

For simplicity, we assume the unidimensional variable  $x_i$  to be income. Building on previous poverty indices including Sen (1976) and Thon (1979), the FGT family of indices is based on the normalized income gap—called the ‘poverty gap’ in the unidimensional poverty literature—which is defined as follows: Given the income distribution  $x$ , one can obtain a *censored* income distribution  $\tilde{x}$  by replacing the values above the poverty line  $z_U$  by the poverty line value  $z_U$  itself and leaving the other values unchanged. Formally,  $\tilde{x}_i = x_i$  if  $x_i < z_U$  and  $\tilde{x}_i = z_U$  if  $x_i \geq z_U$ . Then, the normalized income gap is given by:

$$g_i = \frac{z_U - \tilde{x}_i}{z_U} \quad (2.1)$$

The normalized income gap of person  $i$  is her income shortfall expressed as a share of the poverty line. The income gap of those who are non-poor is equal to 0.<sup>10</sup> The individual income gaps can be collected in an  $n$ -dimensional vector  $g^\alpha = (g_1^\alpha, g_2^\alpha, \dots, g_n^\alpha)$ . Each  $g_i^\alpha$  element is the normalized poverty gap raised to the power  $\alpha \geq 0$  and it can be interpreted as a measure of *individual poverty* where  $\alpha$  is a ‘poverty aversion’ parameter. The class of FGT measures is defined as  $P_\alpha = \sum_{i=1}^n g_i^\alpha / n$ , thus  $P_\alpha$  can be interpreted as the average poverty in the population. The FGT measures can also be expressed in a more synthetic way as  $P_\alpha = \mu(g^\alpha)$ , where  $\mu$  is the mean operator and thus  $\mu(g^\alpha)$  denotes the average or mean of the elements of vector  $g^\alpha$ . This presentation of the FGT indices is useful in understanding the AF class (Alkire and Foster 2011a).

Within the FGT measures, three measures, associated with three different values of the parameter  $\alpha$ , have been used most frequently. The deprivation vector  $g^0$ , for  $\alpha = 0$ , replaces each income below the poverty line with 1 and replaces non-poor incomes with 0. Its associated poverty measure  $P_0 = \mu(g^0)$  is called the headcount ratio, or the mean of the deprivation vector. It indicates the

---

<sup>9</sup> Ravallion (1992) offers an early guidebook on the wide range of possible uses of the FGT measures, and Foster et al. (2010) provide a detailed retrospective of the use and extensions of this class of measures.

<sup>10</sup> An alternative way to define the normalized income deprivation gap not using the censored distribution is that  $g_i = (z_U - x_i) / z_U$  for  $x_i < z_U$ , and  $g_i = 0$  for  $x_i \geq z_U$ .

proportion of people who are poor, also frequently called the incidence of poverty.<sup>11</sup> The normalized gap vector  $g^1$ , for  $\alpha = 1$ , replaces each poor person's income with the normalized income gap and assigns 0 to the rest. Its associated measure  $P_1 = \mu(g^1)$ , the poverty gap measure, reflects the average depth of poverty across the society. The squared gap vector,  $g^2$  for  $\alpha = 2$ , replaces each poor person's income with the squared normalized income gap and assigns 0 to the rest. Its associated measure—the squared gap or distribution-sensitive FGT—is  $P_2 = \mu(g^2)$ ; it emphasizes the conditions of the poorest of the poor as Box 2.1 explains.

The FGT measures satisfy a number of properties, including a subgroup decomposability property that views overall poverty as a population-share weighted average of poverty levels in the different population subgroups.<sup>12</sup> As noted by Sen (1976), the headcount ratio violates two intuitive principles: (1) monotonicity: if a poor person's resource level falls, poverty should rise and yet the headcount ratio remains unchanged; (2) transfer: poverty should fall if two poor persons' resource levels are brought closer together by a progressive transfer between them, and yet the headcount ratio may either remain unchanged or it can even go down. The poverty gap measure satisfies monotonicity, but not the transfer principle; the  $P_2$  measure satisfies both monotonicity and the transfer principle.<sup>13</sup>

**Box 2.1 A numerical example of the FGT measures**

A simple example<sup>14</sup> can clarify the method and these axioms, and will also prove useful in linking the Alkire and Foster methodology (fully described in Chapter 5) to its roots in the FGT class of poverty measures. Consider four persons whose incomes are summarized by vector  $x = (7, 2, 4, 8)$  and the poverty line is  $z_U = 5$ . The headcount ratio  $P_0$ : Consider first the case of  $\alpha = 0$ . Each gap is

<sup>11</sup> In the epidemiological literature there is a clear distinction between the terms incidence and prevalence. Incidence refers to the number or rate of people becoming ill *during a period* of time in a specified population, whereas prevalence refers to the number or proportion of people experiencing an illness in a *particular point* in time (regardless of the moment at which they became ill). In general usage, this distinction is usually ignored and the expression 'poverty  $x_i \geq z_U$  incidence' or 'incidence of poverty' frequently refers to the proportion of poor people in a certain population at a certain point in time (which strictly speaking in epidemiological terms would be poverty prevalence), and not to the proportion of people who became poor over a certain time period (which strictly speaking in epidemiological terms would be incidence). The expression 'poverty prevalence' or 'prevalence of poverty' is also sometimes, although much less frequently, found but refers to the same concept as when incidence is used. In this book we follow the poverty literature and refer to poverty incidence as the poverty rate at a particular point in time.

<sup>12</sup> Note that the population subgroups are mutually exclusive and collectively exhaustive.

<sup>13</sup> Although we address the issue of scales of measurement later on in this chapter, it is worth anticipating that while all three mentioned members of the FGT family ( $P_0$ ,  $P_1$ , and  $P_2$ ) can be applied to cardinal variables (where distances between categories are meaningful) only the headcount ratio can be used with an ordinal variable (where distances between categories are meaningless).

<sup>14</sup> Alkire and Santos (2009).

replaced by a value of 1 if the person is poor and by a value of 0 if non-poor. The deprivation vector is given by:  $g^0 = (0, 1, 1, 0)$ , indicating that the second and third persons in this distribution are poor. The mean of this vector—the  $P_0$  measure—is one half:  $P_0(x, z_I) = \mu(g^0) = 2/4 = 0.5$ , indicating that 50% of the population in this distribution is poor. Undoubtedly, it provides very useful information. However, as noted by Watts (1968) and Sen (1976), the headcount ratio does not provide information on the depth of poverty nor on its distribution among the poor. For example, if the third person became poorer, experiencing a decrease in her income so that the income distribution became  $x' = (7, 2, 3, 8)$ , the  $P_0$  measure would still be one half; that is, it violates monotonicity. Also, if there was a progressive transfer between the two poor persons, so that the distribution was  $x'' = (7, 3, 3, 8)$ , the  $P_0$  measure would not change, violating the transfer principle. This has policy implications. If this was the official poverty measure, a government interested in maximizing the impact of resources on poverty reduction would have an incentive to allocate resources to the least poor, that is, those who were closest to the poverty line, leaving the lives of the poorest of the poor unchanged.

The poverty gap  $P_1$  (or FGT-1): Here  $\alpha = 1$ . Each gap is raised to the power  $\alpha = 1$ , giving the proportion in which each poor person falls short of the poverty line and 0 if the person is non-poor. The normalized gap vector is given by  $g^1 = (0, 3/5, 1/5, 0)$ . The  $P_1$  measure is the mean of this vector.  $P_1(x; z_I) = \mu(g^1) = 4/20$  indicates that the society would require an average of 20% of the poverty line for each person in the society to remove poverty. In fact, \$4 is the overall amount needed in this case to lift both poor persons above the poverty line. Unlike the headcount ratio  $P_0$ , the  $P_1$  measure is sensitive to the depth of poverty and satisfies monotonicity. If the income of the third person decreased so we had  $x' = (7, 2, 3, 8)$  the corresponding normalized gap vector would be  $g^{1'} = (0, 3/5, 2/5, 0)$ , so  $P_1(x'; z_U) = 5/20$ . Clearly,  $P_1(x'; z_U) > P_1(x; z_U)$ . Indeed, all measures with  $\alpha > 0$  satisfy monotonicity. However, a transfer to an extremely destitute person from a less poor person would not change  $P_1$ , since the decrease in one gap would be exactly compensated by the increase in the other. By being sensitive to the depth of poverty (i.e. satisfying monotonicity), the  $P_1$  measure does make policymakers want to decrease the average depth of poverty as well as reduce the headcount. But because of its insensitivity to the distribution among the poor,  $P_1$  does not provide incentives to target the very poorest, whereas the FGT-2 measure does.

The Squared Poverty Gap  $P_2$  (or FGT-2): When we set  $\alpha = 2$ , each normalized gap is squared or raised to the power  $\alpha = 2$ . The squared gap vector in this case is given by  $g^2 = (0, 9/25, 1/25, 0)$ . By squaring the gaps, bigger gaps receive higher weight. Note, for example, that while the gap of the second person (3/5) is three times bigger than the gap of the third person (1/5), the squared gap of the second person (9/25) is nine times bigger than the gap of the third person (1/25). The mean of the  $g^2$  vector—the  $P_2$  measure—is  $P_2(x; z_U) = \mu(g^2) = 10/100$ . The  $P_2$  measure is sensitive to the depth of poverty: if the income of the third person decreases one unit such that  $x' = (7, 2, 3, 8)$ , the squared gap vector becomes  $g^{2'} = (0, 9/25, 4/25, 0)$ , increasing the aggregate poverty level to

$P_2(x'; z_U) = 13/100$ ). It is also sensitive to the distribution among the poor: if there is a transfer of \$1 from the third person to the second one, so  $x = (7, 2, 4, 8)$  becomes  $x'' = (7, 3, 3, 8)$ , the squared gap vector becomes  $g^2 = (0, 4/25, 4/25, 0)$ , decreasing the aggregate poverty level to  $P_2(x''; z_U) = 8/100$ . Squaring the gaps has the effect of emphasizing the poorest poor and providing incentives to policymakers to address their situation urgently. All measures with  $\alpha > 1$  satisfy the transfer property.

## 2.2 Notation and Preliminaries for Multidimensional Poverty Measurement

We now extend the notation to the multidimensional context. We represent achievements as  $n \times d$  dimensional achievement matrix  $X$ , as in the unidimensional framework described in section 2.1. We make two practical assumptions for convenience. We assume that the achievement of person  $i$  in dimension  $j$  can be represented by a non-negative real number, such that  $x_{ij} \in \mathbb{R}_+$  for all  $i = 1, \dots, n$  and  $j = 1, \dots, d$ . Also, we assume that higher achievements are preferred to lower ones.<sup>15</sup> In a multidimensional setting, in contrast to a unidimensional context, the considered *achievements* may not be combinable in a meaningful way into some overall variable. In fact, each dimension can be of a different nature. For example, one may consider a person's income, level of schooling, health status, and occupation, which do not have any common unit of account. As in the unidimensional case, we allow the population size of a society to vary, and we assume  $d$  to denote a fixed set (and number) of dimensions.<sup>16</sup>

We denote the set of all possible matrices of size  $n \times d$  by  $\mathcal{X}_n \in \mathbb{R}_+^{n \times d}$  and the set of all possible achievement matrices by  $\mathcal{X}$ , such that  $\mathcal{X} = \bigcup_n \mathcal{X}_n$ . If  $X \in \mathcal{X}_n$ , then matrix  $X$  contains achievements for  $n$  persons and a fixed set of  $d$  dimensions. Unless specified otherwise, whenever we refer to matrix  $X$ , we assume  $X \in \mathcal{X}$ . The achievements of any person  $i$  in all  $d$  dimensions, which is row  $i$  of matrix  $X$ , are represented by the  $d$ -dimensional vector  $x_i$  for all  $i = 1, \dots, n$ . The achievements in

---

<sup>15</sup> In empirical applications some indicators may not be restricted to the non-negative range, or be scored such that *larger values are worse*, or that the lowest attainable value is strictly positive. For example, the z-scores of children's nutritional indicators may take negative values; in a people-per-room indicator, larger values are worse. And the lowest possible Body Mass Index for human survival is strictly positive. Such indicators may require rescaling.

<sup>16</sup> For simplicity of presentation, in theoretical sections, we use the term dimension to refer to each variable; in empirical presentations often we use the term 'indicator' for the variables, while 'dimension' refers to groupings of indicators.

any dimension  $j$  for all  $n$  persons, which is column  $j$  of matrix  $X$ , are represented by the  $n$ -dimensional vector  $x_{.j}$  for all  $j=1, \dots, d$ .

In multidimensional analysis, each dimension may be assigned a weight or deprivation value based on its relative importance or priority. We denote the relative weight attached to dimension  $j$  by  $w_j$ , such that  $w_j > 0$  for all  $j=1, \dots, d$ . The weights attached to all  $d$  dimensions are collected in a vector  $w=(w_1, \dots, w_d)$ . For convenience we may restrict the weights such that they sum to the total number of considered dimensions, that is:  $\sum_j w_j = d$ . Alternatively, weights may be normalized; in other words, the weights sum to one:  $\sum_j w_j = 1$ .<sup>17</sup>

### 2.2.1 Identifying Deprivations

A common first step in multidimensional poverty assessment in several of the methodologies reviewed in Chapter 3, as well as in the Alkire and Foster (2007, 2011a) methodology, requires defining a threshold in *each* dimension. Such a threshold is the minimum level someone needs to achieve in that dimension in order to be non-deprived. It is called the dimensional **deprivation cutoff**. When a person's achievement is strictly below the cutoff, she is considered deprived. We denote the deprivation cutoff in dimension  $j$  by  $z_j$ ; the deprivation cutoffs for all dimensions are collected in the  $d$ -dimensional vector  $z=(z_1, \dots, z_d)$ . We denote all possible  $d$ -dimensional deprivation cutoff vectors by  $z \in \mathbb{R}_{++}^d$ . Any person  $i$  is considered deprived in dimension  $j$  if and only if  $x_{ij} < z_j$ .

For several measures reviewed in Chapter 3, and for the AF method, it will prove useful to express the data in terms of deprivations rather than achievements. From the achievement matrix  $X$  and the vector of deprivation cutoffs  $z$ , we obtain a deprivation matrix  $g^0$  (analogous to the deprivation vector in the unidimensional context) whose typical element  $g_{ij}^0 = 1$  whenever  $x_{ij} < z_j$  and  $g_{ij}^0 = 0$ , otherwise, for all  $j=1, \dots, d$  and for all  $i=1, \dots, n$ . In other words, if person  $i$  is deprived in dimension  $j$ , then the person is assigned a deprivation status value of 1, and 0 otherwise. Thus, matrix  $g^0(X)$  represents the deprivation status of all  $n$  persons in all  $d$  dimensions in matrix  $X$ . Vector  $g_i^0$  represents the deprivation status of person  $i$  in all dimensions and vector  $g_{.j}^0$  represents

---

<sup>17</sup> Note that the prices used in the unidimensional case provide a particular weighting structure, where the weights do not necessarily sum to  $d$  or 1.

the deprivation status of all persons in dimension  $j$ . From the matrix  $g^0$  one can construct a deprivation score  $c_i$  for each person  $i$  such that  $c_i = \sum_{j=1}^d w_j g_{ij}^0$ . In words,  $c_i$  denotes the sum of *weighted* deprivations suffered by person  $i$ .<sup>18</sup> In the particular case in which weights are equal and sum to the number of dimensions, the score is simply the number of deprivations or deprivation counts that the person experiences. Whenever weights are unequal but sum to the number of dimensions, person  $i$ 's **deprivation score** is defined as the sum of her weighted deprivation counts. The deprivation scores are collected in an  $n$ -dimensional column vector  $c$ .

On certain occasions, it will be useful to use the deprivation-cutoff-censored achievement matrix  $\tilde{X}$ , which is obtained from the corresponding achievement matrix  $X$  in  $\mathcal{X}$ , replacing the non-deprived achievements by the corresponding deprivation cutoff and leaving the rest unchanged. We denote the  $ij^{\text{th}}$  element of  $\tilde{X}$  by  $\tilde{x}_{ij}$ . Then, formally,  $\tilde{x}_{ij} = x_{ij}$  if  $x_{ij} < z_j$ , and  $\tilde{x}_{ij} = z_j$  otherwise. In this way, all achievements greater than or equal to the deprivation cutoffs are ignored in the censored achievement matrix.

When data are cardinally meaningful for all  $i=1, \dots, n$  and all  $j=1, \dots, d$ , and  $z_j \in \mathbb{R}_{++}$ , in other words, when all the achievements take non-negative values and the deprivation cutoffs take strictly positive values, one can construct dimensional gaps or shortfalls from the censored achievement matrix  $\tilde{X}$  as:<sup>19</sup>

$$g_{ij} = \frac{z_j - \tilde{x}_{ij}}{z_j} . \quad (2.2)$$

Each  $g_{ij}$ , or normalized gap, expresses the shortfall of person  $i$  in dimension  $j$  as a share of its deprivation cutoff. Naturally, the gaps of those whose achievement  $x_{ij}$  is above the corresponding dimensional deprivation cutoff  $z_j$  are equal to 0. Generalizing the above, the individual normalized gaps can be collected in an  $n \times d$  dimensional matrix  $g^\alpha$  where each  $g_{ij}^\alpha$  element is the normalized gap defined in (2.2) raised to the power  $\alpha$ ; such normalized gaps can be interpreted as a measure of *individual deprivation* in dimension  $j$ . When  $\alpha = 0$ , we have the  $g^0$  deprivation matrix already defined.

---

<sup>18</sup> Alternative notations for the AF methodology are presented and elaborated in Ch. 5.

