ESSAYS IN MULTIDIMENSIONAL MEASUREMENT: WELFARE, POVERTY, AND ROBUSTNESS

By

Suman Seth

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

August, 2010

Nashville, Tennessee

Approved:

Professor John A. Weymark

Professor Yanqin Fan

Professor Mark A. Cohen

Professor James E. Foster

To my parents, whose persistent inspiritation and honest support convinced me to successfully complete this once seemingly impossible journey.

Acknowledgements

This is a great opportunity for me to express my sincere gratitude to those without whose assistance this dissertation would have been difficult to finish. At the very beginning, I would like to convey my deepest appreciation to my advisor Professor James Foster for his competent guidance, thoughtful advice, and gracious support. He has always encouraged me to work independently, but has been available with his valuable suggestions whenever essential.

I now take the opportunity to thank my co-advisor Professor John Weymark and my dissertation committee members Professor Mark Cohen and Professor Yanqin Fan for their help, advice, and support. I thank to Professor Weymark for being exceedingly generous in spending his valuable time mentoring me to improve the quality of the first chapter of my dissertation. I am grateful to Professor Cohen for his precious support during the second and third years of my doctoral program. I am thankful to Professor Fan for her valuable advice on the part of the dissertation where I have used econometric techniques.

Sabina Alkire, the director of the Oxford Poverty and Human Development Initiative (OPHI), has played a crucial role in determining the composition of my dissertation. Motivated by OPHI's interest in policy analysis, two chapters of my dissertation apply multidimensional techniques to the Indian context. I thank Sabina for her support and amicable advice. I appreciate the financial support from Professor Jennifer Reinganum and the Regional Bureau for Latin America and the Caribbean UNDP that partly funded my research in this dissertation.

I am grateful to Professor Gregory Huffman, Professor Jeremy Atack, and Professor William Collins, who have been the Director of Graduate Studies at different points of time during my stay at Vanderbilt University, for their valuable advice. I am indebted to the faculty members of the economics department who have taught me various tools of economics. I thank Kathleen Finn, Cathy Zebron, and all others administrative staffs who have always been happy to assist me.

On the course of the past six years of my doctoral program, I have been fortunate to have some wonderful friends such as Tamara, Bingyu, Yuling, Valeska, Rosa, Scott, Sujan, Shabana, Alper, Caleb, Sang Soo, Bob, Tommaso, Ali Sina, Suchismita, Indranil, Soumita, Gaston, and Maria Emma, and I am thankful for their generous help and support. I thank David Frisvold, Aniruddha Bagchi, and Rajdeep Sengupta for their advice at different stages of this program. I appreciate the valuable comments that have come from Koen Decancq on the first chapter of my dissertation.

Finally, I express my gratitude to those who did not directly assist me on my dissertation writing, but have been a continual source of inspiration. As always, there has never been a lack of inspiration from my parents and my younger brother, who have continuously motivated me from thousands of miles away. I am also grateful to my aunt Rama Pal and my uncle Sankar Prasad Pal for their help and encouragement. During the last three years of dissertation writing, I have been through several pleasing and cheerful moments and also through a number of moments when I was upset and discontent. The person whom I have always found beside me through all these moments is my wife Barnali. I am grateful to her not only for being such a wonderful partner, but also for being an excellent friend.

Table of Contents

List of Tables

Table | Page

List of Figures

Chapter I

Introduction

Measuring social welfare and deprivation has always been a challenging task for both economic theorists and policy makers. The standard norm has been to use of income for these purposes. Even in the second half of the twentieth century, the welfare or human development or well-being as it is variously known, of a society was predominantly gauged in terms of average income or wealth. Similarly, individuals or households in many countries are still identified as destitute if they fail to acquire incomes above a subsistence threshold. This led to the measurement and analysis of both welfare and poverty being based only on a single component or attribute (also known as 'dimension') of well-being.

However, the proponents of the basic needs approach (Streeten et al., 1981) and later the capability approach (Sen, 1985) have shown that the perception of human welfare and deprivation go beyond income or wealth. The basic needs approach identifies an individual or a household as destitute if they fail to achieve the resources — such as food, shelter, health care, and education — needed to sustain long term physical well-being. The capability approach developed primarily by Amartya Sen, on the other hand, argues that well-being should be based on what individuals are capable of doing and being, and not merely on the commodity bundle that they own. These two approaches have their differences (Anand and Ravallion, 1993), but they are common in at least one aspect: both encourage the measurement of social welfare and poverty to be based on multiple components or attributes of well-being, such as education and health, instead of income alone.

These approaches have inspired several indices of welfare or poverty over the last few decades and motivated many international orgnizations and policy makers to embrace a multidimensional framework for assessing the level of both well-being and deprivation. Examples include, but are not limited to, the well known Human Development Index (HDI) and the Human Poverty Index (HPI) published annually by the United Nations Development Programme (UNDP), various physical quality of life indices, e.g. Morris (1979), and the Human Opportunity Index (de Barros et al., 2009) developed recently by the World Bank researchers. In 2002, the government of India has proposed identifying families below the poverty line using a multidimensional survey (Government of India, 2002). A recent commission appointed by French President Nicolas Sarkozy also recommends using a multidimensional definition of well-being (Stiglitz, Sen, and Fitoussi, 2009). For further instances where the governments of different countries have moved towards proposing a multidimensional definition, see Alkire and Sarwar (2009).

In this dissertation, the indices are classified into the following two categories, when the measurement is based on more than one attribute of well-being: *multidimensional index* and *composite index*. An index that summarizes the state of a society by aggregating achievements of individuals or households based on multiple attributes is called a *multidimensional index*, where an *achievement* refers to the quantity of an attribute obtained by an individual. For a multidimensional index, it is mandatory that the information on all attributes for each individual or household is available from the same data set.

However, it often happens that either (i) the information is collected from different sources or (ii) the achievements in different attributes belong to different sets of individuals within a society or (iii) the information is available only in aggregated form. The first situation occurs when one uses different sample surveys to collect information on different attributes; for example, information on income or consumption expenditure may be collected from an expenditure survey, but information on health may be available from a health survey. Clearly, the same set of individuals is not interviewed in both surveys. The second situation arises when the information on selected attributes is based on different sets of individuals. For example, the life expectancy rate is based on the individuals who passed away, the mortality rate is based on the children under five years of age, the literacy rate is based on the individuals older than fourteen years, the school enrolment rate is based on the individuals who are in the age group of five to fourteen years, etc. Under these circumstances, the achievements across individuals are aggregated to construct an *indicator* for each attribute, and then all these indicators are combined by taking a weighted average to obtain a composite index. Thus, an index that summarizes the state of a society by aggregating the indicators of multiple attributes is called a *composite index*. The composite indices are not only restricted to welfare economics, but are also widely applied to other branches of social science. The well-known examples include, but are not limited to, the Human Development Index, the Human Poverty Index, the Environmental Performance Index, the Global Peace Index, the Index of Economic Freedom, the Child Well-Being Index, and the index for colleges in the U.S. News college rankings.

Note that the definitions of a multidimensional index and a composite index are different, but it may sometime be hard to draw a concrete line of distinction between them. For instance, information on individual achievements may be collected using the same survey and also may be available for the same set of individuals in a society. If the achievements are first aggregated across individuals to obtain an indicator for each attribute, and these indicators are then aggregated to obtain an index, then clearly, this is a multidimensional index by definition. However, if one assumes that the indicators in the first stage are given and aggregates these indicators to obtain an index, then based on the second stage aggregation only, the same index can be interpreted as being a composite index.

One may confront several challenges while constructing an index based on multiple attributes. The major challenges are: to choose an appropriate set of attributes, to collect reasonable data, to determine a suitable set of weights reflecting the importance of each attribute, to select an appropriate aggregation method so that the data can be meaningfully summarized, and to verify the statistical significance of the evaluation generated by the index. This dissertation primarily focuses on various issues related to the method of aggregation so that the interpretation of the indices and the comparisons based on them have meaningful policy implication. In particular, it focuses on three different aspects concerning the aggregation methods of three different sets of indices: multidimensional welfare indices, composite indices, and multidimensional poverty indices. Furthermore, the tools developed in this dissertation are applied to real world data showing how they may influence existing policy decisions.

Chapter II is devoted to developing a class of multidimensional social welfare indices that is sensitive to inequality across individuals because a satisfactory index of social welfare or poverty should be sensitive to the *inequality* in the distributions of the attributes (Atkinson, 1970; Foster and Sen, 1997). Aside from its direct concern, inequality may well have negative indirect effects on social welfare. For example, a high level of inequality may lead to political instability (Alesina and Perotti, 1996; Justino, 2004), tensions among different ethnic groups (Stewart, 2008), an increase in crime rates (Fajnzylber et al., 2002), and feelings of deprivation among the members of society. These consequences, in turn, have adverse effects on the level of social welfare. Thus, the first aspect of aggregation is concerned with developing a class of social welfare indices that is sensitive to inequality across individuals. If two societies have the same level of average achievement in each attribute, then the one with less inequality across individuals should reflect a higher level of social welfare.