<sup>19</sup> This is an analogous construct to the income gaps in the FGT measures. An alternative way to define the deprivation gaps not using the censored distribution is that  $g_{ij} = (z_j - x_{ij}) / z_j$  when  $x_{ij} < z_j$  and  $g_{ij} = 0$  when  $x_{ij} \geq z_j$ .

When  $\alpha = 1$ , we have the  $g^1$  matrix of normalized gaps, and when  $\alpha = 2$ , we have the  $g^2$  matrix of squared gaps. Analogous to the FGT measures,  $\alpha \geq 0$  is a deprivation aversion parameter.

### 2.2.2 Identification and Aggregation in the Multidimensional Case

Sen's (1976) steps of **identification** of the poor and **aggregation** also apply to the multidimensional case. It is clear that the identification of who is poor in the unidimensional case is relatively straightforward. The poverty line dichotomizes the population into the sets of poor and non-poor. In other words, in the unidimensional case, a person is poor if she is deprived. However, in the multidimensional context, the identification of the poor is more complex: the terms 'deprived' and 'poor' are no longer synonymous. A person who is deprived in any particular dimension may not necessarily be considered poor. An identification method, with an associated identification function, is used to define who is poor.

We denote the identification function by  $\rho$ , such that  $\rho(\cdot) = 1$  identifies person  $i$  as poor and  $\rho(\cdot) = 0$  identifies person  $i$  as non-poor. Analogous to the unidimensional case, we denote the number of multidimensionally poor people in a society by  $q$  and the set of poor persons in a society by  $Z$ , such that  $Z = \{i | \rho(\cdot) = 1\}$ . It could be the case that the identification method is based on some 'exogenous' variable, in that it is a variable not included in achievement matrix  $X$ . For example, the exogenous variable could be being the beneficiary of some government programme or living in a specific geographic area. One may also define an identification method based on one particular dimension  $j$  of matrix  $X$ . One may consider the corresponding normative cutoff  $z_j$  to identify the person as poor, in which case the function is  $\rho(x_{ij}; z_j)$ , or one may consider a relative cutoff identifying as poor anyone who is below the median or mean value of the distribution, in which case the function is  $\rho(x_{.j})$ . Alternatively, identification may be based on the whole set of achievements, not necessarily considering dimensional deprivation cutoffs but rather the relative position of each person on the aggregate distribution  $\rho(X)$ .

There are many different ways of identifying the poor in the multidimensional context. A particularly prevalent set of methods consider the person's vector of achievements and corresponding deprivation cutoffs, such that  $\rho(x_i; z) = 1$  identifies person  $i$  as poor and  $\rho(x_i; z) = 0$  identifies

person  $i$  as non-poor.<sup>20</sup> Within this specification of the identification function, at least two approaches can be followed. An approach closely approximating unidimensional poverty is the ‘aggregate achievement approach’, which consists of applying an aggregation function to the achievements across dimensions for each person to obtain an overall achievement value. The same aggregation function is also applied to the dimensional deprivation cutoffs to obtain an aggregate poverty line. As in the unidimensional case, a person is identified as poor when her overall achievement is below the aggregate poverty line. Another method, which we refer to as ‘censored achievement approach’, first applies deprivation cutoffs to identify whether a person is deprived or not in each dimension and then identifies a person considering only the deprived achievements. The ‘counting approach’ is one possible censored achievement approach, which identifies the poor according to the number (count) of deprivations they experience. Note that ‘number’ here has a broad meaning as dimensions may be weighted differently. Chapter 4 and the AF method (Chs. 5–10) use a counting approach. When the scale of the variables allows, other identification methods could be developed using the information on the deprivation gaps.

In counting identification methods, the criterion for identifying the poor can range from ‘union’ to the ‘intersection’. The union criterion identifies a person as poor if the person is deprived in *any* dimension, whereas the intersection criterion identifies a person as poor only if she is deprived in *all* considered dimensions. In between these two extreme criteria there is room for intermediate criteria. Many counting-based measurement exercises since the mid-1970s have used intermediate criteria (see Ch. 4). The AF methodology formally incorporated it into an axiomatic framework.<sup>21</sup>

Once the identification method has been selected, the aggregation step requires selecting a poverty index, which summarizes the information about poverty across society. A poverty index is a function  $P: \mathcal{X} \times \mathcal{z} \rightarrow \mathbb{R}$  that converts the information contained in the achievement matrix  $X \in \mathcal{X}$  and the deprivation cutoff vector  $z \in \mathcal{z}$  into a real number. We denote a poverty index as  $P(X; z)$ .

---

<sup>20</sup> Note that this identification function differs from the one introduced in the unidimensional case in that it depends on the vector of achievements  $x_i$  and the vector of dimensional deprivation cutoffs  $z$ . In the unidimensional case, identification depends on the already-aggregated overall achievement or resource variable  $x_i$  and the aggregate poverty line  $z_U$ , which of course may depend upon the prices of certain commodities.

<sup>21</sup> Within the aggregate achievement approach, the intermediate criterion is operationalized by using the so-called ‘poverty frontier’, defined as the different combinations of the  $d$  achievements that provide the same overall achievement as the aggregate poverty line. Duclos, Sahn, and Younger (2006a) further elaborate the poverty frontier; cf. Atkinson (2003) and Bourguignon and Chakravarty (2003).

An identification and an aggregation method that are used together constitute what we call a multidimensional poverty methodology, and we denote it as  $\mathcal{M} = (\rho, P)$ .

It will prove useful to introduce notation for two further partial indices related to the overall poverty index  $P(X; z)$ . Each of them offers information on different ‘slices’ of the achievement matrix  $X$  as analysed by the corresponding multidimensional poverty methodology  $\mathcal{M} = (\rho, P)$ ; that is, the partial indices are dependent on the selected identification and aggregation methods. One of these partial indices is the poverty level of a particular subgroup within the total population, denoted as  $P(X^\ell; z)$ , where  $X^\ell$  is the achievement matrix of this particular subgroup  $\ell$  (which could be people of a particular ethnicity, for example) contained in matrix  $X$ . Visually, this partial index is based on a horizontal slice of the achievement matrix  $X$  (i.e. a set of rows).

The other partial index is a function of the **post-identification dimensional deprivations**, denoted as  $P_j(x_j; z)$ , where, as stated above, vector  $x_j$  represents the achievements in dimension  $j$  for all  $n$  persons and  $z$  is the deprivation cutoffs vector.<sup>22</sup> This is precisely why the full vector of deprivation cutoffs  $z$  is an argument of the index and not just the particular deprivation cutoff  $z_j$ . Recall that under some identification methods, it is possible to have some people who experience deprivations but are not identified as poor (for example, when a counting approach is used with an identification strategy that it is not union). In such cases, their deprivations will not be considered in the poverty measure and therefore will not be considered in this partial index either. In a visual way, this partial index is based on a vertical slice of the achievement matrix  $X$  (i.e. a column). These two partial indices will be used when introducing the different principles of multidimensional poverty measures in section 2.5.3.

As we shall see, although identification and aggregation have, since Sen (1976), usually been recognized as key steps in poverty measurement, some methods in the multidimensional context do not follow these steps. We will clarify this in Chapter 3.

### 2.2.3 The Joint Distribution

Throughout this book we will frequently refer to the **joint distribution** in contrast to the **marginal distribution** and we will also use the expression **joint deprivations**. The concept of a joint

---

<sup>22</sup> In one of the measures in the AF class, the Adjusted Headcount Ratio, this partial index is called the censored headcount ratio. See section 5.5.3 for a detailed presentation.

distribution comes from statistics where it can be represented using a joint cumulative distribution function.<sup>23</sup> The relevance of the joint distribution in multidimensional analysis was articulated by Atkinson and Bourguignon (1982), who observed that multidimensional analysis was *intrinsically different* because there could be identical dimensional marginal distributions but differing degrees of interdependence between dimensions.<sup>24</sup>

In this book we treat the achievement matrix  $X$  as a representation of the joint distribution of achievements. Each row contains the (vector of) achievements of a given person in the different dimensions, and each column contains the (vector of) achievements in a given dimension across the population. From that matrix, considered with deprivation cutoffs, it is possible to obtain the proportion of the population who are simultaneously deprived in different subsets of  $d$  dimensions. In other words, it is possible to obtain the proportion of people who experience each possible profile of deprivations. This is visually clear in the deprivation matrix  $g^0$ , which represents the joint distribution of deprivations. The higher-order matrices  $g^1$  and  $g^2$  obviously offer further information regarding the joint distribution of the depths of deprivations.

The importance of considering the joint distribution of achievements, which in turn enables us to look at joint deprivations, is best understood in contrast with the alternative of looking at the **marginal distribution** of achievements, and thus, the **marginal deprivations**. The marginal distribution is the distribution in one specific dimension without reference to any other dimension.<sup>25</sup> The marginal distribution of dimension  $j$  is represented by the column vector  $x_j$ . From the marginal distribution of each dimension, it is possible to obtain the proportion of the population deprived with respect to a particular deprivation cutoff. However, by looking at only the marginal distribution, one does not know who is simultaneously deprived in other dimensions.<sup>26</sup>

---

<sup>23</sup> Given two random variables,  $y_1$  and  $y_2$ , the joint distribution can be described with the bivariate cumulative distribution function  $F(b_1, b_2) = \text{Prob}(y_1 \leq b_1, y_2 \leq b_2)$ . In words, the joint distribution gives the proportion of the population with values of  $y_1$  and  $y_2$  lower than  $b_1$  and  $b_2$  correspondingly and simultaneously.

<sup>24</sup> The authors analyse inequality in the two-dimensional case. They introduce the transformation in which there is an increase in the correlation of the achievements, leaving the marginal distributions unchanged—something we discuss in section 0. They extend the conditions for second-order stochastic dominance, noting that such conditions depend on the joint distribution.

<sup>25</sup> Given any random variable  $y_j$ , the marginal distribution can be described with the cumulative distribution function  $F_j(b_j) = \text{Prob}(y_j \leq b_j)$ .

<sup>26</sup> Only in the very particular case in which the two variables are statistically independent, can one obtain the joint distribution from the marginal ones. In such a case, the proportion of people deprived simultaneously in a number of

Table 2.1 illustrates the relevance of the joint distribution in the basic case of  $n$  persons and two dimensions using a contingency table.

**Table 2.1 Joint distribution of deprivation in two dimensions**

		Dimension 2		Total
		Non-deprived	Deprived	
Dimension 1	Non-deprived	$n_{00}$	$n_{01}$	$n_{0+}$
	Deprived	$n_{10}$	$n_{11}$	$n_{1+}$
Total		$n_{+0}$	$n_{+1}$	$n$

We denote the number of people deprived and non-deprived in the first dimension by  $n_{1+}$  and  $n_{0+}$ , respectively; whereas, the number of people deprived and non-deprived in the second dimension are denoted by  $n_{+1}$  and  $n_{+0}$ , respectively. These values correspond to the marginal distributions of both dimensions as depicted in the final row and final column of the table. They could equivalently be expressed as proportions of the total, in which case, for example,  $(n_{1+} / n)$  would represent the proportion of people deprived (or the headcount ratio) in Dimension 1.

The marginal distributions, however, do not provide information about the joint distribution of deprivations, which is described in the four internal cells of the table. In particular, the number of people deprived in both dimensions is denoted by  $n_{11}$ , the number of people deprived in the first but not the second dimension is denoted by  $n_{10}$ , and the number of people deprived in the second and not in the first dimension is denoted by  $n_{01}$ . We know that  $n_{11}$  people are deprived in both dimensions and the sum of  $n_{11} + n_{01} + n_{10}$  is the number of people deprived in at least one dimension. These values correspond to the joint distribution of deprivations.

Consider now the case of four dimensions and four people, to see how valuable information can be added by the joint distribution. Table 2.2<sup>27</sup> presents the deprivation matrix  $g^0$  of two hypothetical distributions,  $X$  and  $X'$ . Such a matrix presents joint distributions of deprivations in a compact way and is used regularly throughout this book.

**Table 2.2 Comparison of two joint distributions of deprivations in four dimensions**

---

variables can be obtained as the product of the proportions of people deprived in each variable. Although this is a topic for further empirical research, *a priori*, it seems unlikely that the independence condition will be satisfied, especially as the number of considered dimensions increases.

<sup>27</sup> Alkire and Foster (2011b). Similar examples on the relevance of considering the joint distribution in the measurement of multidimensional welfare and poverty can be found in Tsui (2002), Pattanaik, Reddy, and Xu (2012), and Seth (2009).

$$g^0(X) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad g^0(X') = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Dimensional  $P_0$ :  $[0.25 \quad 0.25 \quad 0.25 \quad 0.25]$        $[0.25 \quad 0.25 \quad 0.25 \quad 0.25]$

In the table, the marginal distributions of each dimension's  $P_0$  are identical in deprivation matrices  $g^0(X)$  and  $g^0(X')$ . Thus, the proportions of people deprived in each dimension are the same in the two distributions (25%). Yet, while, in distribution  $X$ , one person is deprived in all dimensions and three people experience zero deprivations, in distribution  $X'$ , each of the four persons is deprived in exactly one dimension. In other words, although the marginal distributions are identical, the two joint distributions  $X$  and  $X'$  are very different. We understand that multiple deprivations that are *simultaneously* experienced are at the core of the concept of multidimensional poverty, and this is the reason why the consideration of the joint distribution is important. However, as we shall see, not all methodologies consider the joint distribution. In the next section, we introduce the notation for two methodologies of this type.

#### 2.2.4 Marginal Methods

Some of the methods for multidimensional poverty assessment introduced in Chapter 3 can be called **marginal methods** because they do not use information contained in the joint distribution of achievements. In other words, they ignore all information on links across dimensions. Following Alkire and Foster (2011b), a marginal method assigns the same level of poverty to any two matrices that generate the same marginal distributions. In Table 2.2, a marginal method would assign the same poverty level to distribution  $X$  (four deprivations are experienced by one person) and distribution  $X'$  (each person experiences exactly one deprivation). That is, it would not be able to show whether the deprivations are spread evenly across the population or whether they are concentrated in an underclass of multiply deprived persons. Such marginal methods can also be linked to the order of aggregation while constructing poverty indices (Pattanaik, Reddy and Xu 2012). Specifically, a measure can be obtained by first aggregating achievements or deprivations across people (column-first) within each dimension and then aggregating across dimensions, or it can be obtained by first aggregating achievements or deprivations for each person (row-first) and then aggregating across people. Only measures that follow the second order of aggregation (i.e. first

across dimensions for each person and then across persons) reflect the joint distribution of deprivations (Alkire 2011: 61, Figure 7). Measures that follow the first order of aggregation fall under marginal methods of poverty measurement.

Marginal methods also include cases where achievements for different dimensions are drawn from *different* data sources and/or from *different* reference groups within a population—as occurred, for example, in many indicators associated with the Millennium Development Goals. In this case, rather than having an  $n \times d$  dimensional matrix  $X$  of achievements, one may have a ‘collection’ of vectors  $x^j \in \mathbb{R}_+^{n_j}$ , representing the achievements of  $n_j$  people in dimension  $j$  for all  $j = 1, \dots, d$ . Note that each  $x^j$  vector may refer to *different* sets of people such as children, adults, workers, or females, to mention a few.

Suppose as before, that a deprivation cutoff  $z_j \in \mathbb{R}_{++}$  is defined. Then, we define a *dimensional deprivation index*  $P_j(x^j; z_j)$  for dimension  $j$  by  $P_j: \mathbb{R}_+^{n_j} \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ , which assesses the deprivation profile of  $n_j$  people in dimension  $j$ . The deprivation index might be simply the percentage of people who are deprived in this indicator or some other statistic such as child mortality rates. Note that deprivations are clearly identified in each dimension; however, because the underlying columns of dimensional achievements are not linked, no decision on who is to be considered multidimensionally poor can be made. This is a key difference between the dimensional deprivation index  $P_j(x^j; z_j)$  and the post-identification dimensional deprivation index  $P_j(x_j; z)$ , which depends on the entire deprivation cutoff vector  $z$  (and not just  $z_j$ ) and thus captures the joint distribution of deprivations (see section 2.2.2). Different dimensional deprivation indices  $P_j(x^j, z_j)$  can be considered in a set, constituting the ‘dashboard approach’, or combined by some aggregation function, which is often called a ‘composite index’ (Ch. 3).

### 2.2.5 Useful Matrix and Vector Operations

Throughout the book, we use specific vector and matrix operations. This section introduces the technical notation covering vectors and matrices.

We denote the transpose of any matrix  $X$  by  $X^{\text{tr}}$  where  $X^{\text{tr}}$  has the rows of matrix  $X$  converted into columns. Formally, if  $X \in \mathbb{R}^{n \times d}$  and the  $ij^{\text{th}}$  element of  $X$  is written  $x_{ij}$ , then  $X^{\text{tr}} \in \mathbb{R}^{d \times n}$ , where

$x_{ji}^{\text{tr}} = x_{ij}$  is the  $ij^{\text{th}}$  element of  $X^{\text{tr}}$  for all  $j=1, \dots, d$  and  $i=1, \dots, n$ . The same notation applies to a vector, with  $x^{\text{tr}}$  being the transpose of  $x$ . Thus, if  $x$  is a row vector,  $x^{\text{tr}}$  is a column vector of the same dimension.

As stated in section 2.1 the average or mean of the elements of any vector  $x$  is denoted by  $\mu(x)$ , where  $\mu(x) = \sum_{i=1}^n x_i / n$ . Similarly, the average or mean of the elements of any matrix  $X$  is denoted by  $\mu(X)$ , where  $\mu(X) = \sum_{i=1}^n \sum_{j=1}^d x_{ij} / (nd)$ .

Later in the book we use a related expression, the so-called ‘generalized mean of order’  $\beta \in \mathbb{R}$ . Given any vector of achievements  $y = (y_1, \dots, y_d)$ , where  $y_j \in \mathbb{R}_{++}^d$ , the expression of the weighted generalized mean of order  $\beta$  is given by

$$\mu_{\beta}(y; w) = \begin{cases} \left( \sum_{j=1}^d w_j y_j^{\beta} \right)^{\frac{1}{\beta}} & \text{for } \beta \neq 0 \\ \prod_{j=1}^d (y_j)^{w_j} & \text{for } \beta = 0 \end{cases} \quad (2.3)$$

where  $w_j > 0$  and  $\sum_j w_j = 1$ . When weights are equal,  $w_j = 1/d$  for all  $j$ . Each generalized mean summarizes distribution  $y$  into a single number and can be interpreted as a ‘summary’ measure of well- or ill-being, depending on the meaning of the arguments  $y_j$ . When  $w_j = 1/d$  for all  $j$ , we write  $\mu_{\beta}(y; w)$  simply as  $\mu_{\beta}(y)$ . When  $\beta = 1$ ,  $\mu_{\beta}(y)$  reduces to the arithmetic mean and is simply denoted by  $\mu(y)$ . When  $\beta > 1$ , more weight is placed on higher entries and  $\mu_{\beta}(y)$  is higher than the arithmetic mean, approaching the maximum entry as  $\beta$  tends to  $\infty$ . For  $\beta < 1$  more weight is placed on lower entries, and  $\mu_{\beta}(y)$  is lower than the arithmetic mean, approaching the minimum entry as  $\beta$  tends to  $-\infty$ . The case of  $\beta = 0$  is known as the geometric mean and  $\beta = -1$  as the harmonic mean. Expression (2.3) is also known as a constant elasticity of substitution function, frequently used as a utility function in economics. When generalized means are computed over achievements, it is natural to restrict the parameter to the range of  $\beta < 1$ , giving a higher weight to lower achievements and penalizing for inequality (Atkinson 1970). Likewise when generalized means

are computed over deprivations, it is natural to restrict the parameter to the range of  $\beta > 1$ , giving a higher weight to higher deprivations and also penalizing for inequality. Box 2.2 contains an example of generalized means.

### Box 2.2 Example of generalized means

Consider<sup>28</sup> two distributions  $y$  and  $y'$  with the following distribution of achievements in a particular dimension:  $y = (2, 6, 7)$  and  $y' = (1, 5, 9)$ . We first show how to calculate  $\mu_\beta(y)$  for certain values of  $\beta$  and then compare two distributions with a graph where  $\beta$  ranges from  $-2$  to  $2$ . In this example, we assume that all dimensions are equally weighted:  $w_1 = w_2 = w_3 = 1/3$ .

**Arithmetic Mean:** The arithmetic mean ( $\beta = 1$ ) of distribution  $y$  is:  $\mu_1(y) = (2 + 6 + 7) / 3 = 5$ .

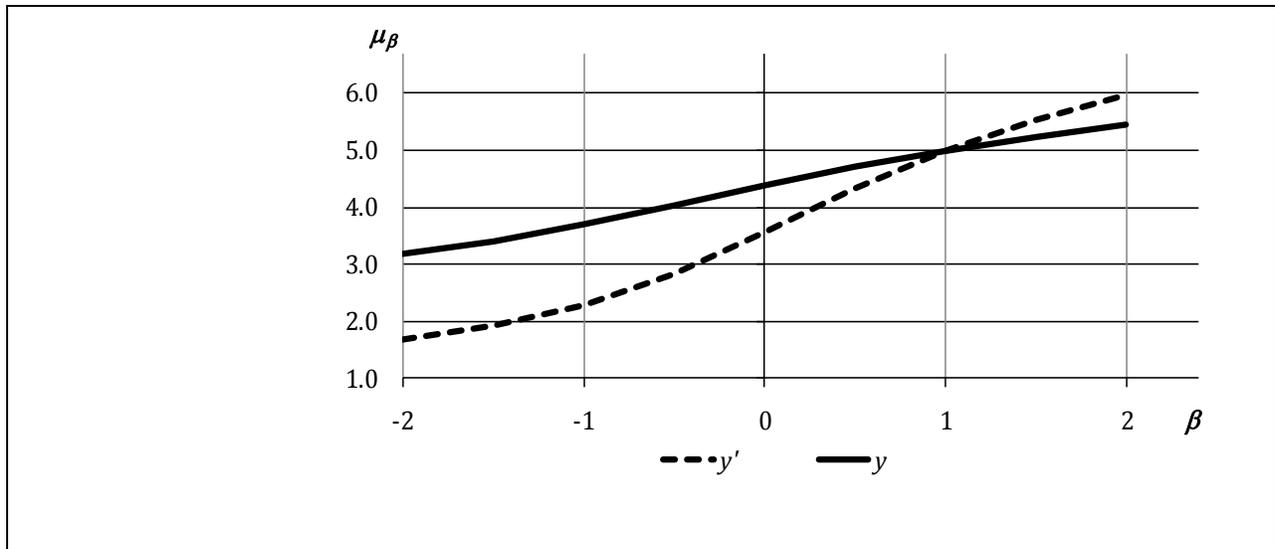
**Geometric Mean:** If  $\beta = 0$ , then  $\mu_\beta$  is the 'geometric mean' and by the formula presented in (2.3) can be calculated as:  $\mu_0(y) = (2)^{1/3} \times (6)^{1/3} \times (7)^{1/3} = 4.38$ .

**Harmonic Mean:** If  $\beta = -1$ , then  $\mu_\beta$  is the harmonic mean and can be calculated as:

$$\mu_{-1}(y) = \left[ \frac{(2)^{-1} + (6)^{-1} + (7)^{-1}}{3} \right]^{-1} = 3.71.$$

The following graph depicts the values of the  $\mu_\beta$  of  $y$  and  $y'$  for different values of  $\beta$ . Note that  $\mu_\beta(y) = \mu_\beta(y')$  when  $\beta = 1$ , given that the two distributions have the same arithmetic mean. In both cases, when  $\beta < 1$ , the generalized means are strictly lower than the arithmetic mean, because the incomes are unequally distributed. Note moreover that for this range,  $\mu_\beta(y') < \mu_\beta(y)$ , because  $y'$  has a more unequal distribution. On the other hand, for  $\beta > 1$ ,  $\mu_\beta(y') > \mu_\beta(y)$ , as the higher incomes receive a higher weight.

<sup>28</sup> Alkire and Santos (2009).



Another matrix transformation that we use is **replication**. A matrix is a replication of another matrix if it can be obtained by duplicating the rows of the original matrix a finite number of times. Suppose the rows of matrix  $Y$  are replicated  $\gamma$  number of times, where  $\gamma \in \mathbb{N} \setminus \{1\}$ . Then the corresponding replication matrix is denoted by  $rep(Y; \gamma)$ . This notation may be used for replication of any vector  $y: rep(y; \gamma)$ . We do not consider column replication, as we consider a fixed set of dimensions.

We also use three types of matrices associated with particular operations: a **permutation matrix**, a **diagonal matrix**, and a **bistochastic matrix**. A permutation matrix, denoted by  $\Pi$ , is a square matrix with one element in each row and each column equal to 1 and the rest of the elements equal to 0. Thus the elements in every row and every column sum to one. We eliminate the special case when a permutation matrix is an ‘identity matrix’ with the diagonal elements equal to 1 and the rest equal to 0. What does a permutation matrix do? If any matrix  $Y$  is pre-multiplied by a permutation matrix, then the rows of matrix  $Y$  are shuffled without their elements being altered. Similarly, if any matrix  $Y$  is post-multiplied by a permutation matrix, then the columns of  $Y$  are shuffled without their elements being altered.

**Example of Permutation Matrix:** Let  $Y = \begin{bmatrix} 6 & 4 & 2 \\ 8 & 6 & 4 \end{bmatrix}$ . Consider  $\Pi = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ . Then  $\Pi Y = \begin{bmatrix} 8 & 6 & 4 \\ 6 & 4 & 2 \end{bmatrix}$ .

Thus, the rows of  $Y$  are merely swapped. Similarly, consider  $\Pi' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ . Then  $\Pi' Y = \begin{bmatrix} 6 & 2 & 4 \\ 8 & 4 & 6 \end{bmatrix}$ . Note

that the second and the third columns have swapped their positions. The first column did not change its position because the first diagonal element in  $\Pi$  is equal to one.

A diagonal matrix, denoted by  $\Lambda$ , is a square matrix whose diagonal elements are not necessarily equal to 1 but all off-diagonal elements are equal to 0. Let us denote the  $ij^{\text{th}}$  element of  $\Lambda$  by  $\Lambda_{ij}$ . Then,  $\Lambda_{ij} = 0$  for all  $i \neq j$ . For our purposes, we require the diagonal elements of a diagonal matrix to be strictly positive or  $\Lambda_{ii} > 0$ . What is the use of a diagonal matrix? If any matrix  $Y$  is post-multiplied by a diagonal matrix, then the elements in each column are changed in the same proportion. Note that different columns may be multiplied by different factors.

**Example of Diagonal Matrix:** Let  $Y = \begin{bmatrix} 6 & 4 & 2 \\ 8 & 6 & 4 \end{bmatrix}$ . Consider  $\Lambda = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ . Then  $\Lambda Y = \begin{bmatrix} 3 & 4 & 4 \\ 4 & 6 & 8 \end{bmatrix}$ .

Each element in the first column has been halved and each element in the third column has been doubled. However, the second column did not change because the corresponding element of  $\Lambda$  is equal to one.

A bistochastic matrix, denoted by  $B$ , is a square matrix in which the elements in each row and each column sum to one. If the  $ij^{\text{th}}$  element of  $B$  is denoted by  $B_{ij}$ , then  $\sum_i B_{ij} = 1$  for all  $j$  and  $\sum_j B_{ij} = 1$  for all  $i$ . Why do we require a bistochastic matrix? If a matrix is pre-multiplied by a bistochastic matrix, then the variability across the elements of each column is reduced while their average or mean is preserved. Note that if a diagonal element in a bistochastic matrix is equal to one, the achievement vector of the corresponding person remains unaffected. If the bistochastic matrix is a permutation matrix or an identity matrix, then the variability remains unchanged.

**Example of Bistochastic Matrix:** Let  $Y = \begin{bmatrix} 6 & 4 & 2 \\ 8 & 6 & 4 \end{bmatrix}$ . Consider  $B = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ . Then  $BY = \begin{bmatrix} 7 & 5 & 6 \\ 7 & 5 & 6 \end{bmatrix}$ .

The bistochastic matrix equalizes the achievements across its elements.

### 2.3 Scales of Measurement: Ordinal and Cardinal Data

An important element of the framework in multidimensional poverty measurement relates to the scales of measurement of the indicators used. Scales of measurement are key because they affect the kind of *meaningful* operations that can be performed with indicators. In fact, as we will observe,

certain types of indicators may not allow a number of operations and thus cannot be used to generate certain poverty measures.

What does **scale of measurement** refer to exactly? Following Roberts (1979) and Sarle (1995), we define a scale of measurement to be a particular way of assigning numbers or symbols to assess certain aspects of the empirical world, such that the relationships of these numbers or symbols replicate or represent certain observed relations between the aspects being measured. There are different classifications of scales of measurement. In this book, we follow the classification introduced by Stevens (1946) and discussed in Roberts (1979). Stevens' classification is consistent with Sen (1970, 1973), which analysed the implications of scales of measurement for welfare economics, distributional analysis, and poverty measurement, and it has largely stood the test of time.<sup>29</sup>

Stevens' (1946) classification relies on four key concepts: assignment rules, admissible transformations, permissible statistics, and meaningful statements. First, the defining feature of a scale is the *rule* or basic empirical operation that is followed for assigning numerals, as elaborated below. Second, each scale has an associated set of *admissible mathematical transformations* such that the scale is preserved. That is, if a scale is obtained from another under an admissible transformation, the rule under the transformed scale is the same as under the original one. Third, a *permissible* statistic refers to a statistical operation that when applied to a scale, produces the same result as when it is applied to the (admissibly) transformed scale. While the word 'permissible' may sound rather strong, it is justifiable under the premise that 'one should only make assertions that are invariant under admissible transformations of scale' (Marcus-Roberts and Roberts 1987: 384).<sup>30</sup> Fourth, a statement is called *meaningful* if it remains unchanged when all scales in the statement are transformed by admissible transformations (Marcus-Roberts and Roberts 1987: 384).<sup>31</sup>

Stevens (1946) considered four basic empirical operations or rules that define four types of scales: equality, rank order, equality of intervals, and equality of ratios. Following them, he defined four main types of scales: nominal, ordinal, interval, and ratio. Stevens' classification is not exhaustive.

---

<sup>29</sup> Stevens' work belongs to a branch of applied mathematics called measurement theory, which is useful in measurement and data analysis.

<sup>30</sup> 'The criterion for the appropriateness of a statistic is invariance under the [admissible] transformations' (Stevens 1946: 678).

<sup>31</sup> The notion of meaningfulness is alluded to in Stevens (1946) and used in Roberts (1979).

For example, it only applies to scales that take real values and which are regular.<sup>32</sup> Also, note that alternative terms are sometimes used for some of Stevens' types. For example, nominal scales are sometimes referred to as categorical scales.

Table 2.3 lists the scale types mentioned above from 'weakest' to 'strongest' in the sense that interval and ratio scales contain much more information than ordinal or nominal scales. The column that presents the rule defining each scale type is cumulative in the sense that a rule listed for a particular scale must be applicable to the scales in rows preceding it. The column that lists the permissible statistics is also cumulative in the same sense. In contrast, the column that lists the admissible transformations goes from general to particular: the particular operation listed in a row is included in the operation listed above.

We now introduce each scale 'type'. The scale pertains to an indicator used to measure dimension  $j$ . The term 'indicator  $j$ ' denotes the indicator of dimension  $j$ . Achievements in indicator  $j$  across the population are represented by vector  $x_{ij} \in \mathbb{R}_+^n$ , where  $x_{ij}$  is the achievement of person  $i$  in the  $j^{\text{th}}$  indicator.

Indicator  $j$  is said to be **nominal** or **categorical** if the scale is based on mutually exclusive categories, not necessarily ordered. Nominal variables are frequently called **categorical** variables. The rule or basic empirical operation behind this type of scale is the determination of equality among observations. A nominal scale is 'the most unrestricted assignment of numerals. The numerals are used only as labels or type numbers, and words or letters would serve as well' (Stevens 1946: 678). That is, numbers assigned to the various achievement levels in this domain are simply placeholders.

Stevens introduces two common types of nominal variables. One uses 'numbering' for identification, such as the identification number of each household in a survey or the line number of individuals living within a household. The other uses numbering for a classification, such that all members of a social group (ethnic, caste, religion, gender, or age) or geographical regions (rural/urban areas, states, or provinces) are assigned the same number. The first type of nominal variable is simply a particular case of the second. There is a wide range of admissible transformations for this type of scale. In fact, any transformation that substitutes or permutes values between groups,

---

<sup>32</sup> An irregular scale does not always generate an acceptable scale from an admissible transformation (see Roberts and Franke 1976, cited in Marcus-Roberts and Roberts 1987: 384).

that is, any one-to-one substitution function  $f$  such that  $x'_{ij} = f(x_{ij})$  for all  $i$ , will leave the scale form invariant.

Given that in a nominal variable, the different categories do not have an order, neither arithmetic operations nor logical operations (aside from equality) are applicable. In terms of relevant statistics, if the nominal variable is simply an identifier, then only the number of categories is a relevant statistic; if the nominal variable contains several cases in each category, then the mode and contingency methods can be implemented, as can hypotheses tests regarding the distribution of cases among the classes (Stevens 1946: 678–9).

Indicator  $j$  is said to be **ordinal** if the order matters but not the differences between values. The rule or basic empirical operation behind this type of scale is the determination of a rank order. Categories can be ordered in terms of ‘greater’, ‘less’, or ‘equal’ (or ‘better’, ‘worse’, ‘preferred’, ‘not preferred’). Admissible transformations consist of any order-preserving transformation, that is, any strictly monotonic increasing function  $f$  such that  $x'_{ij} = f(x_{ij})$  for all  $i$ , as these will leave the scale form invariant. Thus, admissible transformations include logarithmic operation, square root of the values (non-negative), linear transformations, and adding a constant or multiplying by another (positive) constant. Examples of ordinal scales are preference orderings over various categories, or subjective rankings. Given that the true intervals between the scale points are unknown, arithmetic operations are meaningless (because results will change with a change of scale), but logical operations are possible. For example, we can assert that someone reporting a health level of four feels ‘better’ than someone reporting a health level of ‘three’, who in turn feels better than a ‘two’, but we cannot assert whether the difference between level three and four is the same as the difference between level two and three. Nevertheless, some statistics are applicable to ordinal variables, namely, the number of cases, contingency tables, the mode, median, and percentiles.<sup>33</sup> Statistics such as mean and standard deviation cannot be used. Clearly, an ordinal variable is a nominal variable but the converse is not true. Ordinal and nominal (or categorical) variables are also sometimes referred to as qualitative variables.

---

<sup>33</sup> Although Stevens (1946) considers these statistics as ‘permissible’ with these variables, he warns: ‘although percentiles measures may be applied to rank-ordered data ... the customary procedure of assigning a value to a percentile by interpolating linearly within a class interval is, in all strictness, wholly out of bounds. Likewise, it is not strictly proper to determine the mid-point of a class interval by linear interpolation, because the linearity of an ordinal scale is precisely the property which is open to question’ (Stevens 1946: 679).