The second aspect of aggregation is associated with the composite indices that are obtained by taking a weighted average of various indicators of attributes. These indices are frequently used to rank societies or countries. These rankings are often of high national priority. It is a matter of pride for countries to be on the top of the lists. Moreover, various donor countries judge the performance of the debt seeking developing countries based on these composite indices. However, the choice of weights when constructing these indices is crucial because a choice of different weights, other than the one used, may alter the existing rankings and thus leading to ambiguous comparisons. A comparison is ambiguous or not robust if it is reversed when different weights are chosen. On the other hand, a comparison between a pair of countries or societies is completely robust if the comparison is never reversed when weights are changed. Therefore, in addition to making comparison across regions, it is important that one verify how robust these comparisons are. The next two chapters of the dissertation deal with this second aspect of aggregation. In Chapter III, a natural measure for evaluating the level of robustness is proposed and characterized. In Chapter IV, this new measure of robustness is applied to certain real world data sets to examine how the prevalence of robust comparisons varies when the targeted level of robustness is altered. It is shown that this relationship is influenced by the association among the indicators across societies. The research in the Chapter III and IV is jointly conducted with my advisor James Foster and Mark McGillivray.

The Chapters V and VI are devoted to the application of the class of multidimensional welfare indices developed in the first chapter and an application of a class of multidimensional poverty indices developed by Alkire and Foster (2008), respectively, to the Indian context. I find the applications to the Indian context interesting for the following reasons. There has been almost a three-fold increase in the national per-capita gross domestic product of India between 1990-91 and 2007-08. Furthermore, analyses using various poverty measures have suggested a significant fall in poverty when it is defined purely in terms of income. At the same time, the national family health survey and the human development report reveal that more than fifty percent of the rural women are illiterate, fifty seven infants do not survive out of every thousand newborns, nearly ninety percent of the rural households use solid biomass fuel for cooking purposes, and sixty seven percent of the population live without improved sanitation facilities as mandated in the millennium development goals (MDG) by the UNDP. Clearly, an improved performance in terms of income alone fails to reflect improved performance in other attributes of well-being. Moreover, inequality in achievements also remains high across the population and also across different population subgroups, such as across various geographical regions, across religions, and across castes/tribes. Although the dimension-specific averages partly explain these observations, they ignore the existing inter-person inequality. Thus, India happens to be an appropriate context for the application of a multidimensional social welfare index that is also sensitive to inter-personal inequality and the Chapter V is devoted to this objective.

The third aspect of aggregation concerns multidimensional poverty indices. In 2002, the Government of India has proposed identifying families below the poverty line using a multidimensional survey. The survey consists of thirteen questions with five responses each. A household is identified as poor if the household fails to secure a certain score out of these thirteen questions. A poverty index is constructed by just counting the number of poor persons. Consequently, the poverty index is neither sensitive to the depth, nor the breadth of poverty. By saying that the index is not sensitive to the *depth of poverty*, it is meant that the index does not change if a deprived person becomes more deprived in one attribute. Similarly, by saying that the index is not sensitive to the *breadth of poverty*, it is meant that the index does not change if a poor individual becomes deprived in an additional attribute in which (s)he was not deprived before. Moreover, the poverty index does not allow for poverty decomposition across attributes. In other words, it is not possible to calculate the contribution of each attribute to total poverty. A recently proposed class of multidimensional poverty indices by Alkire and Foster (2008) is sensitive to both the depth and the breadth of poverty, and allows for the decomposition of poverty across attributes. In Chapter VI, this new poverty index is applied to the Indian context to analyse the state of multidimensional poverty. The research in this chapter is conducted jointly with Sabina Alkire.

Finally, in Chapter VII, possible extensions are discussed and concluding remarks are provided. Since the research in some of the chapters has been conducted jointly, for the sake of uniformity, third person pronouns are used instead of first person pronouns throughout the rest of this dissertation.

Chapter II

A Class of Distribution and Association Sensitive Multidimensional Welfare Indices

Introduction

In this chapter, we are concerned with the evaluation of social welfare when there are two or more attributes of well-being. To have a common basis for comparison across different societies, we suppose that the set of attributes is fixed. However, to allow for comparisons across societies with different set of individuals, we define our indices for all population sizes. For a society, we summarize the achievement of every individual in every attribute by an achievement matrix. A social welfare index is defined as a real-valued function on the set of possible achievement matrices. We propose a new class of multidimensional social welfare indices and characterize them axiomatically. Indices in this class are constructed in two stages. First, an overall achievement score is obtained for each individual by aggregating over the different attributes of well-being and then these scores are aggregated across individuals. In each stage of this aggregation, we use a generalized mean, which is characterized by a single parameter. Therefore, we refer to our new two-stage welfare indices as the class of *two-parameter generalized mean social welfare indices*.[1] The class includes several indices proposed in the literature, such as those of Foster et al. (2005) and Decancq and Ooghe (2009), as special cases. Indices in our class are particularly amenable for empirical applications because of their simple functional form. Seth (2009) has used this new class of indices to critically evaluate the Human Development Index.

[1] Since writing this dissertation, we have learned that Kockläuner (2006) has proposed a similar class of indices for measuring poverty and has discussed some of its properties. See also Kockläuner (2008).

A satisfactory index of social welfare should be sensitive to the inequality in the distributions of the attributes of well-being (Atkinson, 1970; Foster and Sen, 1997). Aside from its direct concern, inequality may well have negative indirect effects on social welfare. For example, high levels of inequality can lead to political instability, tensions among ethnic groups, increase in crime rates, and feelings of deprivation among the members of society.

When there are multiple attributes of well-being, there are two distinct forms of inequality. The first is concerned with the dispersion across the individual achievements of each attribute (Kolm, 1977) and the second is concerned with the correlation — or more precisely, association — across attributes (Atkinson and Bourguignon, 1982). The first form of inequality is *distribution sensitive inequality* and the second is *association sensitive inequality*. Many multidimensional indices of social welfare, inequality, or poverty, such as the Human Development Index, Human Poverty Index, and various physical quality of life indices, are insensitive to either of these forms of inequality, whereas others, such as those proposed by Hicks (1997), Foster et al. (2005), Gajdos and Weymark (2005), and Alkire and Foster (2008) only take account of distribution sensitive inequality. There have also been a small number of multidimensional indices proposed that take account of both kinds of inequality. See, for example, Tsui (1995, 1999, 2002), Bourguignon (1999), Bourguignon and Chakravarty (2003), Decancq and Lugo (2009), and Decancq and Ooghe (2009).

This class of indices developed in this chapter is most closely related to that of Foster et al. (2005). They also constructed a class of welfare indices by applying a two-stage aggregation procedure in which a generalized mean is used in each stage. However, they used the same generalized mean parameter in both stages. Using a single parameter is quite restrictive because it is then not possible for their indices to be association sensitive. By using two parameters, our indices can be both distribution and association sensitive.

Bourguignon (1999) has also proposed a two-parameter class of indices, albeit in the context of measuring inequality. Each of Bourguignon's indices is a monotonic transform of one of our indices. However, Bourguignon does not provide an axiomatic characterization of his class. Furthermore, as discussed later, the value of his welfare indices can respond to a change in the inequality aversion parameter in a way that is counter intuitive.

The rest of this chapter is organized as follows. In the second section, we introduce our basic definitions and notation. In the third section, we define and discuss the class of two-parameter generalized mean social welfare indices. In the fourth section, we introduce the non-distributional axioms and use them to characterize our class defined in the third section. We then, in the fifth section, introduce our inequality aversion axioms and characterize the subclasses of our class of indices that satisfy them. We consider other subclasses of our indices in the sixth section. In the final section, we discuss possible extensions of our analysis and provide some concluding remarks.

Preliminaries

The set of attributes of well-being is $\mathbf{D} = \{1, \ldots, D\}$, where $D \subset \mathbb{N}$ is the number of attributes.[2] Throughout the analysis D is fixed with $D \geq 2$. For example, the attributes could be income, years of education, and an index of health status. Alternatively, the attributes of well-being could be incomes in different time periods or states of nature. The former would be appropriate for studying income inequality over time, whereas the latter

[2]We use the following standard notation. The set $\mathbb{N}$ is the set of positive integers. The Euclidean k-space is $\mathbb{R}^k$ and it's non-negative and positive orthants are $\mathbb{R}^k_+$ and $\mathbb{R}^k_{++}$, respectively. It is sometime convenient to think of a $j \times k$ real-valued matrix as being a vector in $\mathbb{R}^{jk}$. The D-dimensional simplex is $S^D = \{x \in \mathbb{R}^{D+1}_+ \mid \sum_{i=1}^{D+1} x_i = 1\}$. The interior of S^D is denoted by $Int(S^D)$.

would be appropriate for studying income inequality under uncertainty (Ben Porath et al., 1997). The set of individuals is $\mathbf{N} = \{1, \ldots, N\}$. We let the population size vary, so N can be any integer in $\mathbb{N}$.

The quantity of an attribute obtained by an individual is referred to as an *achievement*. An *achievement matrix* for a population of size N is a matrix $H \in \mathbb{R}^{ND}_{++}$, whose nd^{th} entry is the achievement h_{nd} of attribute d by person n. The n^{th} row $h_{n\cdot}$ of H is the vector listing the achievements of all D attributes by person n. The d^{th} column $h_{\cdot d}$ of H is the vector listing the achievements of all N individuals for attribute d. Let $\mathcal{H}_N$ denote the set of all possible achievement matrices of population size N and let $\mathcal{H} = \cup_{\mathbf{N}\subset\mathbb{N}}\mathcal{H}_N$ be the set of all possible achievement matrices.

A *social welfare index* is a function $W : \mathcal{H} \rightarrow \mathbb{R}$. The social welfare associated with the achievement matrix $H \in \mathcal{H}_N$ is at least as large as the social welfare associated with the achievement matrix $H' \in \mathcal{H}_{N'}$ if and only if $W(H) \geq W(H')$. The matrices H and H' could be for societies with different sets of individuals, as would be the case when making comparison between different countries or regions. Of course, if they are the achievement matrices for a single society, then N must equal N'.

We employ the following operations on vectors and matrices. For all $M \in \mathbb{N}$ and all $x, y \in \mathbb{R}^M$, the *join* of x and y is $(x \vee y) = (\max(x_1, y_1), \ldots, \max(x_M, y_M))$ and the *meet* of x and y is $(x \wedge y) = (\min(x_1, y_1), \ldots, \min(x_M, y_M))$. For all $r, M \in \mathbb{N}$ and all $z \in \mathbb{R}^M$, the *r-replication* of z is the vector $[z]_r = (z, \ldots, z) \in \mathbb{R}^{rM}$ in which z has been replicated r times. Similarly, for all $r, L, M \in \mathbb{N}$ and all $Y \in \mathbb{R}^{LM}$, the *r-replication* of Y is the matrix $[Y]_r \in \mathbb{R}^{L'M}$ in which the rows of Y has been replicated r times, where $L' = r \cdot L$.

The following special vectors and matrices are used in the subsequent discussion. The

M vector whose components are all equal to 1 is $\mathbf{1}_M$. Similarly, the $L \times M$ matrix $\mathbf{1}_{LM}$ is the matrix with a 1 in every entry. The M vector whose components are all equal to $1/M$ is ξ_M.

A Class of Indices

The class of social welfare indices that is introduced here is defined using generalized means. For vectors in $\mathbb{R}^M_{++}$, for all $\gamma \in \mathbb{R}$, and all $a \in \mathbb{R}^M_+$, the *generalized mean of order* γ for the weight vector $a \in S^{M-1}$ is the function $\mu^M_\gamma(\cdot; a)$ on $\mathbb{R}^M_{++}$ defined by setting, for all $x \in \mathbb{R}^M_{++}$,

$$\mu^M_\gamma(x; a) = \begin{cases} \left[\sum_{m=1}^M a_m x_m^\gamma\right]^{1/\gamma} & \text{if } \gamma \neq 0 \\ \prod_{m=1}^M x_m^{a_m} & \text{if } \gamma = 0 \end{cases}. \tag{1}$$

The parameter γ determines the curvature of the level surfaces of μ^M_γ. For $\gamma = 1$, a generalized mean is simply a weighted arithmetic mean. It is a weighted geometric mean and a weighted harmonic mean for $\gamma = 0$ and $\gamma = -1$, respectively. As $\gamma \to \infty$, $\mu^M_\gamma(x; a) \to \max_{m \in M}\{x_m\}$, and as $\gamma \to -\infty$, $\mu^M_\gamma(x; a) \to \min_{m \in M}\{x_m\}$.[3] Of particular interest are generalized means in which all attributes receive the same weight. That is, in (1), the weight vector a is equal to ξ_M. Note that a generalized mean is twice differentiable.

It is common in the literature on multidimensional social welfare and inequality to construct an overall index in two stages. This can be done by either (i) first aggregating across individuals for each attribute and then aggregating across attributes or (ii) first aggregating across attributes for each individual and then aggregating across individuals. Following Pattanaik et al. (2007), the former method is called *column-first two-stage aggregation* and the latter is called *row-first two-stage aggregation*. Pattanaik et al. (2007, Propositions 1

[3] We require that γ be in $\mathbb{R}$ and thereby exclude the limiting cases of $\gamma = \infty$ and $\gamma = -\infty$.

and 2) have shown that the column-first procedure completely ignores interactions across dimensions, which is important if the index is to be association sensitive. Thus, here, we only consider the row-first procedure. In the first stage, achievements are aggregated to obtain an individual's *overall achievement score*. For a population size of $N \in \mathbb{N}$, the overall achievement score for individual n is obtained by applying an aggregation function $Q_n^N : \mathbb{R}_{++}^D \to \mathbb{R}$ for all n in $\mathbf{N}$. Then, in the second stage, these scores are aggregated using a function $\Phi_N : \mathbb{R}^N \to \mathbb{R}$. Formally, the *row-first two-stage aggregation* method can be defined as follows.

Row-First Two-Stage Aggregation For every $\mathbf{N} \subset \mathbb{N}$ and every n in $\mathbf{N}$, there exist functions $\Phi_N : \mathbb{R}^N \to \mathbb{R}$ and $Q_n^N : \mathbb{R}_{++}^D \to \mathbb{R}$ such that for all $H \in \mathcal{H}_N$, the social welfare index W can be written as

$$W(H) = \Phi_N(Q_1^N(h_{1\cdot}), \ldots, Q_N^N(h_{N\cdot})). \tag{2}$$

The indices we propose use generalized means for each stage of the aggregation. For every choice of the parameters α and β in $\mathbb{R}$ and every weight vector a in $Int\left(S^{D-1}\right)$, the *two-parameter generalized mean social welfare index* $W(\cdot; \alpha, \beta, a)$ is defined by setting

$$W(H; \alpha, \beta, a) = \mu_\alpha^N(\mu_\beta^D(h_{1\cdot}; a), \ldots, \mu_\beta^D(h_{N\cdot}; a); \xi_N). \tag{3}$$

for every $\mathbf{N} \subset \mathbb{N}$ and every $H \in \mathcal{H}_N$. Note that (3) is obtained from (2) by setting $\Phi_N(\cdot) = \mu_\alpha^N(\cdot; \xi_N)$ and $Q_n^N(\cdot) = \mu_\alpha^D(\cdot; a)$ for all $\mathbf{N} \subset \mathbb{N}$ and for all n in $\mathbf{N}$. Intuitively, the index is a generalized mean of generalized means. Let $\mathcal{G}$ denote the set of all two-parameter generalized mean social welfare indices.

Following Atkinson (1970), $\mu_\alpha^N(x)$ is referred to as the *equally distributed equivalent overall achievement*, where x is the vector of overall achievements. The parameter α measures society's aversion towards inter-personal inequality in these achievements. That is, α measures the degree to which one individual's overall achievement is substitutable for a second individual's overall achievement in the social welfare index W. Similarly, the parameter β measures the degree of substitutability across the dimensions of well-being of any individual.

In defining the class of indices $\mathcal{G}$, we have not required that they be either distribution or association sensitive. As we shall show, such sensitivity can be achieved by placing restrictions on the parameters that define these indices. In the subsequent sections, under the maintained assumption that we use row-first aggregation, we shall provide an axiomatic characterization of the class of all two-parameter generalized mean social welfare indices, as well as characterizations of the sub-classes that satisfy distribution sensitivity, association sensitivity, or both of these properties together.

Non-Distributional Axioms

In this section, we axiomatically characterize the two-parameter class of generalized mean social welfare indices $\mathcal{G}$ given our assumption that the index is constructed using row-first aggregation, that is, assuming that the social welfare index W has the form in (2). The axioms that we employ are standard in the literature. Furthermore, none of the axioms considered in this section take into account distributional or associational concerns.

The first axiom requires the value of social welfare index to change continuously with a change in the achievement of any person in any dimension.

Continuity (CONT) For every $\mathbf{N} \subset \mathbb{N}$, W is continuous on $\mathbb{R}^{ND}_{++}$.

The next axiom imposes convenient normalizations on the aggregation function Q and the social welfare index W. If an individual has the same achievement in all dimensions, then the overall achievement is equal to this value. Moreover, if everybody has the same overall achievements, then the value of the social welfare index is equal to this common value.

Normalization (NORM) For every $\mathbf{N} \subset \mathbb{N}$, every $\zeta > 0$, and every $H \in \mathcal{H}_N$ such that $H = \zeta 1_{ND}$,

$$Q_n^N(h_{n.}) = \zeta \ \forall n \in \mathbf{N} \quad \text{and} \quad W(H) = \zeta.$$

The social welfare index can be thought of as being a representation of a social preference on the set of achievement matrices. We assume that this preference is homothetic. A preference is *homothetic* if whenever two achievement matrices for the same population are socially indifferent, then so are the achievement matrices obtained by proportionally scaling both of them. By assuming that this preference is homothetic, we are implicitly assuming that we are concerned with relative inequality; that is, there is no change in inequality if an achievement matrix is proportionally scaled.[4]

Homotheticity (HOM) For every $\mathbf{N} \subset \mathbb{N}$, every $\delta > 0$, and every $H, H' \in \mathcal{H}_N$,

$$W(H') = W(H) \Leftrightarrow W(\delta H') = W(\delta H).$$

We assume that the identities of individuals are not ethically significant. This is accomplished by requiring the social welfare index to be symmetric in the sense that the index is

[4] Tsui (1995) introduced a stronger version of homotheticity axiom called *ratio scale invariance*, which has also been used by Decancq and Ooghe (2009). However, this axiom has been questioned by Bourguignon (1999, p. 479). For a related discussion, see Weymark (2006, p. 311).

invariant with respect to permutations of the individual achievement vectors.