Table 2.3 Stevens' classification of scales of measurement

Type of Variable (Scales)	Rule for Creating the Scale	Admissible Transformations	Admissible Mathematical Operations	Admissible Statistics	Particular Cases	Examples Relevant for Poverty Measurement
Qualitative	Nominal or categorical	Determination of equality	Permutation group* $x'_{ij} = f(x_{ij})$ where $f$ is any one-to-one substitution	None	Frequency distribution, mode, contingency correlation	Gender, caste, civil status, ethnicity
	Ordinal	Determination of greater or less	Isotonic group** $x'_{ij} = f(x_{ij})$ where $f$ is any monotonic increasing function	None	Median percentiles	Ordered categorical or weak ordinal Type of source of drinking water, sanitation facility, cooking fuel, floor. Levels of schooling
Quantitative or cardinal	Interval scale	Determination of equality of intervals or differences	General linear group*** $x'_{ij} = a x_{ij} + b$ , with $a > 0$	Add, subtract	Mean, Standard deviation, Rank-order correlation, Product-moment correlation.	Dichotomous variables z-scores of nutritional indicators (ex. weight for age) Body Mass Index
	Ratio scale	Determination of equality of ratios	Similarity group $x'_{ij} = a x_{ij}$ , with $a > 0$	Divide, multiply	Coefficient of variation	Count variables Dichotomous variables Income, consumption expenditure, number of deaths a mother experienced, years of schooling,‡ numbers of Jersey milk cows owned

Source: Stevens (1946). The columns on 'Particular cases' and 'Examples' have been added by the authors.

\* A permutation group here refers to the composition of permutations that can be performed with elements of a group.

\*\* An isotonic group here refers to the set of transformations which preserve order.

\*\*\* A general linear group here refers to the set of linear transformations (with the specifications stated above).

\*\*\*\* A similarity group here refers to the set of ratio transformations (with the specifications stated above).

‡: Note that years of schooling may also be interpreted in an ordinal sense, depending on the meaning attached.

Unordered categorical variables—such as eye colour—are not relevant for the construction of poverty measures. Relevant categorical variables are those that can be exhaustively and non-trivially partitioned into at least two sets according to some exogenous condition, and in which those sets can be arranged in a complete ordering. There will be fewer sets than there are categorical responses, or else the original variable would already have been ordinal. If a set contains multiple elements, it may not be possible to rank those elements against one another. Hence the resulting construction would be a ‘semi-order’ (Luce 1956) or ‘quasi-order’ (Sen 1973). Additionally, it may be possible to distinguish set(s) that are considered to be adequate achievements from those that are inadequate, forming a ‘weak order’, that is, some pairs of responses can be ranked as ‘preferred to’ and some others cannot be ranked.<sup>34</sup> For example, because it is difficult to assess whether it is better to have access to a public tap than to a borehole or a protected well as sources of drinkable water, the Millennium Development Goal indicator considers all three of them to be adequate sources of drinkable water (unrankable). Similarly, while one cannot rank access to an unprotected spring versus access to rainwater, both sources are considered inadequate by MDG standards. A variable thus constructed is often called an ‘ordered categorical’; we might also call the variables obtained as a weak order of categories in a nominal variable, a ‘weak-ordinal’ variable. Admissible transformations of weak ordinal variables include any transformations that partition the categorical variables into the relevant sets (safe water sources) in the same order; any apparent ordering of elements within the relevant sets can vary freely.

Indicator  $j$  is said to be of *interval scale* if the rule or basic empirical operation behind its scale is the determination of equality of intervals or differences. Importantly, interval scales do not have a predefined zero point. The admissible transformations of interval scale consist of the linear transformation  $x'_{ij} = a x_{ij} + b$  (with  $a > 0$ ), as this preserves the differences between categories. While the *difference* between two values of an interval-scale variable is meaningful, the ratios are not. The most cited example in this literature refers to two temperature scales: Celsius (°C) and Fahrenheit (°F). While the difference between 15°C and 20°C is the same as the difference between 20°C and 25°C, one cannot say that 20°C is twice as hot as 10°C because 0°C does not mean ‘no temperature’. That is, the Celsius scale (and Fahrenheit scale) lack a natural zero. Also, the difference between 15°C and 20°C and between 20°C and 25°C is also precisely the same if measured in

---

<sup>34</sup> Relatedly, Luce (1956) distinguished a weak order from a semi-order over the same set of elements. In a weak order the indifference relation is transitive, but in a semi-order it is not.

Fahrenheit (59°F and 68°F vs 68°F and 77°F) although the value of the difference is nine rather than five degrees. An interval scale allows addition and subtraction and the computation of most statistics, namely, number of cases, mode, contingency correlations, median, percentiles, mean, standard deviation, rank-order correlation, and product-moment correlation, but it is not meaningful to compute the coefficient of variation or any other ‘relative’ measure.

In multidimensional poverty measurement, one indicator that is usually of interest is the z-score of under 5-year-old children’s nutritional achievement. We consider a nutritional z-score to be of interval-scale type. Box 2.3 provides a more detailed explanation of how to compute z-scores. Z-scores range from negative to positive values, spaced in (the reference population’s) standard deviation units, and the zero value means that the child’s nutritional achievement is at the median of what is considered a healthy population.

Indicator  $j$  is said to be of **ratio scale** if the rule or basic empirical operation behind its scale is the determination of equality of ratios. Such a rule requires the scale to have a ‘natural zero’: namely the value 0 means ‘no quantity’ of that indicator. In other words, the value 0 is the absolute lowest value of the variable. Admissible transformations of interval-scale variables consist of functions such as  $x'_j = a x_j$  (with  $a > 0$ ), as this preserves the ratio differences. Examples of ratio-scale variables are age, height, weight, and temperature in Kelvin, as 0° Kelvin means ‘no temperature’; 200 pounds is twice as much as 100 pounds, sixty years as thirty years, and so on. Ratio-scale variables allow statements such as ‘a value is twice as large as another’, and they allow any type of mathematical operation, as well as the computation of any statistic (number of cases, mode, contingency correlations, median, percentiles, mean, standard deviation, rank-order correlation, product-moment correlation, and coefficient of variation). Interval- and ratio-scale variables constitute what are commonly referred to as **cardinal** variables.

It is interesting to observe that in the order presented, from nominal to ratio scales, the admissible transformations become more restricted but the meaningful statistics become more unrestricted, ‘suggesting that in some sense the data values carry more information’ (Velleman and Wilkinson 1993: 66). Stevens (1959: 24) provided an insightful example of how measurement can progress from weaker to stronger scales. Early humans probably could only distinguish between cold and warm and thus used a nominal scale. Later, degrees of warmer and colder had been introduced and so the use of an ordinal scale gained prominence. The introduction of thermometers led to the use

of an interval scale. Finally, the development of thermodynamics led to the ratio scale of temperature by introducing the Kelvin scale.

### Box 2.3 Children's nutritional z-scores

The nutritional status of children under 5 years old is assessed with three anthropometric indicators: weight-for-age, also called 'underweight'; weight-for-height, also called 'wasting'; and height-for-age, also called 'stunting'. The indicators are constructed as follows. The World Health Organization has measured the height and weight of a reference population of children from different ethnicities, which is considered to constitute a standard of well nourishment (WHO 2006). From that population, a distribution of weights according to each age, a distribution of heights according to each age, and a distribution of height according to each weight are obtained. Each of these is discriminated by gender.

How is a child's nutrition assessed? Let us consider the case of the weight-for-age ( $w/a$ ) indicator. Once the weight and age of the child have been documented, this information can be expressed in her weight-for-age z-score. This is computed as the child's observed weight minus the median weight of children of the same sex and age in the reference population, divided by the standard deviation of the reference population. That is:

$$zS_{w/a} = \frac{(y_{i,w/a} - \bar{Y}_{w/a})}{\sigma_{w/a}}$$

where  $zS_{w/a}$  is the z-score of weight-for-age,  $y_{i,w/a}$  is the observed weight of child  $i$ ,  $\bar{Y}_{w/a}$  is the median weight of children of the same sex and age as child  $i$  in the reference population (healthy children), and  $\sigma_{w/a}$  is the standard deviation of the weight of children of that age in the reference population.

The z-scores for weight-for-height ( $w/h$ ) and height-for-age ( $h/a$ ) are computed in an analogous way. Thus,  $zS_{i,j} = (y_{i,j} - \bar{Y}_j) / \sigma_j$  for all  $i$  and all  $j = w/a, w/h$ , and  $h/a$ .

Thus, for example, suppose 14-month-old Anna weighs 8.3 kilograms. The median weight in the reference population of children of that age is 9.4 and the standard deviation is 1.\* Thus, the z-score of Anna is  $-1.1$ , meaning that Anna is about one standard deviation below the median weight of healthy children.

It is considered that children with z-scores that are more than two standard deviations below the median of the reference population suffer moderate undernutrition, and if their z-score is more than three standard deviations below, they suffer severe undernutrition (underweight, wasting, or stunting, correspondingly). Children with a z-score of weight-for-height above +2 standard deviations above the median are considered to be overweight (*WHO Programme of Nutrition* 1997).

An alternative way to assess the nutritional status of children is to use percentiles rather than z-scores, but z-scores present a number of advantages. Most importantly, they can be used to compute summary statistics such as a mean and standard deviation, which cannot be meaningfully done with percentiles (O'Donnell et al. 2008).

Note that if we take a linear transformation of the z-score for weight-for-age such that  $zS'_{i,w/a} = a \times zS_{i,w/a} +$

$\ell$ , where  $a = 1/\sigma_{w/a}$  and  $\ell = \bar{Y}_{w/a}/\sigma_{w/a}$ , then  $zS'_{w/a} = y_{i,w/a}$ . Note that the difference  $y_{i,w/a} - y_{i',w/a}$  (or  $zS'_{i,w/a} - zS'_{i',w/a}$ ) has the same implication as the difference  $zS_{i,w/a} - zS_{i',w/a}$ . This equivalence would hold for any linear transformation, exhibiting the characteristics of an interval-scale indicator.

\*These values were taken from WHO's reference tables:

[http://www.who.int/childgrowth/standards/sft\\_wfa\\_girls\\_z\\_0\\_5.pdf](http://www.who.int/childgrowth/standards/sft_wfa_girls_z_0_5.pdf)

Having introduced the scales of measurement, it is worth making a few clarifications regarding other frequently mentioned types of indicators. First, Stevens' classification makes no reference to **continuous** versus **discrete** variables, for example. Continuous variables can take any value on the real line within a range. Discrete variables, in contrast, can only take a finite or countably infinite number of values.<sup>35</sup> Ordinal variables are discrete variables, but cardinal ones (interval and ratio scale) can be either discrete or continuous.<sup>36</sup> Second, note that **count** variables such as counts of publications, number of children in a household, or number of chickens, are particular cases of ratio-scale variables (Stevens 1946), such that the only admissible transformation is the identity function, i.e.  $a = 1$ . Roberts (1979) refers to the counting scale type as an **absolute scale**.

Third, **dichotomous** (also called binary) variables can be of different scales, depending on the meaning of their categories. When the two values simply refer to unordered, mutually exclusive categories, such as being male or female, the variable is of nominal scale. When there is an order between the categories, such as being deprived or not in a specific dimension, the variable is **cardinal**. If the two values refer to having or lacking the same thing, such as a fully functional method for wood smoke ventilation, the variable may be interpreted as of ratio scale.<sup>37</sup>

Fourth, the reader may wonder where do **Likert scales**—introduced by Likert (1932) and often used in social sciences—fit? Likert scales are obtained from responses to a set of (carefully phrased) statements to which each respondent expresses her level of agreement on a scale such as one to five: strongly disagree, disagree, neutral, agree, or strongly agree. Each statement and its responses are known as a Likert item. A Likert scale is obtained by summing or averaging the responses to each

<sup>35</sup> Countably infinite means that the values of the discrete variable have one-to-one correspondence with the natural numbers.

<sup>36</sup> Note that other authors equate the distinction between qualitative/ordinal vs quantitative/cardinal with the distinction between discrete vs continuous variables (e.g. Bossert, Chakravarty, and D'Ambrosio 2013). In our definitions, cardinal variables can be either continuous or discrete, so the two pairs are not equivalent.

<sup>37</sup> Dichotomous variables can also be obtained from nominal ones. For example, given a nominal variable on age intervals, a dummy variable can be created for each age interval ('belongs' or 'does not belong' to that particular age range). More commonly, one can dichotomize variables with categorical responses into deprived and non-deprived states; for example, classifying 'sources of water' into two exhaustive groups reflecting 'safe' and 'unsafe' water.

item so that a score is acquired for each person. Likert scales are frequently treated as interval scales, under the assumption (at times empirically verified) that there is an equal distance between categories (Brown 2011; Norman 2010). Thus, descriptive statistics (like means and standard deviations) and inferential statistics (like correlation coefficients, factor analysis, and analysis of variance) are regularly implemented with Likert scales. However, this has been criticized as being ‘illegitimate to infer that the intensity of feeling between ‘strongly disagree’ and ‘disagree’ is equivalent to the intensity of feeling between other consecutive categories on the Likert scale’ (Cohen et al. 2000, cited in Jamieson 2004: 1217). Thus, there is ongoing disagreement about whether Likert scales should be treated as ordinal scales (Pett 1997; Hansen 2003; Jamieson 2004). Often empirical psychometric tests are performed to ‘verify’ whether the assumption that the scale can be treated as cardinal holds for a particular dataset.

Stevens’ (1946) landmark work sparked a body of literature on measurement theory, which raised strong warnings regarding the applicability of statistics to different scales, including Stevens (1951, 1959), Luce (1959), and Andrews et al. (1981). This engendered an extensive and ongoing debate across literatures from psychology to statistics. Some social scientists and statisticians consider that although, in theory, certain statistics such as the mean and standard deviations are inappropriate for ordinal data, they can still be ‘fruitfully’ applied if the problem in question and data structure (and comparability, if relevant) have been tested and seem to warrant assumptions that they can be treated as interval scale. The arguments in favour of this position are interestingly articulated by Velleman and Wilkinson (1993). For example, it is stated that ‘the meaningfulness of a statistical analysis depends on the questions it is designed to answer’ (Lord 1953; Guttman 1977), that in the end ‘every knowledge is based on some approximation’ (Tukey 1961), and that ‘... if science was restricted to probably meaningful statements it could not proceed’ (Velleman and Wilkinson 1993). Another argument is that parametric methods are highly robust to violation of assumptions and thus can be implemented with ordinal data (Norman 2010). Although we acknowledge this ongoing debate, in what follows we do endorse the requirement of **meaningfulness**—as defined above—in order to compute statistics or mathematical measures in each scale type; hence, we limit the operations with ordinal data to those that cohere with its definition and point out deviations from this.

As this section suggests, the indicators’ scale and the analysts’ considered response to scales of measurement must be articulated before selecting a methodology to measure and assess

multidimensional poverty. In Chapter 3, we clarify whether each method surveyed can be used with ordinal data. In general, when an indicator is of ratio scale, such as income, consumption, or expenditure (all expressed in monetary units), arithmetic operations with elements of the scale are permissible. Thus, it is meaningful to compute the normalized deprivation gaps as defined in Expression (2.2). When all considered indicators are of ratio scale, multidimensional poverty measures based on normalized deprivation gaps can be used. If the indicator is of interval scale, gaps may be redefined appropriately. When indicators are ordinal, measures based on normalized deprivation gaps are meaningless because results would vary under different admissible transformations of the indicators' scales. The Adjusted Headcount Ratio presented in Chapters 5–10 can be meaningfully implemented using ordinal indicators, provided that issues of comparability are addressed, as we will now clarify.

## 2.4 Comparability across People and Dimensions

The last section established the scales of measurement by which we can rigorously compare achievement levels in one variable, and the mathematical and statistical operations that can be performed on that variable. The discussion enabled us to identify the scale of measurement of each single indicator. Yet multidimensional measures seek to compare people's achievements or deprivations *across* indicators, in ways that respect the scale of measurement of each indicator. This is by no means elementary. As Sen (1970) pointed out, cardinally meaningful variables may not necessarily be cardinally comparable—across people or, in multidimensional measurement, across dimensions. This section scrutinizes how these comparisons can legitimately proceed. That is, it takes a step back from the material presented thus far, to make explicit assumptions that have usually been implicit in work on multidimensional poverty measurement.

The issue of comparability across dimensions raised in this section has potentially significant empirical implications for quantitative methods beyond measurement. For example, as Chapter 3 will show, dichotomous data are regularly used in techniques that implicitly attribute cardinal meaning *and* comparability to the 0–1 values from several dimensions. If, in fact, the associated deprivation values of these data differ significantly from one another, then results based on techniques that treat each variable as cardinally equivalent may be affected; such exercises should, strictly, employ the 0- $w$  variables because it is their relative weights or deprivation values that create

cardinal comparability across dimensions. Hence the issues raised in this section, in a preliminary and intuitive way at this stage, have far-reaching implications for multidimensional analyses.

The properties we will present in section 2.5 define certain characteristics that a poverty measure may fulfill. Underlying many properties is an assumption that the poverty measure itself is cardinally meaningful. Based on this assumption, a change in the underlying  $n \times d$  achievement matrix will change the poverty measure in desired and predictable ways according to the properties. In order to generate a cardinally meaningful multidimensional poverty index, it is necessary to treat indicators correctly according to their scale of measurement, as we have seen already. But it is also necessary to compare and aggregate a set of indicators (i) across people and (ii) across dimensions. Neither of these steps is trivial.<sup>38</sup>

It must be recalled that variables used to construct unidimensional poverty measures, as defined in section 2.1, already entail comparisons across themselves as component dimensions and across people and households. In terms of interpersonal comparisons, the same household income level is normally assumed to be associated with the same level of individual welfare or poverty.<sup>39</sup> Components of such measures (sources of income, or consumption / expenditure on different goods) are usually assumed to be additive as cardinal variables having a common unit (usually, a currency like pounds or rupees), hence prices (adjusted where necessary) are used as weights. Equivalence scales may be used to augment comparability across households.

Multidimensional poverty measures, like income poverty measures, entail a basic assumption that the indicators are interpersonally comparable. Additionally, counting-based measures further assume that the same deprivation score is associated with the same level of poverty for different people. This assumption implies that deprivations have been made comparable. Comparability across dimensions must be obtained in order to generate cardinally meaningful deprivation scores and associated multidimensional poverty measures. But how? Multidimensional poverty measures may contain fundamentally distinct components that are not measured in the same units and may have no natural means of conversion into a common variable.

---

<sup>38</sup> There is a very large literature on interpersonal comparisons and partial comparisons, stemming from Sen (1970). Basu (1980) raises comparability across dimensions in the context of government preferences and helpfully distinguishes comparability and measurability (ch. 6, 74–5).

<sup>39</sup> Sen (1979, 1985, 1992, 1997) has powerfully observed how the same level of resources may in fact be associated with different levels of well-being because of differences in people's ability to convert resources into well-being.

Empirically, the mechanics by which apparent comparability has been created in counting-based measures are clear (Chapter 4). When data are dichotomous and interpersonally comparable, the application of *deprivation values* is understood to create cardinal comparability across dimensions. When data are ordinal or ordered categorical, *deprivation cutoffs* are used to dichotomize the data; and those cutoffs, together with *deprivation values*, establish cardinal comparability across deprivations. In the case of appropriately scaled cardinal data, comparability across deprivations is created by the *weights* and the *deprivation cutoffs*.

Let us start with the most straightforward case: the deprivation matrix. This presents dichotomous values either because the original indicator is dichotomous (access to electricity), or because a cardinal, ordinal, or ordered categorical variable has been dichotomized by the application of a deprivation cutoff into two categories: deprived and non-deprived.

As mentioned above, we can consider dichotomous deprivations to be (trivially) cardinal.<sup>40</sup> More precisely, the deprivation cutoff establishes a ‘natural zero’ in the sense that any person whose achievement meets or exceeds the natural zero is non-deprived, and anyone who does not, is deprived. We require the natural zeros to be comparable states across dimensions for the purposes of poverty measurement. But how are the ‘one’ values, reflecting deprived states, comparable? Deprivation values (the vector of relative weights) create an explicit value for each of the deprived states, across the set of possible finite dimensional comparisons. After the deprivation values have been applied, deprivations may be cardinally compared across dimensions.

The reason we draw the reader’s attention at this stage to the assumptions underlying cardinal comparisons across people and across dimensions is in order that they might observe how differently multidimensional poverty measurement techniques, such as those surveyed in Chapter 3, undertake and justify such comparisons, and also so that some readers might be encouraged to explore these important issues further—both theoretically and empirically.

---

<sup>40</sup> Earlier we defined a cardinal variable to be interval-scale type if the rule or basic empirical operation behind its scale is the determination of equality of intervals or differences. Consider a variable having exactly two points, neither of which is a natural zero. In this case, they can be understood to be equally spaced along any scale, hence trivially satisfy this definition. If either of the points occurs at a natural zero then the dichotomous variable is ‘trivially ratio scale’.

## 2.5 Properties for Multidimensional Poverty Measures

In selecting one poverty measurement methodology from a set of options, a policymaker thinks through how a poverty measure should behave in different situations in order to be a ‘good’ measure of poverty and support policy goals. Then she asks which measure meets these requirements. For example, should the poverty measure increase or decrease if the achievement of a poor person rises while the achievements of other people remain unchanged? Should poverty comparisons change when achievements are expressed in different units of measurement? Should the measure of poverty in a more populous country with a larger number of poor people be higher than the poverty measure in a small country with a smaller number of, but proportionally more, poor people?

A policymaker seeking to ameliorate poverty should have a good understanding of the various normative principles that her chosen poverty measure embodies, just as a pilot of a plane must have a sound understanding of how a particular plane responds to different operations. If the policymaker has a good understanding of the principles embodied by alternative measures then she will be able to choose the measure most closely reflecting publicly desirable ethical principles and most appropriate for the application—just as a pilot will choose the best way to fly the plane for a particular journey.

The normative judgements embodied by a poverty measure are reflected in its mathematical properties, including its structure and its response to changes in its argument. The axiomatic method, formally introduced to this field by Sen (1976),<sup>41</sup> refers to measures that have been designed based on principles that are taken by the researcher as **axioms**, i.e. as ‘statements that are accepted true without proof.’<sup>42</sup> In this section, we define and discuss the various properties proposed in the

---

<sup>41</sup> Watts (1968) offered an early intuitive (non-formal) justification for selecting the functional form of a poverty measure according to the properties it should satisfy.

<sup>42</sup> Within the poverty measurement literatures, there are essentially two procedures for constructing measures in the axiomatic framework. In the first, known as **characterization**, a number of principles that are considered desirable are introduced and then the *entire* class of measures (one or many) that embody these principles is determined. This procedure entails a sufficiency condition, which shows that the measure satisfies these principles, and, simultaneously, a necessity condition, which shows that this is the *only* measure (or the family of measures) that satisfies the set of desirable principles. Studies that follow this procedure include Sen (1976), Tsui (2002), Chakravarty, Mukherjee, and Ranade (1998), Bossert, Chakravarty, and D’Ambrosio (2013), Chakravarty, Deutsch and Silber (2008), Bossert, Chakravarty, and D’Ambrosio (2013), Hoy and Zheng (2011), and Porter and Quinn (2013). Second, studies may introduce a number of properties that are considered desirable and then propose a measure or family of measures satisfying these properties, without claiming it to be the only measure or family of measures to do so. Studies following this procedure include

literature on multidimensional poverty measures. A consensus may emerge around some properties, thus they may be considered as axioms, while others may remain optional. In what follows, we use the words ‘principles’ and ‘properties’ interchangeably.

The set of properties for multidimensional measurement has been built upon its unidimensional counterpart. However, as noted by Alkire and Foster (2011a), there is one vitally important difference: in the multidimensional context, the identification step is no longer elementary and properties must be viewed as restrictions on the overall poverty methodology  $\mathcal{M} = (\rho, P)$ , that is, as **joint restrictions** on the identification and the aggregation method. For simplicity, in what follows, we state the properties in terms of the poverty index  $P(X; z)$ , which entails two assumptions. First, that any multidimensional poverty index is associated with an identification function. Second, that the identification function relies on a functional form of the type  $\rho(x_i; z)$ .<sup>43</sup>

We classify the properties into four categories.<sup>44</sup> The first set of properties requires that a poverty measure should not change under certain transformations of the achievement matrix. We refer to these as **invariance properties**, which are **symmetry**, **replication invariance**, **scale invariance** and two alternative focus properties, **poverty focus** and **deprivation focus**. The second set requires poverty to either increase or decrease with certain changes in the achievement matrix. We refer to these as **dominance properties**, which are **monotonicity**, **transfer**, **rearrangement**, and **dimensional transfer** properties. The third set of principles relates overall poverty to either groups of people or groups of dimensions and thus is called **subgroup properties**. Other properties, that guarantee that the measure behaves within certain usual, convenient parameters, we refer to as **technical properties**. Each of the four following sections provides a formal outline and intuitive interpretations of each set of properties.

### 2.5.1 Invariance Properties

The first invariance principle is **symmetry**. Symmetry requires that each person in a society is treated anonymously so that only deprivations matter and not the identity of the person who is deprived. Hence this property is also often referred to as **anonymity**. As long as the deprivation

---

Bourguignon and Chakravarty (2003), Calvo and Dercon (2009), Foster (2009), Alkire and Foster (2011a), and Foster and Santos (2013).

<sup>43</sup> Other possible identification methods may violate some of the properties stated in this section. Future research may develop a set of properties for the identification function in the multidimensional context.

<sup>44</sup> This classification follows Foster (2006).

profile of the entire society remains unchanged, swapping achievement vectors across people should not change overall poverty.<sup>45</sup> This type of rearrangement can be obtained by pre-multiplying the achievement matrix by a permutation matrix of appropriate order.

**Symmetry:** If an achievement matrix  $X'$  is obtained from achievement matrix  $X$  as  $X' = \Pi X$ , where  $\Pi$  is a permutation matrix of appropriate order, then  $P(X'; z) = P(X; z)$ .

**Example:** Suppose the initial achievement matrix is  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and the permutation matrix is  $\Pi = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ . Then  $X' = \Pi X = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 4 \\ 4 & 4 & 2 \\ 8 & 6 & 4 \end{bmatrix}$ . Note that

the first and the second person (rows) in matrix  $X$  have swapped their positions. However, as long as the deprivation cutoffs remain unchanged, there is no reason for the level of poverty to be different for these two societies. Hence,  $P(X'; z) = P(X; z)$ .

The second invariance principle, **replication invariance**, requires that if the population of a society is replicated or cloned with the same achievement vectors a finite number of times, then poverty should not change.<sup>46</sup> In other words, the replication invariance property requires the level of poverty in a society to be standardized by its population size so that societies with different population sizes are comparable to each other, as are societies whose populations change over time. Thus, this property is also known as the **principle of population**.<sup>47</sup>

**Replication Invariance:** If an achievement matrix  $X'$  is obtained from another achievement

<sup>45</sup> The principle of anonymity should not be misunderstood as making poverty measurement blind to subgroup inequalities or identities. Often disadvantaged groups are discriminated against because of their religion, ethnicity, gender, age, or some other characteristic. In such cases, two people, one from the disadvantaged group and the other from a non-disadvantaged group, may have exactly the same deprivation profile in the  $d$  considered dimensions. Yet one person from the disadvantaged group may experience an *effectively* higher deprivation than another. Such situations can be addressed in different ways. One way is analysing poverty by subgroup (this is where the subgroup consistency and decomposability properties addressed below become useful). Another is to incorporate another dimension that reflects the disadvantage of this group such that the deprivation profiles of the two people are no longer equal.

<sup>46</sup> This principle was first suggested by Dalton (1920) in the context of inequality measurement.

<sup>47</sup> In the context of welfare measurement, Foster and Sen (1997) referred to this as the 'symmetry for population'. Chakravarty, Mukherjee, and Ranade (1998), Bourguignon and Chakravarty (2003), and Deutsch and Silber (2005) call it the 'principle of population'. Bossert, Chakravarty, and D'Ambrosio (2013) introduce a separate principle called the 'poverty Wicksell population principle' to compare societies with different population sizes. This property requires that if a person is added to the society with the same level of poverty as the aggregate poverty of the society, overall poverty should not change.

matrix  $X$  by replicating  $X$  a finite number of times, then  $P(X'; z) = P(X; z)$ .

**Example:** Suppose we are required to compare the level of poverty in two societies with the

initial achievement matrices  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and  $Y = \begin{bmatrix} 8 & 8 & 8 \\ 9 & 9 & 9 \end{bmatrix}$ . Note that the population

sizes in these societies are different. In order to make them comparable, we replicate  $X$  twice to obtain

$X' = \begin{bmatrix} 4 & 3 & 8 & 4 & 3 & 8 \\ 4 & 5 & 6 & 4 & 5 & 6 \\ 2 & 4 & 4 & 2 & 4 & 4 \end{bmatrix}^{\text{tr}}$  and  $Y$  thrice to obtain  $Y' = \begin{bmatrix} 8 & 9 & 8 & 9 & 8 & 9 \\ 8 & 9 & 8 & 9 & 8 & 9 \\ 8 & 9 & 8 & 9 & 8 & 9 \end{bmatrix}^{\text{tr}}$ . By

replication invariance, we know that  $P(X; z) = P(X'; z)$  and  $P(Y; z) = P(Y'; z)$ . Thus, it is equivalent comparing  $X$  to  $Y$  and  $X'$  to  $Y'$ .

The third invariance principle, **scale invariance**,<sup>48</sup> requires that the evaluation of poverty should not be affected by merely changing the scale of the indicators. For example, if the duration of completed schooling is an indicator, then deprivation in education, thus overall poverty, should be the same regardless of whether duration is measured in years or in months, provided the deprivation cutoff is correspondingly adjusted.

The scale of any indicator in an achievement matrix can be altered by post-multiplying the achievement matrix by a diagonal matrix  $\Lambda$  of appropriate order ( $d$ , the number of dimensions). If a diagonal element is equal to one, then the scale of the respective indicator does not change. The diagonal elements of  $\Lambda$  need not be the same because different indicators may have different scales and units of measurement. A weaker version of the scale invariance principle, referred to as ‘unit consistency’, has been proposed by Zheng (2007) in the context of unidimensional poverty measurement and extended to the multidimensional context by Chakravarty and D’Ambrosio (2013). This principle requires that poverty comparisons, but not necessarily poverty values, should

---

<sup>48</sup> Most of the studies, such as Chakravarty, Mukherjee, and Ranade (1998), Bourguignon and Chakravarty (2003), and Deutsch and Silber (2005), have used the term ‘scale invariance’, whereas Tsui (2002) uses the term ‘ratio-scale invariance’.

not change if the scales of the dimensions are altered.<sup>49</sup> The scale invariance property implies the unit consistency property, but the converse does not hold.

**Scale Invariance:** If an achievement matrix  $X'$  is obtained by post-multiplying the achievement matrix  $X$  by a diagonal matrix  $\Lambda$  such that  $X' = X\Lambda$ , and the deprivation cutoff vector  $z'$  is obtained from  $z$  such that  $z' = z\Lambda$ , then  $P(X'; z') = P(X; z)$ .

**Unit Consistency:** For two achievement matrices  $X$  and  $X''$  and two deprivation cutoff vectors  $z$  and  $z''$ , if  $P(X''; z'') < P(X; z)$ , then  $P(X''\Lambda; z''\Lambda) < P(X\Lambda; z\Lambda)$ .

**Example:** Suppose the initial achievement matrix is  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and the deprivation

cutoff vector is  $z = [5, 6, 4]$ . Matrix  $X$  is post-multiplied by the diagonal matrix

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 2.2 \end{bmatrix} \text{ to obtain } X' = X\Lambda = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 2.2 \end{bmatrix} = \begin{bmatrix} 4 & 48 & 4.4 \\ 3 & 60 & 8.8 \\ 8 & 72 & 8.8 \end{bmatrix}.$$

Similarly, vector  $z$  is post-multiplied by  $\Lambda$  to obtain  $z' = [5, 72, 8.8]$ . Note that the first indicator and its deprivation cutoff did not change at all because the first diagonal element in  $\Lambda$  is 1. The second diagonal element in  $\Lambda$  is 12, as in 1 year = 12 months, and the third diagonal element in  $\Lambda$  is 2.2, as in 1 kilogram = 2.2 pounds. Scale invariance requires that this transformation does not change overall poverty. Hence,  $P(X'; z') = P(X; z)$ .

Now suppose that there exists a hypothetical achievement matrix  $X''$  such that  $P(X''; z) < P(X; z)$ . Then, the unit consistency property requires that the comparison does not change if both matrices and the vector  $z$  of deprivation cutoffs are multiplied by a diagonal matrix. Hence,  $P(X''\Lambda; z\Lambda) < P(X\Lambda; z\Lambda)$ .

---

<sup>49</sup> Both the scale invariance and the unit consistency principles refer to cases in which achievements are changed in a certain proportion (which may differ or not across achievements). A different principle known as 'translation invariance', popularized by Kolm (1976a,b), requires the poverty level to remain the same if each achievement and its corresponding deprivation cutoff are changed by adding the same constant for every person (although the constant added can differ across dimensions). Technically, if an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  so that  $X' = X + \Gamma$ , where  $\Gamma \in \mathbb{R}_+^{n \times d}$  and  $\Gamma_i = \Gamma_{i'}$  for all  $i \neq i'$ , and  $z' = z + \Gamma_i$ , then  $P(X'; z') = P(X; z)$ .

The fourth invariance principle is **focus**. The primary difference between the measurement of welfare and inequality and the measurement of poverty is that the latter is concerned with the base or bottom of the distribution whereas welfare and inequality are concerned with the entire distribution. The focus principle is crucial because it requires the poverty measure to respond only to the achievements of the poor. The **poverty focus** principle requires that poverty should not change if there is an improvement in any achievement of a non-poor person. This principle is a natural extension of the focus principle in the single dimensional context, which requires that overall poverty remains unchanged with an increase in the income of the non-poor.

**Poverty Focus:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  such that  $x'_{ij} > x_{ij}$  for some pair  $(i, j) = (i', j')$  where  $i' \notin Z$ , and  $x'_{ij} = x_{ij}$  for every other pair  $(i, j) \neq (i', j')$ , then  $P(X'; z) = P(X; z)$ .