Anonymity (ANON) For every $\mathbf{N} \subset \mathbb{N}$, every $H, H' \in \mathcal{H}_N$, and for every permutation matrix $P \in \mathbb{R}_+^{NN}$ such that $H' = PH$,

$$W(H') = W(H).$$[5]

The preceding axioms do not place any restrictions on the value of the index for achievement matrices for societies with different population sizes. We assume that if an achievement matrix is replicated an arbitrary number of times, then the value of the social welfare index is unchanged. Thus, social welfare is being measured in per capita terms.

Population Replication Invariance (POPRI) For every $r \in \mathbb{N}$ and every $H, H' \in \mathcal{H}$ such that $H' = [H]_r$,

$$W(H') = W(H).$$

We assume that each attribute of well-being contributes positively to social welfare. It is, therefore, natural to assume that the value of the social welfare function increases if the value of some attribute for some individual increases with no decrease in the value of any attribute for any individual.

Monotonicity (MON) For every $\mathbf{N} \subset \mathbb{N}$ and every $H, H' \in \mathcal{H}_N$ such that $H' \geq H$ and $H' \neq H$,

$$W(H') > W(H).$$

[5] A permutation matrix is a square matrix with each row and column having exactly one element equal to one and the rest equal to zero.

The restriction of the social welfare index to achievement matrices in $\mathcal{H}_N$ provides an index of social welfare for any group of size N. We assume that social welfare increases if the social welfare of a subgroup of the society increases, while that of the rest of the population is unchanged. This increase in subgroup social welfare may be accompanied by both increases and decreases in achievements of individuals in the subgroup. Our monotonicity axiom does not apply to such comparisons.

Subgroup Consistency (SUBCON) For every $N_1, N_2, N \in \mathbb{N}$ such that $N_1 + N_2 = N$, every $H_1, H_1' \in \mathcal{H}_{N_1}$, and every $H_2, H_2' \in \mathcal{H}_{N_2}$, if $W(H_1') > W(H_1)$ and $W(H_2') = W(H_2)$, then $W(H_1', H_2') > W(H_1, H_2)$.

It is common in empirical analysis for an individual's overall achievement score to be obtained by taking a weighted sum of his achievements in each dimension. These weights could measure the relative importance of the different achievements. See, for example, Decancq and Lugo (2008). Alternatively, they can be used to convert the units for each dimension into a common scale. Suppose that the set of achievements $\mathbf{D}$ is partitioned into two disjoint subsets $\mathbf{D}_1$ and $\mathbf{D}_2$. For given values of the achievements in $\mathbf{D}_2$, the aggregation function Q_n^N for person n in a row-first two-stage aggregation procedure defines a conditional ordering of achievement vectors for the attributes in $\mathbf{D}_1$. When fixed weights are used to aggregate the attributes in $\mathbf{D}$, this conditional ordering is independent of the values in $\mathbf{D}_2$. We do not assume *a prori* that fixed weights are used in this aggregation. However, we do assume that for every partition of $\mathbf{D}$ into disjoint subsets $\mathbf{D}_1$ and $\mathbf{D}_2$, the aggregation function Q_n^N defines a conditional ordering of achievement vectors for the attributes in $\mathbf{D}_1$ that is independent of the values of the attributes in $\mathbf{D}_2$. That is, Q_n^N is assumed to be *completely strictly separable.* More precisely, we assume that Q_n^N is additively separable for all n in $\mathbf{N}$.[6]

Additive Separability (ADDSEP) For every $\mathbf{N} \subset \mathbb{N}$ and every $n \in \mathbf{N}$, the aggregation function Q_n^N can be written as

$$Q_n^N(h_{n\cdot}) = U_n(V_1^n(h_{n1}) + \cdots + V_D^n(h_{nD})) \tag{4}$$

for all $h_{n\cdot} \in \mathbb{R}_{++}^D$, where $U_n : \mathbb{R} \to \mathbb{R}$ is a continuous and increasing function, and $V_d^n : \mathbb{R}_{++} \to \mathbb{R}$ is a continuous function for all d in $\mathbf{D}$.

For row-first two-stage aggregation, Theorem 1 shows that the non-distributional axioms introduced in this section characterize the set of two-parameter generalized mean social welfare indices $\mathcal{G}$.

Theorem 1 *An index $W : \mathcal{H} \to \mathbb{R}$ is a two-parameter generalized mean social welfare index if and only if W is obtained using row-first two-stage aggregation and satisfies CONT, NORM, HOM, ANON, POPRI, MON, SUBCON, and ADDSEP.*

Proof. See Appendix A. ■

Inequality Sensitivity Axioms

In this section, we introduce axioms that are concerned with the sensitivity of the social welfare indices to the two forms of inequality described above. First, we introduce a distribution sensitivity axiom that ensures that the social welfare index takes account of the spread

[6] Additive separability of Q_n^N is equivalent to complete strict separability if $D \geq 3$. However, for $D = 2$, additive separability is a somewhat stronger assumption than complete strict separability. See Blackorby et al. (1978, Section 4.4).

of the multidimensional distribution and we then characterize the subclass of the class of two-parameter generalized mean social welfare indices $\mathcal{G}$ that satisfies this axiom. Next, we introduce two alternative association sensitivity axioms and we characterize the subclasses of $\mathcal{G}$ that satisfy each of these axioms. Finally, we characterize the subclasses of $\mathcal{G}$ that satisfy both our distribution sensitivity axiom and one of our association sensitive axioms.

Distribution Sensitive Inequality

Distributional sensitivity of the social welfare index W is obtained by requiring that the value of the index increases if an achievement matrix is subjected to a common smoothing. For every $\mathbf{N} \subset \mathbb{N} \backslash \{1\}$ and every $H', H \in \mathcal{H}_N$, H' is obtained from H by a *common smoothing* if there exists a bistochastic matrix B such that $H' = BH$ and H' is not a permutation of H.[7] Note that the same bistochastic matrix is being applied to each attribute. Formally, we require our social welfare index to satisfy the following axiom due to Kolm (1977).

Increasing under Common Smoothing (ICS) For every $\mathbf{N} \subset \mathbb{N} \backslash \{1\}$ and every $H', H \in \mathcal{H}_N$ such that H' is obtained from H by a common smoothing,

$$W(H') > W(H).^{8}$$

When there is only one dimension of well-being, H' and H are distributions of a single attribute, and the requirement that H' be obtained from H by a common smoothing is equivalent to saying that H' can be obtained from H by a sequence of Pigou-Dalton transfers, possibly supplemented by permutations of some of the distributions in this sequence.

[7] A bistochastic matrix is a non-negative square matrix whose row and column sums are both equal to one.

[8] This axiom is also known as the Uniform Majorization Principle. See Kolm (1977) and Weymark (2006) for further discussion of this and related distribution sensitivity axioms.

Theorem 2 characterizes the subclass of $\mathcal{G}$ that satisfies ICS.

Theorem 2 *A two-parameter generalized mean social welfare* index $W(H; \alpha, \beta, a)$ *satisfies ICS if and only if* $\alpha < 1$ *and* $\beta < 1$.

Proof. See Appendix B. ■

In the definition of a generalized mean μ_γ^M, the parameter γ determines the curvature of the level surfaces (iso-achievement curves) of μ_γ^M. The restriction $\beta < 1$ implies that the aggregation function Q is strictly quasi-concave and thus has a strictly convex upper contour set. Consequently, the overall achievement score increases when one achievement vector is obtained from the second by a strictly convex combination of the achievements of the latter. Note that the first stage aggregation function is analogous to the constant elasticity of substitution (CES) function in the utility analysis. Similarly, if $\alpha < 1$, then the aggregation function Φ is also strictly quasi-concave in its arguments which are the overall achievements of the individuals.

Association Sensitive Inequality

We now consider the sensitivity of the social welfare index W to a change in the association between dimensions while leaving the marginal distributions unaltered.[9] Association sensitivity was introduced into the literature on multidimensional social welfare by Atkinson and Bourguignon (1982) and has subsequently been considered by Tsui (1995, 1999, 2002), Bourguignon (1999), Bourguignon and Chakravarty (2003), and Decancq and Lugo (2009),

[9] What we refer to association between dimensions is often called dependence in the statistics literature and correlation in the literature on economic inequality. We do not employ the term 'correlation' here so as to emphasize that we are not restricting our attention to the correlation coefficient used in statistics.

among others. There are various ways in which the different dimensions of well-being may be interdependent, with the consequence that there are a number of different concepts of association sensitivity. See Joe (1997, Chapter 2) for a discussion.