However, in the multidimensional framework, the terms ‘deprived’ and ‘poor’ are not synonymous. Someone can be poor yet not deprived in every single indicator. The poverty focus principle does not cover the situation where a poor person’s achievement increases in a dimension in which that person is not deprived. If poverty were to change in this case, the non-deprived attainments would compensate deprived attainments, creating a form of substitutability across dimensions. Practically, this could encourage a policymaker who is enthusiastically interested in reducing the poverty figures to assist the poor in their non-deprived dimensions, instead of addressing the dimensions in which they are deprived. Thus, a second focus principle might be relevant. The **deprivation focus** principle requires that overall poverty not change if there is an increase in achievement in a non-deprived dimension, regardless of whether it belongs to a poor or a non-poor person. This property prevents poverty from falling when a poor person’s achievements increase in non-deprived dimensions. Note that focus properties motivate the construction of a censored achievement matrix  $\tilde{X}$ , as described in section 2.1.1.<sup>50</sup>

**Deprivation Focus:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  such that  $x'_{ij} > x_{ij}$  for some pair  $(i, j) = (i', j')$  where  $x_{ij} \geq z_j$  (whether or not  $i'$  happens to be poor), and  $x'_{ij} = x_{ij}$  for every other pair  $(i, j) \neq (i', j')$ , then  $P(X'; z) = P(X; z)$ .

---

<sup>50</sup> Bourguignon and Chakravarty (2003) refer the deprivation focus as ‘strong focus’ and the poverty focus as ‘weak focus’. Chakravarty, Mukherjee, and Ranade (1998) and Tsui (2002) only used the deprivation focus and did not consider the poverty focus.

**Example:** Suppose the initial achievement matrix is  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and the deprivation cutoff

vector is  $z = [5, 6, 4]$ . Clearly, the third person is not deprived in any dimension, thus cannot be

considered poor. If matrix  $X'$  is obtained from  $X$  such that  $X' = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 5 \end{bmatrix}$ , then by poverty

focus,  $P(X'; z) = P(X; z)$ . The second person is deprived in two out of three dimensions. If the second person is considered poor and  $X''$  is obtained from  $X$  by increasing the person's

achievement in the non-deprived dimension (third dimension) such that  $X'' = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 5 \\ 8 & 6 & 4 \end{bmatrix}$ , then

by the deprivation focus principle,  $P(X''; z) = P(X; z)$ .

The focus principle is one example in which it can be verified that the properties of multidimensional poverty measures are, as stated at the beginning of this section, joint restrictions on the identification and the aggregation methods. For example, for the deprivation focus principle to be satisfied, the identification method cannot follow the aggregate achievement approach. Also, as Alkire and Foster note (2011a: 481), the relevance of the two focus principles is connected to the criterion used to identify the poor (within a counting approach to identification). When a union criterion is used to identify the poor, the deprivation focus principle implies the poverty focus principle, whereas when an intersection criterion is used to identify the poor, the poverty focus principle implies the deprivation focus principle. When intermediate criteria are used, neither of the two principles implies the other.

The last invariance property **ordinality** relates to the type of scale of the particular indicator used for measuring each dimension. As we explained in section 2.3, the scale of an indicator affects the type of operations and statistics that can be meaningfully applied, in the sense that statements remain unchanged when all scales in the statement are transformed by admissible transformations. Building

on this notion of meaningfulness, ordinality requires the poverty estimate not to change under admissible transformations of the scales of the indicators that compose the poverty measure.<sup>51</sup>

More formally, we say that  $(X'; z')$  is obtained from  $(X; z)$  as an *equivalent representation* if there exist appropriate admissible transformations  $f_j: R_+ \rightarrow R_+$  for  $j=1, \dots, d$  such that  $x'_{ij} = f_j(x_{ij})$  and  $z'_j = f_j(z_j)$  for all  $i=1, \dots, n$ . In other words, an equivalent representation can comprise a set of admissible transformations, each of which is appropriate for each scale type, and which assigns a different set of numbers to the same underlying basic data. For example, an equivalent representation can include monotonic increasing transformations for ordinal variables, linear transformations for interval-scale variables, and proportional changes for ratio-scale variables. We now state the ordinality axiom as follows:

**Ordinality:** Suppose that  $(X'; z')$  is obtained from  $(X; z)$  as an equivalent representation, then  $P(X'; z') = P(X; z)$ .

**Example:** Suppose the initial achievement matrix is  $X = \begin{bmatrix} 4,100 & 2 & 0 \\ 3,500 & 3 & 0 \\ 8,200 & 4 & 1 \end{bmatrix}$  and the

deprivation cutoff vector is  $z = [5,000, 4, 1]$ , where the first dimension is measured with a ratio-scale indicator, say income in dollars \$; the second dimension is measured with an ordinal-scale indicator, say self-rated health on a scale of one to five; and the third indicator is measured with a dichotomous 0–1 variable, say access to electricity. Suppose that income is now expressed in some other currency  $D$ , such that  $D = \$0.5$  (note that  $x'_{ij} = 0.5x_{ij}$  in this case), self-rated health is now expressed with the scale (1, 4, 9, 16, and 25) (note that  $x'_{ij} = (x_{ij})^2$ ), and access to electricity is now coded as 1 for no access and 2 for access (i.e.  $x'_{ij} = x_{ij} + 1$ ). Thus matrix  $X'$  obtained from  $X$  is such that

$X' = \begin{bmatrix} 2,050 & 4 & 1 \\ 1,750 & 9 & 1 \\ 4,100 & 16 & 2 \end{bmatrix}$ , and  $z = [2,500, 16, 2]$ , then the ordinality property requires that

$P(X'; z') = P(X; z)$ .

---

<sup>51</sup> A nice theorem would be to prove that the only poverty measure invariant to admissible transformations of the nominal or ordinal variables is one based on dichotomous variables.

### 2.5.2 Dominance Properties

This section covers six principles, each of which has a stronger version and a weaker version. The stronger version requires that a poverty measure strictly moves in a particular direction, given certain transformations in the achievements of the poor. The weaker version does not *require* a poverty measure to move in a particular direction but ensures that the poverty measure does not move in the opposite (wrong) direction under certain transformations of the achievements.<sup>52</sup> The first dominance principle, **monotonicity**, requires that if the achievement of a poor person in a deprived dimension increases while other achievements remain unchanged, then overall poverty should decrease. Normatively, this principle considers that improvements in deprived achievements of the poor are good and should be reflected by producing a reduction in poverty. Its weaker version, referred to as weak monotonicity, ensures that poverty should not increase if there is an increase in any person's achievement in the society.<sup>53</sup>

**Monotonicity:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  such that  $x'_{ij} < \min\{x_{ij}, z_j\}$  for some pair  $(i, j) = (i', j')$  where  $i' \in Z$ , and  $x'_{ij} = x_{ij}$  for every other pair  $(i, j) \neq (i', j')$ , then  $P(X'; z) < P(X; z)$ .

**Weak Monotonicity:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  such that  $x'_{ij} > x_{ij}$  for some pair  $(i, j) = (i', j')$  and  $x'_{ij} = x_{ij}$  for every other pair  $(i, j) \neq (i', j')$ , then  $P(X'; z) \leq P(X; z)$ .

**Example:** Suppose the initial achievement matrix is  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and the deprivation cutoff vector is  $z = [5, 6, 4]$ . Clearly, the first person is deprived in all three dimensions

---

<sup>52</sup> In the unidimensional case, Chakravarty (1983a) defined the weak version of the transfer principle differently (as requiring poverty to change in a particular dimension under the relevant transformation of the achievement but such that the number of poor people did not change). In this book the weak version of a principle differs from the strong one in stating a weak inequality as opposed to a strict one.

<sup>53</sup> Alkire and Foster (2011a) distinguished the monotonicity principle from the weak monotonicity principle. Others, including Chakravarty, Mukherjee, and Ranade (1998), Tsui (2002), Bourguignon and Chakravarty (2003), and Deutsch and Silber (2005) imply weak monotonicity by their monotonicity principle. Bossert, Chakravarty, and D'Ambrosio (2013) did not introduce a weak monotonicity principle.

and is considered poor. If matrix  $X'$  is obtained from  $X$  such that  $X' = \begin{bmatrix} 4 & 4 & 3 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$ , then

by the monotonicity principle,  $P(X'; z) < P(X; z)$ . On the other hand, weak monotonicity requires that  $P(X'; z) \leq P(X; z)$ ; it ensures that poverty does not increase due to the increase in achievement.

The monotonicity and weak monotonicity principles are natural extensions of the analogous concepts in the unidimensional poverty analysis. However, the next principle in this category, referred to as **dimensional monotonicity**, is specific to the multidimensional context. This principle was introduced by Alkire and Foster (2011a). Dimensional monotonicity requires that if a poor person, who is not deprived in all dimensions, becomes deprived in an additional dimension then poverty should increase. This principle ensures that we are not only concerned with the number of poor in a society but also with the extent to which the poor are deprived in multiple dimensions—what we call the intensity of their deprivation. A measure that satisfies monotonicity also satisfies dimensional monotonicity.

**Dimensional Monotonicity:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  such that  $x'_{ij} < z_j \leq x_{ij}$  for some pair  $(i, j) = (i', j')$  where  $i' \in Z$  and  $x'_{ij} = x_{ij}$  for every other pair  $(i, j) \neq (i', j')$ , then  $P(X'; z) > P(X; z)$ .

**Weak Dimensional Monotonicity:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  such that  $x'_{ij} < z_j \leq x_{ij}$  for some pair  $(i, j) = (i', j')$  where  $i' \in Z$  and  $x'_{ij} = x_{ij}$  for every other pair  $(i, j) \neq (i', j')$ , then  $P(X'; z) \geq P(X; z)$ .

**Example:** Suppose the initial achievement matrix is  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and the deprivation

cutoff vector is  $z = [5, 6, 4]$ . Suppose the second person is identified as poor by some criterion but is not deprived in all dimensions. If matrix  $X'$  is obtained from  $X$  by the second person becoming deprived in a principle additional dimension such that

$X' = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 3 \\ 8 & 6 & 4 \end{bmatrix}$ , then by dimensional monotonicity,  $P(X'; z) > P(X; z)$ .

The third principle in the category of dominance principles, **transfer**, is concerned with inequality among the poor. This property has been borrowed from the multidimensional inequality measurement literature, and governs how a poverty measure should behave when the distribution of achievements among the poor becomes more or less equal while their average achievements remain the same. There are different ways of reducing inequality within a multidimensional distribution (see Marshall and Olkin 1979). We follow the approach known as **uniform majorization** Kolm (1977). A uniform majorization among the poor is a transformation in which the achievements among the poor are averaged across them or, equivalently, the original bundles of achievements of poor individuals are replaced by a convex combination of them. It is worth noting that the ‘averaging’ occurs within dimensions, across people.<sup>54</sup> Mathematically, this transformation is obtained by pre-multiplying the achievement matrix by a bistochastic matrix. The transfer principle requires that if an achievement matrix is obtained from another achievement matrix by reducing inequality among the poor, while the average achievement among the poor remains the same, then poverty decreases.<sup>55</sup> The **weak transfer** principle ensures that poverty does not increase when achievements among the poor become more equal.

**Transfer:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  such that  $X' = BX$  where  $B$  is a bistochastic matrix which is not a permutation or an identity matrix and  $B_{ii} = 1$  for all  $i \notin Z$ , then  $P(X'; z) < P(X; z)$ .

**Weak Transfer:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  such that  $X' = BX$  where  $B$  is a bistochastic matrix which is not a permutation or an identity matrix and  $B_{ii} = 1$  for all  $i \notin Z$ , or  $P(X'; z) \leq P(X; z)$ .

---

<sup>54</sup> See Fleurbaey (2006a) and Duclos et al. (2011) for discussion on axioms based on uniform majorization.

<sup>55</sup> Note that it is not possible for a multidimensional poverty measure to satisfy the deprivation focus principle and the transfer principle simultaneously (Tsui 2002). For example, suppose the initial achievement matrix is  $X = \begin{bmatrix} 2 & 3 & 8 \\ 2 & 3 & 10 \end{bmatrix}$  and the deprivation cutoff vector is  $z = [5, 6, 4]$  and both of them are identified as poor by some criteria. Consider the bistochastic matrix  $B = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ . Then,  $X' = BX = \begin{bmatrix} 2 & 3 & 9 \\ 2 & 3 & 9 \end{bmatrix}$ . The transfer principle now requires that  $P(X'; z) < P(X; z)$ , but by the deprivation focus principle, we should have  $P(X'; z) = P(X; z)$ .

**Example:** Suppose the initial achievement matrix is  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and the deprivation

cutoff vector is  $z = [5, 6, 4]$ . The first person is indeed poor—being deprived in all three dimensions. Suppose the second person is also identified as poor by some criterion.

Consider the bistochastic matrix  $B = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ , where the achievements are

distributed equally across the two poor persons so that  $X' = \begin{bmatrix} 3.5 & 4.5 & 3 \\ 3.5 & 4.5 & 3 \\ 8 & 6 & 4 \end{bmatrix}$ . The transfer

principle requires that  $P(X'; z) < P(X; z)$ , whereas the weak transfer principle requires that  $P(X'; z) \leq P(X; z)$ .

The transfer principle in the multidimensional context is similar to its unidimensional counterpart, which is also concerned with the spread of the distribution. There is a second form of inequality among the poor that is only relevant in the multidimensional context, and depends on how dimensional achievements are associated across the population. This second form of inequality corresponds to the joint distribution of achievements, and was introduced by Atkinson and Bourguignon (1982): ‘in the study of multiple deprivation, investigators have been concerned with the ways in which different forms of deprivation (...) tend to be associated ...’ (p. 183). Authors working on this issue have used both the term ‘correlation’ and the term ‘association’. Correlation refers to the degree of linear association between two variables, whereas association is a broader term that includes linear association and also encompasses other forms of association such as quadratic or simply rank association.<sup>56</sup> Given a monotonic transformation of a variable, it is possible that while some form of association, such as rank association, remains invariant, the degree of correlation changes. Thus, here we prefer to use the broader concept of association to define the related properties.

---

<sup>56</sup> Rank association refers to the degree of agreement between two rankings. In the context of the properties discussed here, perfect rank association would occur if person  $i'$ , having higher achievement than person  $i$  in dimension  $j'$ , also has higher achievements in all the other dimensions  $j \neq j'$ . That is:  $x_{i'j} \geq x_{ij}$  for all  $j = 1, \dots, d$ .

The principles that require a measure to be sensitive to the association between dimensions refer to a specific type of rearrangement of the achievements across the population that we call ‘association decreasing rearrangement’.<sup>57</sup> The intuition is as follows. Imagine that originally person  $i'$  is at least as well off in all dimensions as person  $i$ . Then, there is a switch in the achievement of one or more dimensions, but not in all dimensions, between the two persons such that, after the switch, person  $i'$  no longer has achievements equal to or higher than person  $i$  in *all* dimensions but only in some. Suppose also that the achievements of everyone else remain unchanged. Such a transformation constitutes an association-decreasing rearrangement. Formally, given two persons  $i$  and  $i'$  in  $Y \in \mathcal{X}_n$  such that  $y_{i'j} \leq y_{ij}$  for all  $j$ , if matrix  $X$  is obtained from  $Y$  such that  $x_{ij} = y_{i'j}$  and  $x_{i'j} = y_{ij}$  for some dimension  $j$  and  $x_{i'j} = y_{i'j}$  for all  $i'' \neq i, i'$  and all  $j' \neq j$ , and  $X$  is not a permutation of  $Y$ , then  $X$  is stated to be obtained from  $Y$  by an association-decreasing rearrangement. The requirement that  $X$  is not a permutation of  $Y$  prevents the switch from taking place in dimensions where both people have equal achievements.

**Example:** Suppose the initial achievement matrix is  $Y = \begin{bmatrix} 3 & 4 & 2 \\ 4 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$ . Suppose  $X$  is

obtained from  $Y$  so that  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$ . Note that the persons in the first two rows

have switched their achievements in the first dimension so that the person in the first row no longer has a lower achievement than the next person in all three dimensions. Thus,  $X$  is obtained from  $Y$  by an association-decreasing rearrangement. In fact, note that the achievement vectors of the first and the second person are comparable by vector dominance in  $Y$  (each element in one vector is equal to or greater than the same element in the other vector) but not in  $X$ . Hence, the overall association between dimensions in  $X$  is lower.

To be relevant to the analysis of poverty, there is one further qualification: the transformation must take place among two poor persons. Thus, in the example above, both the first and the second person must have been identified as poor in order for this transformation to affect the poverty

---

<sup>57</sup> This transformation was motivated by Boland and Proschan (1988).

measure. For example, if the deprivation cutoff vector  $z = [5, 6, 4]$ , then initially in  $Y$  the first person is deprived in three dimensions, the second person is deprived only in the first two dimensions, and the third person is non-deprived in all three dimensions. Suppose that the identification function identifies the first two persons as poor. When  $X$  is obtained from  $Y$ , the first and the second person have switched their achievements in the first dimension so that now the first person is no longer more deprived than the second person in all three dimensions. Thus, in this case,  $X$  is said to be obtained from  $Y$  by an association-decreasing rearrangement *among the poor*.<sup>58</sup>

Should poverty increase or decrease due to an association-decreasing rearrangement among the poor? One possible intuitive view is that poverty should go down or at least not increase because the association-decreasing rearrangement seems to reduce inequality among the poor. In the numerical example above, the first person was originally more deprived in all dimensions than the second person, and after the rearrangement, she is less deprived in one dimension. This was the argument provided by Tsui (2002). However, in line with Atkinson and Bourguignon (1982), Bourguignon and Chakravarty (2003) argued that the change in overall poverty should be contingent on the relation between dimensions, namely, whether they are substitutes or complements. When dimensions are thought to be substitutes for one another, poverty should not increase under the association-decreasing rearrangement. The intuition is that if dimensions are substitutes, an association-decreasing rearrangement helps both people compensate for their meagre achievements in some dimensions with higher achievements in other, a capacity that was limited for one of them before the rearrangement. When indicators are thought to be complements, poverty should not decrease under the described transfer. The intuition is that the association-decreasing rearrangement has reduced the ability of one of the persons to combine achievements and reach a certain level of well-being.<sup>59</sup> Based on the arguments above, the following properties can be defined. As with the case of monotonicity and transfer, we may define a strong and a weak version of each of the properties.