Here, association sensitivity of W is obtained by requiring that the value of the index increases if an achievement matrix is subjected to an association increasing transfer. For every $\mathbf{N} \subset \mathbb{N}\backslash\{1\}$ and every $H, H' \in \mathcal{H}_N$, H' is obtained from H by an *association increasing transfer* if $H' \neq H$, H' is not a permutation of H, and there exist two individuals n_1 and n_2 such that $h'_{n_1\cdot} = (h_{n_1\cdot} \vee h_{n_2\cdot})$, $h'_{n_2\cdot} = (h_{n_1\cdot} \wedge h_{n_2\cdot})$, and $h'_{n\cdot} = h_{n\cdot}$ for all $n \in \mathbf{N}\backslash\{n_1, n_2\}$.[10] To interpret this definition, consider two individuals and an achievement matrix such that neither individual has at least as much of every attribute than the other. If for each attribute, we reallocate their achievements between these two individuals so that one of them has at least as much of every achievement as the other, then the resulting achievement matrix has been obtained from the former by an association increasing transfer. As emphasized by Bourguignon and Chakravarty (2003), whether an association increasing transfer is socially beneficial depends on whether the attributes are substitutes or complements in W. As a consequence, we have the following two different association sensitivity axioms, the choice of which depends on which of these two cases apply.

Decreasing under Increasing Association (DIA) For every $\mathbf{N} \subset \mathbb{N}\backslash\{1\}$ and every $H', H \in \mathcal{H}_N$ such that H' is obtained from H by a finite sequence of association increasing transfers,

$$W(H') < W(H).$$

[10] The concept of an association increasing transfer was introduced by Tsui (1999) under the name of a correlation increasing transfer. Tsui's concept was in turn was based on the idea of a basic rearrangement due to Boland and Proschan (1988). These concepts are closely related to the correlation increasing switches considered by Bourguignon and Chakravarty (2003). For formal definitions of these concepts, see the articles cited above and for a discussion of the relationship between them, see Chakravarty (2009) and Seth (2009).

Increasing under Increasing Association (IIA) For every $\mathbf{N} \in \mathbb{N}\backslash\{1\}$ and every $H', H \in \mathcal{H}_N$ such that H' is obtained from H by a finite sequence of association increasing transfers,

$$W(H') > W(H).$$

Theorem 3 characterizes the subclasses of $\mathcal{G}$ that satisfy these axioms.

Theorem 3 (i) *A two-parameter generalized mean social welfare index $W(H;\alpha,\beta,a)$ satisfies DIA if and only if $\alpha < \beta$.* (ii) *A two-parameter generalized mean social welfare index $W(H;\alpha,\beta,a)$ satisfies IIA if and only if $\alpha > \beta$.*

Proof. See Appendix C. ■

After an association increasing transfer takes place, one of the two individuals affected by the transfer has at least as much of every attribute as the other affected individual. If the attributes are substitutes (resp. complements) from the perspective of social welfare, then such a transfer should decrease (resp. increase) the value of the social welfare index, which requires that α is less than (resp. larger than) β. For example, if two of the attributes are income and some indicator of health status, then it is natural to regard them as being substitutes because an individual with poor health can better deal with his condition if he has sufficient funds to help ameliorate this situation. On the other hand, if quality of health and housing infrastructure are two attributes of well-being, then good health is better enjoyed by an individual whose housing infrastructure is improved as well. In this situation, these two attributes are complements to each other.

Sensitivity to Both Forms of Inequality

By combining Theorems 2 and 3, we obtain the subclasses of $\mathcal{G}$ that are both distribution and association sensitive.

Theorem 4 (i) *A two-parameter generalized mean social welfare index* $W(H;\alpha,\beta,a)$ *satisfies ICS and DIA if and only if* $\alpha<\beta<1$*.* (ii) *A two-parameter generalized mean social welfare index* $W(H;\alpha,\beta,a)$ *satisfies ICS and IIA if and only if* $\beta<\alpha<1$*.*

To illustrate the significance of the parameter restrictions in Theorem 4, we consider the problem of a policy maker who needs to decide which person to allocate a marginal transfer T in his budget so as to maximize the increase in social welfare. For simplicity, in the following discussion we suppose that both α and β are non-zero. For any $\mathbf{N}\subset\mathbb{N}$ and any $H\in\mathcal{H}_N$, if the transfer T is provided to person n to improve her achievement in dimension d, then the increment in social welfare is:

$$\frac{\partial W(H;\alpha,\beta,a)}{\partial T}=\left(a_d h_{nd}^{\beta-1}C_n^{\alpha-\beta}\mathbf{C}\right)c_{nd},$$

where a_d is the weight of dimension d in the calculation of the overall achievement scores, $c_{nd}=\partial h_{nd}/\partial T$ is the increase in achievement h_{nd} due to the transfer, $C_n=\mu_\beta\left(h_{n\cdot};a\right)$ is the overall achievement score of person n, and $\mathbf{C}=\frac{1}{N}W(H;\alpha,\beta,a)^{1-\alpha}$. Note that $\mathbf{C}$ is identical across all individuals. Let $\omega_{nd}=a_d h_{nd}^{\beta-1}C_n^{\alpha-\beta}c_{nd}$ for all n and all d. To maximize the increase in social welfare, the policy maker should assist person n to increase her achievement in dimension d if

$$\omega_{nd}>\omega_{n'd'}\ \forall n'\in\mathbf{N}/\{n\}\text{ and }\forall d'\in\mathbf{D}/\{d\}. \tag{5}$$

First, to illustrate the role that the restriction $\alpha < 1$ plays, we consider the situation in which $h_{nd} = \bar{h}_n$ for all d and $c_{nd} = \bar{c}$ for all d and all n. In this case, $\omega_{nd} = h_n^{\alpha-1}\bar{c}$. Consider the problem of determining which individual the budget increase should be spent on. Because $\alpha < 1$, the policy maker should provide the transfer to the individual or individuals for which $\bar{h}_n$ is minimal.

Second, we consider the role that the restriction $\beta < 1$ plays. This role is most clearly seen when $a_d = \bar{a}$ for all d and $c_{nd} = \bar{c}$ for all d and all n. Consider the problem of determining which attribute the budget increase should be spent on conditional on individual n being the person receiving the transfer. Because C_n does not depend of d and because $\beta < 1$, it follows from (5) that the transfer should be spent on the attribute or attributes for which h_{nd} is minimal.

Third, to illustrate how the substitutability and complementarity between attributes affects the allocation of the transfer, we again consider the situation in which $a_d = \bar{a}$ for all d and $c_{nd} = \bar{c}$ for all d and all n. We already know that if individual n receives a transfer, the transfer should be spent on the attribute or attributes for which h_{nd} is minimal. If the social welfare index is not association sensitive, then $\alpha = \beta$ and thus $\omega_{nd} = \bar{a}\bar{c}h_{nd}^{\beta-1}$. Hence, the transfer should be allocated to the individuals and attributes for which h_{nd} is minimal regardless of what anybody's overall achievement score is. If, however, the social welfare index is association sensitive, then $\alpha \neq \beta$ and thus $\omega_{nd} = \bar{a}\bar{c}h_{nd}^{\beta-1}C_n^{\alpha-\beta}$ and the transfer should be allocated to those individuals and attributes for which $h_{nd}^{\beta-1}C_n^{\alpha-\beta}$ are maximal. Suppose that $h_{nd} = h_{n'd'}$, where d (resp. d') is the attribute with minimal achievement for individual n (resp. n'). Then, the transfer should not go to individual n' if the attributes are substitutes ($\alpha < \beta$) and $C_{n'} > C_n$. Similarly, the transfer should not go to individual n' if the attributes are complements ($\alpha > \beta$) and $C_{n'} < C_n$. When the attributes are substitutes (resp. complements), then higher (resp. lower) association is detrimental to social welfare and, thus, the individual with the lower (resp. higher) overall achievement score should be favored whenever they have the same minimal achievements.

Related Social Welfare Indices

Foster et al. (2005) have proposed a one-parameter class of generalized mean social welfare indices, which we refer to as the FLS class. The FLS class is the subclass of our two-parameter generalized means $\mathcal{G}$ obtained by setting $\alpha = \beta \leq 1$. The FLS indices exibit distribution sensitivity, but as can be seen from Theorem 4, they are not association sensitive. When $\alpha = \beta = 1$, the social welfare index is simply the arithmatic mean across individuals of weighted arithmatic means across attributes. This index is neither association nor distribution sensitive. Several well-known indices are simple means of weighted arithmetic means. For example, the Human Development Index (United Nations Development Programme, 2006) and the Morris (1979) physical quality of life index have this functional form.

For the FLS class, both column-first two-stage aggregation and row-first two-stage aggregation yield an identical evaluation. This invariance property is called *path independence*.

Path Independence (PATHIN) For every $\mathbf{N} \subset \mathbb{N}$, there exist functions $\Phi : \mathbb{R}^N_{++} \to \mathbb{R}_{++}$ and $Q : \mathbb{R}^D_{++} \to \mathbb{R}_{++}$ such that for all $H \in \mathcal{H}_N$,

$$\Phi(Q(h_{1\cdot}), \ldots, Q(h_{N\cdot})) = Q(\Phi(h_{\cdot 1}), \ldots, \Phi(h_{\cdot D})).$$

Note that the class of two-parameter generalized mean social welfare indices cannot be simultaneously association sensitive and path independent. If the data for different attributes are available at different levels of aggregation, we do not have enough information to consider association among attributes. For example, education data may be available at the individual level, income data may be available at the household level, and health data may be available at the municipality level. In such circumstances, it may be appropriate to require the social welfare index to be path independent. Subclasses of $\mathcal{G}$ that satisfy PATHIN are characterized in Theorem 5.[11]

Theorem 5 (i) *A two-parameter generalized mean social welfare index $W(H;\alpha,\beta,a)$ satisfies PATHIN if and only if $\alpha=\beta$.* (ii) *A two-parameter generalized mean social welfare index $W(H;\alpha,\beta,a)$ satisfies PATHIN and ICS if and only if $\alpha=\beta<1$.*

Proof. For any $N\in\mathbf{N}$ and any $H\in\mathcal{H}_N$, let $W_1=\mu_\alpha^N(\mu_\beta^D(h_{1\cdot};a),\ldots,\mu_\beta^D(h_{N\cdot};a);\xi_N)$ and $W_2=\mu_\beta^D(\mu_\alpha^N(h_{\cdot 1};\xi_N),\ldots,\mu_\alpha^N(h_{\cdot D};\xi_N),a)$. It is straightforward to show that if $\alpha=\beta$, then $W_1=W_2$. By Hardy et al. (1934, Theorem 26), $W_1>W_2$ if $\beta<\alpha$ and $W_1<W_2$ if $\beta>\alpha$.[12] Hence, $W_1\neq W_2$ if $\alpha\neq\beta$. Part (ii) of the theorem follows directly by combining part (i) with Theorem 2. ■

The subclass of $\mathcal{G}$ for which $\alpha\in(0,1)$ and $\beta<1$ shares the same ordinal properties as the class of welfare indices proposed by Bourguignon (1999). For $a\in Int(S^{D-1})$, $\alpha\in(0,1)$, and $\beta<1$, the Bourguignon social welfare index is defined as

$$W_B(H;\alpha,\beta,a)=\frac{1}{N}\sum\nolimits_{n=1}^{N}(\mu_\beta^D(h_{n\cdot};a))^\alpha=(W(H;\alpha,\beta,a))^\alpha\,, \tag{6}$$

[11] For a class of path independent standard of living indices, see Dutta et al. (2003).

[12] Although Hardy et al. (1934) assume that both α and β are positive, their proof can be easily extended for all α and β in $\mathbb{R}$.

for all $N \subset \mathbf{N}$ and all $H \in \mathcal{H}_N$. Thus, our index $W(H;\alpha,\beta,a)$ is a monotonic transformation of the corresponding Bourguignon index.

By using the inequality aversion parameter α to transform $W(H;\alpha,\beta,a)$ as in (6), it is unclear how to interpret a comparison of welfare levels for different values of α. To see why, consider any $N \in \mathbf{N}$ and suppose that there are two societies with achievement vectors H, H' $\in \mathcal{H}_N$ such that $h_{n\cdot} = h'_{n\cdot} = h$ for all n. In this situation, $W_B(H;\alpha,\beta,a) \neq W_B(H';\alpha',\beta,a)$ for any $\alpha \neq \alpha'$. However, it is not clear why differences in inequality aversion should result in different levels of social welfare when everybody has the same achievement vector.

Bourguignon has used his welfare index to construct an inequality index by setting $I_B(H;\alpha,\beta,a) = 1 - W_B(H;\alpha,\beta,a)/W_B(\bar{H};\alpha,\beta,a)$, where $\bar{H} = BH$ and $B = \mathbf{1}_{NN}/N$. It is shown in Seth (2009) that for some $\alpha > \alpha' > \alpha''$, $I_B(H;\alpha,\beta,a) < I_B(H;\alpha',\beta,a) > I_B(H;\alpha'',\beta,a)$. Thus, with this index, inequality is not monotonically increasing in the inequality aversion parameter for a given achievement matrix.

Recently, Decancq and Ooghe (2009) have proposed a class welfare indices that are also constructed using a row-first two-stage aggregation procedure. In the first stage, they use the geometric mean μ_0^D to aggregate across attributes and in the second stage, they use a generalized mean μ_α^N with $\alpha < 0$ to aggregate across individuals. This procedure implicitly assumes that attributes are substitutes and thus their indices can only satisfy IDA but not IIA. Note that the Decancq-Ooghe class is a subclass of $\mathcal{G}$.

Conclusion

In this chapter, we have proposed a class of two-parameter generalized mean social welfare indices and characterized it axiomatically. Under appropriate parametric restrictions, we have shown that these indices are both distribution and association sensitive. Because of their simple functional structure, our indices are easy to implement empirically. We have also shown that the indices proposed by Foster et al. (2005) and Decancq and Ooghe (2009), as well as the Human Development Index, are subclasses of our indices. We have also discussed how indices are related to the Bourguignon class of indices.

Our indices proposed here assume that the degree of substitution between each pair of attributes is the same. As a consequence, all attributes are either substitutes or complements to each other. A natural extension of our analysis would be to construct a more general class of indices that would treat some attributes as substitutes, while simultaneously treating other attributes as complements.

Following Tsui (1995), we have only considered association increasing transfers of the kind introduced by Boland and Proschan (1988). Alternative concepts of dependence among attributes could be used to construct indices based on them. Decancq (2009) has done this for positive orthant dependence.

Seth (2009) has used the indices proposed in this chapter to measure social welfare in Mexico using 2000 census data and has found that the ranking of Mexican states differs when association sensitivity is taken into account than when it is not. In Chapter four, I apply the index to the Indian context showing how the consideration for inequality may alter the state level rankings.

Appendix

A. Proof of Theorem 1

Proof. The sufficiency part of the proof is straightforward. To prove necessity, suppose that W is obtained using row-first two-stage aggregation, i.e., W takes the form (2), and that it satisfies CONT, NORM, HOM, ANON, POPRI, MON, SUBCON, and ADDSEP.

Consider any $\mathbf{N}$ and any $\hat{H} \in \mathcal{H}_N$ such that $\hat{h}_{nd} = \hat{x}_n$ for every $d \in \mathbf{D}$. By NORM, $Q_n^N(\hat{h}_{n\cdot}) = \hat{x}_n$ for every $n \in \mathbf{N}$ and, hence, $W(\hat{H}) = \Phi_N(\hat{x})$, where $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_N)$. Let $\bar{H} = P\hat{H}$ for some permutation matrix P. Reasoning as above, $W(\bar{H}) = \Phi_N(\bar{x})$, where $\bar{x}^T = P\hat{x}^T$ and $\hat{x}^T$ is the transpose of $\hat{x}$. ANON implies $W(\hat{H}) = W(\bar{H})$ and therefore $\Phi_N(\hat{x})$ $= \Phi_N(\bar{x})$. Thus, Φ_N is symmetric in its arguments. It follows from NORM that Φ_N is a reflexive function, i.e., $\Phi_N(\zeta, \ldots, \zeta) = \zeta$ for all $\zeta \in \mathbb{R}_{++}$. Consider any $H \in \mathcal{H}_N$ and let $\delta =$ $W(H)$. Define $H^0 \in \mathcal{H}_N$ by setting $h^0_{nd} = \delta$ for all n in $\mathbf{N}$ and d in $\mathbf{D}$. By NORM, it follows that $W(H^0) = W(H)$. Now consider any $\lambda > 0$. Then by HOM we have $W(\lambda H^0) = W(\lambda H)$, and by NORM it follows that $\lambda\delta = W(\lambda H^0)$. We conclude that $\lambda W(H) = W(\lambda H)$ for any λ > 0 and any H in $\mathcal{H}_N$, and so W is homogeneous of degree one. Using the vector $\hat{x}$ defined above, it further follows that $\Phi_N(\lambda\hat{x}) = \lambda\Phi_N(\hat{x})$ and therefore Φ_N is also homogeneous of degree one.