---

<sup>58</sup> Note that if  $Y$ , on the contrary, is obtained from  $X$ , then it is called ‘basic rearrangement’ by Boland and Proschan (1988). In multidimensional poverty measurement, it is referred to as ‘basic rearrangement-increasing transfer’ by Tsui (2002), ‘correlation increasing switch’ by Bourguignon and Chakravarty (2003), and ‘correlation increasing arrangement’ by Deutsch and Silber (2008). In multidimensional welfare analysis, an analogous concept has been called ‘association increasing transfer’ (Seth 2013), and in multidimensional inequality analysis it has been called ‘correlation increasing transfer’ by Tsui (1999) and ‘unfair rearrangement principle’ by Decancq and Lugo (2012).

<sup>59</sup> In the multidimensional measurement literature the substitutability and complementarity relationship between indicators is defined in terms of the second cross-partial derivative of the poverty measure with respect to any two dimensions being positive or negative. This obviously requires the dimensions to be cardinal and the poverty measure to be twice differentiable. Practically, given two dimensions  $j$  and  $j'$ , substitutability implies that poverty decreases less

**Weak Rearrangement (Substitutes):** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by an association-decreasing rearrangement among the poor, then  $P(X'; z) \leq P(X; z)$ .

**Converse Weak Rearrangement (Complements):** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by an association-decreasing rearrangement among the poor, then  $P(X'; z) \geq P(X; z)$ .

**Strong Rearrangement (Substitutes):** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by an association-decreasing rearrangement among the poor, then  $P(X'; z) < P(X; z)$ .

**Converse Strong Rearrangement (Complements):** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by an association-decreasing rearrangement among the poor, then  $P(X'; z) > P(X; z)$ .

The weak versions of these properties have been previously defined; the strict versions have not.<sup>60</sup> Note that the properties above are applicable when the identification function uses the deprived as well as the non-deprived dimensions to identify poor people. In other words, a poor person's identification status is allowed to change even when their achievements in non-deprived dimensions change while their achievements in the deprived dimensions remain unchanged.

The rearrangement set of properties could be made more precise when the identification of the poor respects the deprivation-focus property as well as the poverty-focus property. Identification that respects deprivation focus occurs when identification is solely based on dimensions in which poor persons are deprived, not on dimensions in which poor persons are not deprived. For example, these properties cannot distinguish situations when a poverty measure satisfying the deprivation-

with an increase in achievement in dimension  $j$  for people with higher achievements in dimension  $j'$  (Bourguignon and Chakravarty 2003: 35). Conversely, complementarity implies that poverty decreases more with an increase in achievement in dimension  $j$  for people with higher achievements in dimension  $j'$ . If the dimensions are independent, the second cross-partial derivative is zero and poverty should not change under the described transformation. This corresponds to the Auspitz–Lieben–Edgeworth–Pareto (ALEP) definition and differs from Hick's definition, traditionally used in the demand theory (which relates to the properties of the indifference contours) (Atkinson 2003: 55). See Kannai (1980) for critiques of the ALEP definition. For a critique of Bourguignon and Chakravarty's (2003) association axiom, see Decancq (2012).

<sup>60</sup> For various weak versions of the sensitivity to rearrangement properties in poverty measurement literature, see Tsui (2002), Chakravarty (2009) (which contains a modified version of the properties in Bourguignon and Chakravarty (2003), and Alkire and Foster (2011a). For different statements of the stronger versions of the property in the measurement of welfare and inequality, see Tsui (1995), Gajdos and Weymark (2005), Decancq and Lugo (2012), and Seth (2013).

focus property should be strictly or weakly sensitive to the joint distribution of achievements among the poor. Let us consider the following two examples where the deprivation cutoff vector is

$z = [5, 6, 4]$ . In the first example, suppose the achievement matrix  $X = \begin{bmatrix} 3 & 4 & 5 \\ 4 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  is obtained from

$Y = \begin{bmatrix} 3 & 4 & 4 \\ 4 & 5 & 5 \\ 8 & 6 & 4 \end{bmatrix}$  by switching the achievements between the two poor persons in the third

dimension. Clearly, an association-decreasing rearrangement has taken place between the two poor persons in  $X$ , but this switch should not affect overall poverty as none of these two persons is deprived in the third dimension.

In the second example, suppose  $X = \begin{bmatrix} 3 & 4 & 5 \\ 4 & 4 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  is obtained from  $Y = \begin{bmatrix} 4 & 4 & 5 \\ 3 & 4 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  by switching the

achievements between the two poor persons in the first dimension. Again, certainly an association-increasing rearrangement has taken place, but if we look at the achievements in the deprived dimensions of the two poor persons, they appear to be permutations of each other and thus overall poverty should not change. In order to make the transformations relevant in this situation, we need to ensure that the association-decreasing rearrangements occur only among the deprived dimensions of the poor. Thus, there is a need to define a new set of properties that is compatible with the deprivation-focus property, which can be done by defining the properties in terms of the censored achievement matrices.

In this book, we define an additional set of new rearrangement properties by defining a transformation called **association-decreasing deprivation rearrangement among the poor**. Let  $\tilde{Y}$  and  $\tilde{X}$  denote the censored achievement matrices for  $Y$  and  $X$ , respectively (defined in section 2.2.5). Consider two poor persons  $i$  and  $i'$  in  $Y \in \mathcal{X}_n$  such that  $y_{ij} \leq y_{i'j}$  for all  $j$ . If matrix  $X$  is obtained from  $Y$  such that  $x_{ij} = y_{ij}$  and  $x_{i'j} = y_{ij}$  for some dimension  $j$ , and  $x_{i'j} = y_{i'j}$  for all  $i'' \neq i, i'$  and all  $j' \neq j$ , and  $\tilde{X}$  is not a permutation of  $\tilde{Y}$ , then  $X$  is stated to be obtained from  $Y$  by an association-decreasing deprivation rearrangement among the poor. The requirement of  $\tilde{X}$  not being a permutation of  $\tilde{Y}$  has two analogous implications as in case of the association decreasing rearrangement. It prevents the two cases presented in the previous paragraph. Thus, it does not

consider the cases where the switch of achievements between the two (poor) persons takes place in their non-deprived dimensions instead of the deprived dimension. Also, it prevents the censored deprivation vectors from being permutations of each other due to an association-decreasing rearrangement. The following example illustrates the transformation.

**Example:** Suppose the initial achievement matrix is  $Y = \begin{bmatrix} 3 & 4 & 2 \\ 4 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and the deprivation cutoff

vector is  $z = [5, 6, 4]$ . Thus, the first person is deprived in three dimensions, the second person is deprived only in the two first dimensions, and the third person is non-deprived in all three dimensions. Moreover, the first person is more deprived than the second one in every dimension.

The corresponding censored achievement matrix is given by  $\tilde{Y} = \begin{bmatrix} 3 & 4 & 2 \\ 4 & 5 & 4 \\ 5 & 6 & 4 \end{bmatrix}$ . Now, suppose  $X$  is

obtained from  $Y$  such that  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$ . Note that the first and the second person have switched

their achievements in the first dimension so that now the first person is no longer more deprived than the second person in all three dimensions. The corresponding censored achievement matrix is

given by  $\tilde{X} = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 5 & 6 & 4 \end{bmatrix}$  which is clearly not a permutation of  $\tilde{Y}$ . In this case,  $X$  is said to be

obtained from  $Y$  by association-decreasing deprivation rearrangement among the poor.

We define the following four additional properties using the same concept of substitutability and complementarity between dimensions discussed previously, but require the association-decreasing rearrangement to take place between the deprived dimensions of the poor. Note that, due to the transformation, the set of poor remains unchanged.

**Weak Deprivation Rearrangement (Substitutes):** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by an association-decreasing deprivation rearrangement among the poor, then  $P(X'; z) \leq P(X; z)$ .

**Converse Weak Deprivation Rearrangement (Complements):** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by an association-decreasing

deprivation rearrangement among the poor, then  $P(X'; z) \geq P(X; z)$ .

**Strong Deprivation Rearrangement (Substitutes):** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by an association-decreasing deprivation rearrangement among the poor, then  $P(X'; z) < P(X; z)$ .

**Converse Strong Deprivation Rearrangement (Complements):** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by an association-decreasing deprivation rearrangement among the poor, then  $P(X'; z) > P(X; z)$ .

How are deprivation rearrangement properties related to or different from the rearrangement properties? First, if a poverty measure satisfies the (converse) weak deprivation rearrangement property, then the poverty measure will satisfy the (converse) weak rearrangement property, and the converse is true as well. Also, a poverty measure that satisfies the (converse) strong deprivation rearrangement property automatically satisfies the (converse) strong rearrangement property. But a poverty measure that satisfies the (converse) strong rearrangement property does not necessarily satisfy the (converse) strong deprivation rearrangement property. Therefore, the main difference between these two set of properties lies in their strong versions.

Although the rearrangement properties show technically how the change in poverty is related to association between dimensions, further research is required to understand the practicalities of rearrangement properties. Importantly, note that these properties require a uniform assumption across dimensions: either they are *all* substitutes, or they are *all* complements, which may be highly constraining. On the empirical side, there does not seem to be a standard procedure for determining the extent of substitutability and complementarity across dimensions of poverty. Moreover, it is not entirely clear that any interrelationships across variables must be incorporated into the overarching methodology for evaluating multidimensional poverty. Instead, the interconnections might plausibly be the subject of separate empirical investigations that supplement, but do not constitute, the underlying poverty measure.

A related property, which is consistent with the ordinality property discussed in section 2.5.1, is dimensional transfer. The association-decreasing rearrangement, as well as the association-decreasing deprivation rearrangement among poor people, requires the achievements of poor people to be rearranged. However, some rearrangements, even when achievement matrices are not permutations of each other, may not alter the deprivation status of the poor and thus the corresponding

deprivation matrices may either be identical or a permutation of each other. Therefore, the rearrangement properties discussed above are not useful for judging whether an ordinal poverty measure (as we discuss in section 3.6.1) is strictly or weakly sensitive to data transformations when deprivations are transferred between poor persons. Let us show with an example how an association-decreasing rearrangement among the poor may cause no change in the deprivation

matrices. Suppose two achievement matrices  $Y = \begin{bmatrix} 3 & 4 & 2 \\ 4 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  and  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$  with deprivation

cutoff vector  $z = [5, 6, 4]$ , where  $X$  is obtained from  $Y$  by an association-decreasing rearrangement among the poor. These two achievement matrices have identical corresponding deprivation

matrices, such that  $g^0(Y) = g^0(X) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ .

A dimensional rearrangement among the poor is an association-decreasing rearrangement among the poor (in achievements) that is simultaneously an association-decreasing rearrangement in deprivations. In other words, the initial deprivation vectors (and achievement vectors) are ranked by vector dominance, while the final deprivation vectors (and achievement vectors) are not. The extra condition ensures that the person with a lower level of achievements is actually deprived in some dimensions in which the other person is not and that, through the rearrangement, one or more of these deprivations (but not all) are traded for non-deprived levels. More formally, let  $g^0(Y)$  and  $g^0(X)$  denote the deprivation matrices for  $Y$  and  $X$ , respectively (defined in section 2.2.1). Consider two poor persons  $i \in Z$  and  $i' \in Z$  (according to some identification method  $\rho$ ) in  $Y \in \mathcal{X}_n$  such that  $g_{i'j}^0(Y) \leq g_{ij}^0(Y)$  for all  $j$ . If matrix  $X$  is obtained from  $Y$  such that  $g_{ij}^0(X) = g_{ij}^0(Y)$  and  $g_{i'j}^0(X) = g_{ij}^0(Y)$  for some dimension  $j$ , and  $g_{i'j}^0(X) = g_{i'j}^0(Y)$  for all  $i'' \neq i, i'$  and all  $j' \neq j$ , and  $g^0(X)$  is not a permutation of  $g^0(Y)$ , then we define  $X$  to be obtained from  $Y$  by a **dimensional rearrangement among the poor**. A dimensional rearrangement among the poor does not affect the number of poor persons, and neither does a dimensional increment among the poor. This transformation can be interpreted as a progressive transfer in that it transforms an initial ‘spread’ in joint deprivations between two poor persons into a moderated situation where neither person has unambiguously more than the other. The overall achievement levels in society are unchanged, but

the correlation between dimensions (and hence inequality) has been reduced. The following property requires poverty to decrease when there is a dimensional rearrangement among the poor.

Dimensional rearrangement among the poor, as defined above, covers only switches of achievements and deprivations in deprived dimensions among people who are and remain poor, and excludes permutations of corresponding deprivation matrices.

**Dimensional Transfer:** If an achievement matrix  $X'$  is obtained from another achievement matrix  $X$  by a dimensional rearrangement among the poor, then  $P(X'; z) < P(X; z)$ .<sup>61</sup>

### 2.5.3 Subgroup Properties

The next set of principles is concerned with the link between overall poverty and poverty in different subgroups of the population, and the link between overall poverty and dimensional deprivations.

The first principle—**subgroup consistency**—ensures that the change in overall poverty is consistent with the change in subgroup poverty.<sup>62</sup> For example, suppose the entire society is divided into two population subgroups: Group 1 and Group 2. Poverty in Group 1 remains unchanged while poverty in Group 2 decreases. One would expect overall poverty to decrease. If overall poverty did not reflect subgroup poverty, there would be an inconsistency, which would be conceptually and politically problematic. As a result, national poverty estimates would not reflect regional successes in poverty reduction. A related principle with a stronger requirement is **population subgroup decomposability**. This principle requires overall poverty to be equal to a weighted sum of subgroups' poverty, noted as  $P(X^\ell; z)$  in section 2.2.2, where the weight attached to each subgroup's poverty is the population share of that subgroup.

Suppose an achievement matrix  $X$  is divided into  $m \geq 2$  subgroups, such that  $X^\ell$  and  $n^\ell$  denote, correspondingly, the achievement matrix and the population size of subgroup  $\ell$ , for all  $\ell = 1, \dots, m$ , and the subgroups are mutually exclusive and collectively exhaustive:

$$\sum_{\ell=1}^m n^\ell = n.$$

---

<sup>61</sup> For a different statement of the strong dimensional transfer property using an association-increasing rearrangement, see Seth and Alkire (2013).

<sup>62</sup> The concept of subgroup consistency in poverty measurement has been motivated by Foster and Shorrocks (1991).

**Subgroup Consistency:** If an achievement matrix  $X'$  is obtained from the achievement matrix  $X$  such that  $P(X^{\ell'}; z) < P(X^{\ell}; z)$  but  $P(X^{\ell'}; z) = P(X^{\ell}; z)$  for all  $\ell \neq \ell'$ , and total population, as well as subgroup population, remains unchanged, then  $P(X'; z) < P(X; z)$ .

**Population Subgroup Decomposability:** 
$$P(X; z) = \sum_{\ell=1}^m (n^{\ell} / n) P(X^{\ell}; z)$$

**Example:** Suppose the initial achievement matrix is  $X = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$ . Let us divide  $X$

into two subgroups  $X^1 = \begin{bmatrix} 4 & 4 & 2 \end{bmatrix}$  with population size  $n^1 = 1$  and  $X^2 = \begin{bmatrix} 3 & 5 & 4 \\ 8 & 6 & 4 \end{bmatrix}$

with population size  $n^2 = 2$ . Now, suppose  $X^1$  changes to  $X^1 = \begin{bmatrix} 4 & 4 & 3 \end{bmatrix}$  and  $X^2 = X^2$ .

We know that by the monotonicity principle,  $P(X^1; z) < P(X^1; z)$ . Then the subgroup consistency principle requires that  $P(X'; z) < P(X; z)$ .

The additive decomposability principle requires that overall poverty can be expressed as

$$P(X; z) = \frac{1}{3} P(X^1; z) + \frac{2}{3} P(X^2; z).$$

The population subgroup decomposability property has been one of the most attractive properties for policy analysis as it can be particularly useful for targeting and monitoring progress in different subgroups. It is worth noting that a poverty measure that satisfies population subgroup decomposability necessarily satisfies subgroup consistency. However, the converse is not true, which means subgroup consistency does not necessarily imply population subgroup decomposability.

The other form of decomposition that is of tremendous relevance in the policy analysis of multidimensional poverty refers to the possibility of breaking down poverty by deprivations across dimensions among the poor. This property, called **dimensional breakdown**, requires overall poverty to be equal to a weighted sum of the dimensional deprivations *after identification*  $P_j(x_j; z)$  introduced in section 2.2.2. It creates a consistency between the post-identification dimensional deprivations and overall poverty.

In the particular case in which a counting approach using a union criterion is followed for identification, then  $P_j(x_j; z) = P_j(x_j; z_j)$ , provided the base population  $n_j = n$  for all  $j$ . In other

words, the dimensional deprivation of any dimension  $j$  after identification coincides with the dimensional deprivation level before identification. In this special case, the property of dimensional breakdown coincides with the property of ‘factor decomposability’ introduced by Chakravarty, Mukherjee, and Ranade (1998) and referred to by Bossert, Chakravarty, and D’Ambrosio (2013) as ‘additive decomposability in attributes’. In other cases, dimensional breakdown is done after the identification of the poor and shows only the deprivations faced by the poor.

**Dimensional Breakdown:** For an  $n \times d$  dimensional achievement matrix  $X$ ,  $P(X; z) = \sum_{j=1}^d w_j P_j(x_j; z)$ , where  $w_j$  is the weight attached to dimension  $j$  and  $P_j(x_j; z)$  is the dimensional deprivation index after identification in dimension  $j$ .