Let $X_N \in \mathbb{R}^N_{++}$ denote the set of all vectors of overall achievement scores with the fixed population size N and let $X = \cup_{N\subset\mathbb{N}} X_N$. Define $\Phi : X \to \mathbb{R}$ so that $\Phi_N(x) = \Phi(x)$ for all N and all $x \in X_N$. The function Φ inherits continuity from W. Furthermore, Φ inherits the analogue of subgroup consistency from W. For any $r \in \mathbb{N}$, let $\tilde{H} = [\hat{H}]_r$, with the same $\hat{H}$ defined earlier. By POPRI, $W(\tilde{H}) = W(\hat{H})$ and therefore Φ satisfies replication invariance because $\Phi(\hat{x}) = \Phi(\tilde{x})$, where $\tilde{x}^T = [\hat{x}^T]_r$ for all $r \in \mathbb{N}$. We have shown that Φ satisfies all the assumptions of the Theorem in Foster and Székely (2008, p. 1149). Thus, there exists a scalar $\alpha \in \mathbb{R}$ such that Φ can be written as

$$\Phi(x) = \begin{cases} \left(\frac{1}{N}\sum_{n=1}^{N} x_n^{\alpha}\right)^{1/\alpha} & \text{if } \alpha \neq 0 \\ \left(\prod_{n=1}^{N} x_n\right)^{1/N} & \text{if } \alpha = 0 \end{cases} \tag{7}$$

for all $x \in X$, where N is the number of components in x.

We now prove that Q_n^N is also a generalized mean. First, for any $\mathbf{N}$, we show that $Q_n^N = Q_{n'}^N$ for all $n, n' \in \mathbf{N}$. Consider any $n, n' \in \mathbf{N}$ and any $\bar{h} \in \mathbb{R}_{++}^D$. Let $H \in \mathcal{H}_N$ be such that $h_{n\cdot} = \bar{h}$ and $h_{\hat{n}\cdot} = \mathbf{1}_D$ for all $\hat{n} \neq n$ and let $H' \in \mathcal{H}_N$ be such that $h'_{n'\cdot} = \bar{h}$ and $h_{\hat{n}\cdot} = \mathbf{1}_D$ for all $\hat{n} \neq n'$. Using NORM and (7), $W(H) = \mu_\alpha^N(1, \ldots, 1, Q_n^N(\bar{h}), 1, \ldots, 1; \xi_N)$ and $W(H')$ $= \mu_\alpha^N(1, \ldots, 1, Q_{n'}^N(\bar{h}), 1, \ldots, 1; \xi_N)$. By ANON, $W(H) = W(H')$. Using the formula for a generalized mean of order α, it now follows that $Q_n^N(\bar{h}) = Q_{n'}^N(\bar{h})$. Hence, $Q_n^N = Q_{n'}^N$ for all $n, n' \in \mathbf{N}$. We denote this common function by Q^N.

Next, we prove that $Q^N = Q^{N'}$ for all $N, N' \in \mathbb{N}$. Consider any $H \in \mathcal{H}_1$. Note that H $= h$ for some $h \in \mathbb{R}_{++}^D$. By (7), $W(H) = Q^N(h)$. Consider any $N \in \mathbb{N}$ and let $\bar{H} = [h]_N$. By (7), $W(\bar{H}) = Q^N(h)$. POPRI implies that $W(\bar{H}) = W(H)$. Hence, $Q^N(h) = Q^1(h)$ for all $h \in \mathbb{R}_{++}^D$ and all $N \in \mathbb{N}$. Therefore, $Q^1 = Q^N$ for all $\mathbf{N} \subset \mathbb{N}$. We denote this common function by Q.

Because $W(\cdot) = Q(\cdot)$ when $N = 1$, Q inherits the properties of continuity, monotonicity, and homogeneity of degree one from W. For all $h \in \mathbb{R}_{++}^D$, ADDSEP implies that $Q(h)$ $= U(\sum_{d=1}^{D} V_d(h))$, where $U : \mathbb{R} \to \mathbb{R}$ is continuous and increasing and $V_d : \mathbb{R}_{++} \to \mathbb{R}$ is continuous for all d. The monotonicity of Q implies that each V_d is also increasing.

Hence, by Eichhorn (1978, Theorem 2.4.1), there exists a scaler $\beta \in \mathbb{R}$ and a weight vector $a \in Int(S^{D-1})$ such that Q can be written as:

$$Q(h) = \begin{cases} \left(\sum_{d=1}^{D} a_d h_d^{\beta}\right)^{1/\beta} & \text{if } \beta \neq 0 \\ \prod_{d=1}^{D} h_d^{a_d} & \text{if } \beta = 0 \end{cases} \tag{8}$$

for all $h \in \mathbb{R}_{++}^{D}$. In other words, the first-stage aggregation function Q is a generalized mean of order β. Therefore, W is a two-parameter generalized mean social welfare index. ■

B. Proof of Theorem 2

The proof of Theorem 2 is based on Lemma B1.

Lemma B1 *For any* $\mathbf{N} \subset \mathbb{N} \backslash \{1\}$*, if* H' *is obtained from* $H \in \mathcal{H}_N$ *by a common smoothing, then (i)* $\sum_{n=1}^{N} G(h'_{n\cdot}) > \sum_{n=1}^{N} G(h_{n\cdot})$ *for strictly concave* G *and (ii)* $\sum_{n=1}^{N} G(h'_{n\cdot}) < \sum_{n=1}^{N} G(h_{n\cdot})$ *for strictly convex* G*.*

Proof. The proof is similar to the proof of Marshall and Olkin (1979, Theorem B.1., p. 433). Consider any $N \in \mathbb{N} \backslash \{1\}$ and suppose that H' is obtained from $H \in \mathcal{H}_N$ by a common smoothing. Thus, $H' = BH$ for some bistochastic matrix B. Denote row n of B by $b_{n\cdot}$. Because H' is not a permutation of H, there exist two individuals n_1 and n_2 such that $h'_{n_1\cdot} \neq h_{n_1\cdot}$ and $h'_{n_2\cdot} \neq h_{n_2\cdot}$. Let $G : \mathbb{R}_{++}^{D} \to \mathbb{R}$ be strictly concave. Strict concavity of G implies $G(h'_{n\cdot}) = G(\sum_{\hat{n}=1}^{N} b_{n\hat{n}} h_{\hat{n}\cdot}) > \sum_{\hat{n}=1}^{N} b_{n\hat{n}} G(h_{\hat{n}\cdot})$ for $n = n_1, n_2$. Because for all $n \in \mathbf{N} \backslash \{n_1, n_2\}$, either $h'_{n\cdot} = h_{n\cdot}$ or $h'_{n\cdot} = \sum_{\hat{n}=1}^{N} b_{n\hat{n}} h_{\hat{n}\cdot}$, it follows that $G(h'_{n\cdot}) \geq \sum_{\hat{n}=1}^{N} b_{n\hat{n}} G(h_{\hat{n}\cdot})$ for all $n \in \mathbf{N} \backslash \{n_1, n_2\}$. Hence, $\sum_{n=1}^{N} G(h'_{n\cdot}) > \sum_{n=1}^{N} \sum_{\hat{n}=1}^{N} b_{n\hat{n}} G(h_{\hat{n}\cdot}) = \sum_{\hat{n}=1}^{N} \sum_{n=1}^{N} b_{n\hat{n}} G(h_{\hat{n}\cdot}) =$

$\sum_{\hat{n}=1}^{N} G(h_{\hat{n}\cdot}) = \sum_{n=1}^{N} G(h_{n\cdot})$. The second part of the lemma can be proved in a similar manner. ∎

Proof of Theorem 2. (a) We first establish sufficiency. That is, we show that if $\alpha < 1$ and $\beta < 1$, then $W(\cdot;\alpha,\beta,a)$ defined in (3) satisfies ICS. We consider four cases.

Case 1. We first suppose that $\alpha \neq 0$ and $\beta \neq 0$. In this case,

$$W(H;\alpha,\beta,a) = \Psi\left(\sum_{n=1}^{N} G_1(h_{n\cdot})\right), \text{ where } G_1(h_{n\cdot}) = \frac{1}{N}\mu_\beta^D(h_{n\cdot};a)^\alpha \text{ and } \Psi(x) = (x)^{\frac{1}{\alpha}}.$$

The first partial of G_1 is $\partial G_1(h_{n\cdot})/\partial h_{nd} = \frac{1}{N}\alpha a_d h_{nd}^{\beta-1} X^{\frac{\alpha}{\beta}-1}$ where $X = \sum_{d=1}^{D} a_d h_{nd}^{\beta}$, and the second partial is:

$$(G_1)_{dd} = \frac{1}{N} a_d \alpha (\beta - 1) h_{nd}^{\beta-2} X^{\alpha/\beta-1} + \frac{1}{N} a_d^2 \alpha (\alpha - \beta) h_{nd}^{2\beta-2} X^{\alpha/\beta-2}.$$

The second cross partial derivative of G_1 is:

$$(G_1)_{dd'} = \frac{1}{N}\alpha (\alpha - \beta) a_d a_{d'} h_{nd}^{\beta-1} h_{nd'}^{\beta-1} X^{\alpha/\beta-2}.$$