Given that the dimensional breakdown property requires additivity in the deprivations, it is not consistent with the properties of association sensitivity in their strict form—that is, with requiring decreasing or increasing poverty under an association-decreasing rearrangement.<sup>63</sup>

#### 2.5.4 Technical Properties

Finally, we introduce certain technical principles, which ensure that the poverty measure is meaningful. These principles are **non-triviality**, **normalization**, and **continuity**.

The **non-triviality** principle requires that a poverty measure takes at least two different values. This property may appear to be trivial by its name, but it is important: unless a measure takes two different values, it is not possible to distinguish a society with poverty from a society with no poverty. Note that when a measure satisfies the strong version of at least one of the dominance principles, this property is automatically satisfied (by definition, poverty will take at least two different values). However, when a measure only satisfies the weak version of all dominance principles, this property becomes necessary.

The **normalization** principle requires that the values of a poverty measure lie within the 0–1 range. It takes a minimum value of 0 when there is no poverty in a society, and it takes a maximum value of 1 when poverty is at its maximum. The **continuity** property prevents a poverty measure from changing abruptly, given marginal changes in achievements.

---

<sup>63</sup> For a formal discussion of this inconsistency, see Alkire and Foster (2013).

**Non-triviality:** A poverty measure should take at least two distinct values.

**Normalization:** A poverty measure should take a minimum value of 0 and a maximum value of 1.

**Continuity:** A poverty measure should be continuous on the achievements.

It is worth noting that not all properties defined above are applicable across all scales of measurement, just as not all mathematical operations are admissible for all scales of measurement. Thus, some of these properties may need to be adapted according to the requirements of different scales. The next chapter outlines various poverty measurement methodologies based on the framework introduced in this chapter and discusses which scales of measurement they use and which properties they satisfy.

## References

- Alkire, S. and Santos, M. E. (2009). “Poverty and Inequality Measurement”, in S. Deneulin and L. Shanai (eds.) *An Introduction to the Human Development and Capability Approach*, Abingdon, Earthscan Ltd., Ch. 6.
- Alkire, S. (2011). ‘Multidimensional Poverty and its Discontents’, in *Measure for Measure: How Well Do We Measure Development?* Proceedings of the 8th AFD-EUDN Conference, 55-90.
- Alkire, S. and Foster, J. (2007). ‘Counting and Multidimensional Poverty Measurement’. *OPHI Working Paper 7*, Oxford University.
- Alkire, S. and Foster, J. (2011a). ‘Counting and Multidimensional Poverty Measurement’. *Journal of Public Economics*, 95(7–8): 476–487.
- Alkire, S. and Foster, J. (2011b). ‘Understandings and Misunderstandings of Multidimensional Poverty Measurement’. *Journal of Economic Inequality*, 9: 289–314.
- Alkire, S. and Foster, J. (2011b). ‘Understandings and Misunderstandings of Multidimensional Poverty Measurement’. *Journal of Economic Inequality*, 9: 289–314.
- Andrews et al.: Andrews, F. M., Klem, L., Davidson, T. N., O'Malley, P. M., and Rodgers, W. L. (1981). *A Guide for Selecting Statistical Techniques for Analyzing Social Science Data*. University of Michigan, Institute for Social Research.
- Atkinson, A. B. (1970). ‘On the Measurement of Inequality’. *Journal of Economic Theory*, 2(3): 244–263.
- Atkinson, A. B. (1989). *Poverty and Social Security*. Harvester Wheatsheaf.
- Atkinson, A. B. (2003). ‘Multidimensional Deprivation: Contrasting Social Welfare and Counting Approaches’. *Journal of Economic Inequality*, 1(1): 51–65.
- Atkinson, A. B. and Bourguignon, F. (1982). ‘The Comparison of Multi-dimensioned Distributions of Economic Status’. *Review of Economic Studies*, 49(2): 183–201.
- Basu, K. (1980). *Revealed Preference of Government*. Cambridge UP.
- Boland, P. J. and Proschan, F. (1988). ‘Multivariate Arrangement Increasing Functions with Applications in Probability and Statistics’. *Journal of Multivariate Analysis*, 25(2): 286–298.
- Booth, C. (1894). *The Aged Poor: Condition*. Macmillan.
- Booth, C. (1903). *Life and Labour of the People in London*. Macmillan.
- Bossert, W., Chakravarty, S., and D'Ambrosio, C. (2013) ‘Multidimensional Poverty and Material Deprivation With Discrete Data’. *Review of Income and Wealth*, 59(1): 29–43.
- Bossert, W., Chakravarty, S., and D'Ambrosio, C. (2013) ‘Multidimensional Poverty and Material Deprivation With Discrete Data’. *Review of Income and Wealth*, 59(1): 29–43.
- Bourguignon, F., and Chakravarty, S. R. (2003). ‘The Measurement of Multidimensional Poverty’. *Journal of Economic Inequality*, 1(1): 25–49.
- Bourguignon, F., and Chakravarty, S. R. (2003). ‘The Measurement of Multidimensional Poverty’. *Journal of Economic Inequality*, 1(1): 25–49.

- Bowley, A. L. and Burnett-Hurst, A. R. (1915). *Livelihood and Poverty, A Study in the Economic and Social Conditions of Working Class Households in Northampton, Warrington, Stanley, Reading (and Bolton)*. King.
- Brown, J. D. (2011). 'Questions and Answers about Language Testing Statistics: Likert Items and Scales of Measurement?'. *SHIKEN: JALT Testing & Evaluation SIG Newsletter* 15(1): 10–14.
- Calvo C. and Dercon, S. (2013). 'Vulnerability to Individual and Aggregate Poverty'. *Social Choice and Welfare* 41(4): 721–740.
- Chakravarty, S. R. (1983a). 'Ethically Flexible Measures of Poverty'. *Canadian Journal of Economics* 16(1): 74-85.
- Chakravarty, S. R. (2009). *Inequality, Polarization and Poverty: Advances in Distributional Analysis*, vol. VI. Springer.
- Chakravarty, S. R. and D'Ambrosio, C. (2013). 'A Family of Unit Consistent Multidimensional Poverty Indices', in V. Bérenger and F. Bresson (eds.), *Poverty and Social Exclusion around the Mediterranean Sea*. Springer, 75–88.
- Chakravarty, S. R., Deutsch, J., and Silber, J. (2008). 'On the Watts Multidimensional Poverty Index and Its Decomposition'. *World Development* 36 (6): 1067-1077.
- Chakravarty, S. R., Mukherjee, D., and Ranade, R. R. (1998). 'On the Family of Subgroup and Factor Decomposability Measures of Multidimensional Poverty', in D. J. Slottje (ed.), *Research on Economic Inequality* 8. JAI Press, 175–194.
- Cohen et al.: Cohen, L., Manion, L., Morrison, K. (2000). *Research Methods in Education*. (5th ed.). Routledge Falmer.
- Dalton, H. (1920). 'The Measurement of the Inequality of Incomes'. *The Economic Journal*, 31(121): 348-361.
- Decancq, K. (2012). 'Elementary Multivariate Rearrangements and Stochastic Dominance on a Fréchet Class'. *Journal of Economic Theory*, 147(4): 1450–1459.
- Decancq, K. and Lugo, M. A. (2012). 'Weights in Multidimensional Indices of Well-Being: An Overview'. *Econometric Reviews*, 32(1): 7–34.
- Deutsch, J. and Silber, J. (2005). 'Measuring Multidimensional Poverty: An Empirical Comparison of Various Approaches'. *Review of Income and Wealth*, 51(1): 145–174.
- Deutsch, J. and Silber, J. (2008). 'The Order of Acquisition of Durable Goods and the Measurement of Multidimensional Poverty', in N. Kakwani and J. Silber (eds.), *Quantitative Approaches to Multidimensional Poverty Measurement*, London, Palgrave-Macmillan, pp. 226-243.
- Duclos et al.: Duclos, J. Y., Sahn, D. E., and Younger, S. D. (2011). 'Partial Multidimensional Inequality Orderings'. *Journal of Public Economics*, 95(3–4): 225–238.
- Duclos, J. Y., Sahn, D. E., and Younger, S. D. (2006a). 'Robust Multidimensional Poverty Comparisons'. *The Economic Journal*, 116(514): 943–968.
- Fleurbaey, M. (2006a). 'Social Welfare, Priority to the Worst-Off and the Dimensions of Individual Well-Being', in F. Farina and E. Savaglio (eds.), *Inequality and Economic Integration*. Routledge, 225–268.

- Foster *et al.* (2010): Foster, J., Greer, J., and Thorbecke, E. (2010). 'The Foster–Greer–Thorbecke (FGT) Poverty Measures: 25 Years Later'. *Journal of Economic Inequality*, 8(4): 491–524.
- Foster J. and Santos M. E. (2013). 'Measuring Chronic Poverty', in G. Betti and A. Lemmi (eds.), *Poverty and Social Exclusion: New Methods of Analysis*. Routledge, 143–165.
- Foster J., *et al.* (2013): Foster, J. E., Seth S., Lokshin, M., and Sajaia Z. (2013). *A Unified Approach to Measuring Poverty and Inequality: Theory and Practice*. The World Bank.
- Foster J., Greer, J., and Thorbecke, E. (1984). 'A Class of Decomposable Poverty Measures'. *Econometrica*, 52(3): 761–766.
- Foster, J. (2006). 'Poverty Indices', in A. de Janvry and R. Kanbur (eds.), *Poverty, Inequality and Development: Essays in Honor to Erik Thorbecke*. Springer, 41–66.
- Foster, J. (2009). 'A Class of Chronic Poverty Measures', in T. Addison, Hulme, D., and R. Kanbur (eds.), *Poverty Dynamics: Interdisciplinary Perspectives*. OUP, 59–76.
- Foster, J. and Sen, A. K. (1997). 'On Economic Inequality after a Quarter Century', an annexe to A. Sen, *On Economic Inequality*. OUP, 107–220.
- Foster, J. and Shorrocks, A. F. (1991). 'Subgroup Consistent Poverty Indices'. *Econometrica*, 59(3): 687–709.
- Gajdos, T. and Weymark, J. A. (2005). 'Multidimensional Generalized Gini Indices'. *Economic Theory*, 26(3): 471–496.
- Guttman, L. (1977). 'What Is Not What in Statistics'. *The Statistician*, 26(2): 81–107.
- Hansen, J. P. (2003). 'CAN'T MISS—Conquer any Number Task by Making Important Statistics Simple. Part 1: Types of Variables, Mean, Median, Variance and Standard Deviation.' *Journal for Healthcare Quality*, 25(4): 19–24.
- Hoy, M. and Zheng, B. (2011). 'Measuring Lifetime Poverty'. *Journal of Economic Theory*, 146(6): 2544–2562.
- Jamieson, S. (2004). 'Likert Scales: How to (Ab)use Them'. *Medical Education* 38(12): 1217–1218.
- Kannai, Y. (1980). 'The ALEP Definition of Complementarity and Least Concave Utility Functions'. *Journal of Economic Theory*, 22(1): 115–117.
- Kolm, S. C. (1976a). 'Unequal Inequalities. I'. *Journal of Economic Theory*, 12(3): 416–442.
- Kolm, S. C. (1976b). 'Unequal Inequalities. II'. *Journal of Economic Theory*, 13(1): 82–111.
- Kolm, S. C. (1977). 'Multidimensional Egalitarianisms'. *The Quarterly Journal of Economics*, 91(1): 1–13.
- Likert, R. (1932). 'A Technique for the Measurement of Attitudes'. *Archives of Psychology*, 140: 1–55.
- Lord, F. M. (1953). 'On the Statistical Treatment of Football Numbers'. *American Psychologist* 8: 750–751.
- Luce, R. D. (1956). 'Semioorders and a Theory of Utility Discrimination'. *Econometrica*, 24: 178–191.
- Luce, R. D. (1959). 'On the Possible Psychophysical Laws'. *Psychological Review*, 66(2): 81–95.
- Marcus-Roberts, H. and Roberts, F. S. (1987). 'Meaningless Statistics'. *Journal of Educational Statistics* 12(4): 383–394.

- Marshall, A. W. and Olkin, I. (1979). *Theory of Majorization and its Applications*. Academic.
- Norman, G. (2010). 'Likert Scales, Levels of Measurement and the "Laws" of Statistics'. *Advances in Health Science Education* 15(5): 625–632.
- O'Donnell, O., van Doorslaer, E., Wagstaff, A., and Lindelow, M. (2008). *Analyzing Health Equity Using Household Survey Data. A Guide to Techniques and Their Implementation*. World Bank.
- Pattanaik, P., Reddy, S., and Xu, Y. (2012). 'On Measuring Deprivation and Living Standards of Societies in a Multi-Attribute Framework'. *Oxford Economic Papers* 64: 43–56.
- Pett, M. A. (1997). *Non-parametric Statistics for Health Care Research*. SAGE Publications.
- Porter, C. and Quinn, N. (2013). 'Measuring Intertemporal Poverty: Policy Options for the Poverty Analyst', in G. Betti and A. Lemmi (eds.), *Poverty and Social Exclusion: New Methods of Analysis*. Routledge, 166–193.
- Ravallion, M. (1992). 'Poverty Comparisons: A Guide to Concepts and Methods'. *Living Standards Measurement Study, Working Paper* 88. World Bank.
- Roberts, F. S. (1979). *Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences*, Vol. 7. CUP.
- Roberts, F. S. and Franke, C. H. (1976). 'On the Theory of Uniqueness in Measurement'. *Journal of Mathematical Psychology* 14(3): 211–218.
- Rowntree, B. S. (1901). *Poverty: A Study of Town Life*. Macmillan.
- Sarle, W. S. (1995). 'Measurement Theory: Frequently Asked Questions'. *Disseminations of the International Statistical Applications Institute*. <ftp://ftp.sas.com/pub/neural/measurement.html>. Accessed 2 August 2013.
- Sen, A. (1970). *Collective Choice and Social Welfare*. Oliver and Boyd.
- Sen, A. (1973). *On Economic Inequality*, OUP.
- Sen, A. (1976). 'Poverty: An Ordinal Approach to Measurement'. *Econometrica*, 44(2): 219–231.
- Sen, A. K. (1979). 'Equality of What?', in S. McMurrin (ed.), *Tanner Lectures on Human Values*. CUP.
- Sen, A. K. (1985). 'Well-Being, Agency and Freedom: The Dewey Lectures 1984'. *The Journal of Philosophy*, 82(4): 169–221.
- Sen, A. K. (1992). *Inequality Re-Examined*. OUP.
- Sen, A. K. (1997). *On Economic Inequality. Expanded Edition with a Substantial Annex by James E Foster and Amartya Sen*. OUP.
- Seth, S. (2009). 'Inequality, Interactions, and Human Development'. *Journal of Human Development and Capabilities*, 10(3): 375–396.
- Seth, S. (2013). 'A Class of Distribution and Association Sensitive Multidimensional Welfare Indices'. *Journal of Economic Inequality*, 11(2): 133–162.
- Seth, S. and Alkire, S. (2013). 'Measuring and Decomposing Inequality among the Multidimensionally Poor using Ordinal Variables: A Counting Approach'. *OPHI Working paper* 68, Oxford University.
- Stevens, S. S. (1946). 'On the Theory of Scales of Measurement'. *Science*, NS 103 (2684): 677–680.

- Stevens, S. S. (1951). 'Mathematics, Measurement and Psychophysics', in S. S. Stevens (ed.), *Handbook of experimental psychology*. Wiley: 1–49.
- Stevens, S. S. (1959). 'Measurement, Psychophysics, and Utility', in C. W. Churchman and P. Ratoosh (eds.), *Measurement: Definitions and theories*. Wiley: 18–63.
- Stevens, S. S. (1959). 'Measurement, Psychophysics, and Utility', in C. W. Churchman and P. Ratoosh (eds.), *Measurement: Definitions and theories*. Wiley: 18–63.
- Thon, D. (1979). 'On Measuring Poverty'. *Review of Income and Wealth*, 25: 429–440.
- Townsend, P. (1954). 'Measuring Poverty'. *British Journal of Sociology*, 5: 130–137.
- Tsui K.-Y. (1995). 'Multidimensional Generalizations of the Relative and Absolute Inequality Indices: The Atkinson-Kolm-Sen Approach'. *Journal of Economic Theory*, 67(1): 251–265.
- Tsui K.-Y. (1999): 'Multidimensional Inequality and Multidimensional Entropy Measures: An Axiomatic Derivation?'. *Social Choice and Welfare*, 16(1): 145-157.
- Tsui, K. (2002). 'Multidimensional Poverty Indices'. *Social Choice and Welfare* 19(1): 69–93.
- Tukey, J. W. (1961). 'Data Analysis and Behavioral Science or Learning to Bear the Quantitative Man's Burden by Shunning Badmandments', in L. V. J. Belmont (ed.), *The Collected works of John W. Tukey*, Vol. III (1986). Wadsworth, 391–484.
- Velleman, P. F. and Wilkinson, L. (1993). 'Nominal, Ordinal, Interval, and Ratio Typologies are Misleading'. *The American Statistician*, 47(1): 65–72.
- Watts, H.W. (1968). 'An Economic Definition of Poverty', in D. P. Moynihan (ed.), *On Understanding Poverty*. Basic Books.
- WHO Programme of Nutrition. (1997). *Global Database on Child Growth and Malnutrition*. Compiled by M. de Onis and M. Blössner. World Health Organization.
- WHO Multicentre Growth Reference Study Group. (2006). *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for- Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development*. Geneva: World Health Organization.
- Zheng B. (1997). 'Aggregate Poverty Meas Measures'. *Journal of Economic Surveys*, 11(2): 123–162.