The Hessian matrix Q_1 of G_1 can be written as:

$$Q_1 = Q_1^1 + Q_1^2;$$

where Q_1^1 is a $D \times D$ diagonal matrix with the d^{th} diagonal element being equal to

$$Q_1^1(d,d) = \frac{1}{N} a_d \alpha (\beta - 1) h_{nd}^{\beta-2} X^{\frac{\alpha}{\beta}-1} \text{ for all } d = 1,\ldots,D,$$

and

$$Q_2^1 = \begin{bmatrix} \frac{1}{N} a_1^2 \alpha \left(\alpha - \beta\right) h_{n1}^{2\beta-2} X^{\alpha/\beta - 2} & \dots & \frac{\alpha(\alpha-\beta)}{N} a_1 a_D h_{n1}^{\beta-1} h_{nD}^{\beta-1} X^{\alpha/\beta-2} \\ \vdots & \ddots & \vdots \\ \frac{\alpha(\alpha-\beta)}{N} a_1 a_D h_{n1}^{\beta-1} h_{nD}^{\beta-1} X^{\alpha/\beta-2} & \dots & \frac{1}{N} a_D^2 \alpha \left(\alpha - \beta\right) h_{nD}^{2\beta-2} X^{\alpha/\beta-2} \end{bmatrix}.$$

Therefore, for a non-zero vector $z = (z_1, \dots, z_D) \in \mathbb{R}^D$,

$$zQ_1 z' = zQ_1^1 z' + zQ_1^2 z'; \tag{9}$$

where

$$zQ_1^1 z' = \sum_{d=1}^{D} z_d^2 \left(\frac{1}{N} a_d \alpha \left(\beta - 1\right) h_{nd}^{\beta-2} X^{\frac{\alpha}{\beta}-1} \right) = \frac{1}{N} \alpha \left(\beta - 1\right) X^{\frac{\alpha}{\beta}} \left[\sum_{d=1}^{D} \frac{a_d h_{nd}^{\beta}}{X} \left(\frac{z_d}{h_{nd}} \right)^2 \right],$$

and

$$zQ_1^2 z' = \frac{\alpha(\alpha-\beta) X^{\alpha/\beta}}{N} \left[z_1 \dots z_D \right] \begin{bmatrix} \left(\frac{a_1 h_{n1}^{\beta}}{X h_{n1}} \right)^2 & \dots & \frac{a_1 h_{n1}^{\beta}}{X h_{n1}} \frac{a_D h_{nD}^{\beta}}{X h_{nD}} \\ \vdots & \ddots & \vdots \\ \frac{a_1 h_{n1}^{\beta}}{X h_{n1}} \frac{a_D h_{nD}^{\beta}}{X h_{nD}} & \dots & \left(\frac{a_D h_{nD}^{\beta}}{X h_{nD}} \right)^2 \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_D \end{bmatrix}$$

or,

$$zQ_1^2 z' = \frac{1}{N} \alpha \left(\alpha - \beta\right) X^{\alpha/\beta} \left(\sum_{d=1}^{D} \frac{a_d h_{nd}^{\beta}}{X} \frac{z_d}{h_{nd}} \right)^2.$$

From (9), it follows that,

$$zQ_1 z' = \frac{1}{N} \alpha \left(\beta - 1\right) X^{\frac{\alpha}{\beta}} \left[\sum_{d=1}^{D} \frac{a_d h_{nd}^{\beta}}{X} \left(\frac{z_d}{h_{nd}} \right)^2 \right] + \frac{1}{N} \alpha \left(\alpha - \beta\right) X^{\alpha/\beta} \left(\sum_{d=1}^{D} \frac{a_d h_{nd}^{\beta}}{X} \frac{z_d}{h_{nd}} \right)^2.$$

Hence,

$$zQ_1z' = \alpha \frac{X^{\frac{\alpha}{\beta}}}{N} \left[(\beta - 1) \left(\sum_{d=1}^{D} \frac{X_d}{X} \left(\frac{z_d}{h_{nd}} \right)^2 \right) + (\alpha - \beta) \left(\sum_{d=1}^{D} \frac{X_d}{X} \frac{z_d}{h_{nd}} \right)^2 \right],$$

where $X_d = a_d h_{nd}^{\beta}$ and, as defined earlier, $X = \sum_{d=1}^{D} X_{nd}$. By Jensen's inequality, $\sum_{d=1}^{D}(X_d/X)(z_d/h_{nd})^2 \geq (\sum_{d=1}^{D}(X_d/X)(z_d/h_{nd}))^2$. As $\alpha < 1$ and $\beta < 1$, we have $(\beta - 1)\sum_{d=1}^{D}(X_d/X)(z_d/h_{nd})^2 + (\alpha - \beta)(\sum_{d=1}^{D}(X_d/X)(z_d/h_{nd}))^2 < 0$. There are two subcases: (i) $0 < \alpha < 1$ and (ii) $\alpha < 0$.

In subcase (i), $zQ_1z' < 0$. Hence, $G_1(\cdot)$ is strictly concave. Therefore, if H' is obtained from H by common smoothing, then by part (i) of Lemma B1, we have $\sum_{n=1}^{N} G_1(h'_{n\cdot}) > \sum_{n=1}^{N} G_1(h_{n\cdot})$. Because $\Psi(\cdot)$ is increasing for $\alpha > 0$, it follows that $W(\cdot; \alpha, \beta, a)$ satisfies ICS.

In subcase (ii), $zQ_1z' > 0$. Hence, $G_1(\cdot)$ is strictly convex. Part (ii) of Lemma B1 then implies that $\sum_{n=1}^{N} G_1(h'_{n\cdot}) < \sum_{n=1}^{N} G_1(h_{n\cdot})$ if H' is obtained from H by common smoothing. Because $\Psi(\cdot)$ is decreasing for $\alpha < 0$, $W(\cdot; \alpha, \beta, a)$ satisfies ICS in this subcase as well.

Case 2. We now suppose that $\alpha \neq 0$ and $\beta = 0$. In this case,

$$W(H; \alpha, \beta, a) = \Psi \left(\sum_{n=1}^{N} G_2(h_{n\cdot}) \right), \text{ where } G_2(h_{n\cdot}) = \frac{1}{N} \mu_0^D(h_{n\cdot}; a)^{\alpha} \text{ and } \Psi(x) = (x)^{\frac{1}{\alpha}}.$$

The first partial of G_2 is $\partial G_2(h_{n\cdot})/\partial h_{nd} = \frac{1}{N}\alpha a_d h_{nd}^{-1} Y$ where $Y = \prod_{d=1}^{D} h_{nd}^{\alpha a_d}$. The second partial of G_2 is $(G_2)_{dd} = \frac{1}{N}\alpha a_d(\alpha a_d - 1) h_{nd}^{-2} Y$ and the second cross partial is $(G_2)_{dd'} = \frac{1}{N}\alpha^2 a_d a_{d'} h_{nd}^{-1} h_{nd'}^{-1} Y$. Let the Hessian matrix be denoted by Q_2. For a non-zero vector $z =$

$(z_1, \ldots, z_D) \in \mathbb{R}^D$,

$$zQ_2z' = [z_1 \ldots z_D] \begin{bmatrix} \frac{\alpha a_1(\alpha a_1 - 1)}{N} h_{n1}^{-2} Y & \ldots & \frac{\alpha^2 a_1 a_D}{N} h_{n1}^{-1} h_{nD}^{-1} Y \\ \vdots & \ddots & \vdots \\ \frac{\alpha^2 a_1 a_D}{N} h_{n1}^{-1} h_{nD}^{-1} Y & \ldots & \frac{\alpha a_D(\alpha a_D - 1)}{N} h_{nD}^{-2} Y \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_D \end{bmatrix}$$

$$\text{or, } zQ_2z' = \frac{\alpha Y}{N} [z_1 \ldots z_D] \begin{bmatrix} -a_1 h_{n1}^{-2} + \alpha a_1^2 h_{n1}^{-2} & \ldots & \alpha a_1 a_D h_{n1}^{-1} h_{nD}^{-1} \\ \vdots & \ddots & \vdots \\ \alpha a_1 a_D h_1^{-1} h_D^{-1} & \ldots & -a_D h_D^{-2} + \alpha a_D^2 h_D^{-2} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_D \end{bmatrix}.$$

Hence,

$$zQ_2z' = \frac{\alpha Y}{N} \left[-\sum_{d=1}^{D} a_d \frac{z_d^2}{h_{nd}^2} + \alpha \left(\sum_{d=1}^{D} a_d \frac{z_d}{h_{nd}} \right)^2 \right].$$

By Jensen's inequality, $\sum_{d=1}^{D}(a_d z_d^2/h_{nd}^2) \geq (\sum_{d=1}^{D} a_d z_d/h_{nd})^2$. Because $\alpha < 1$, we have $-\sum_{d=1}^{D}(a_d z_d^2/h_{nd}^2) + \alpha(\sum_{d=1}^{D} a_d z_d/h_{nd})^2 < 0$. There are also two subcases: (i) $0 < \alpha < 1$ and (ii) $\alpha < 0$. Reasoning as in Case 1, it follows that $W(\cdot; \alpha, \beta, a)$ satisfies ICS.

Case 3. Next, we suppose that $\alpha = 0$ and $\beta \neq 0$. In this case,

$$W(H; \alpha, \beta, a) = \left(\prod_{n=1}^{N} G_3(h_{n\cdot}) \right)^{1/N}, \text{ where } G_3(h_{n\cdot}) = \mu_\beta^D(h_{n\cdot}; a).$$

Taking the logarithm on each side of this equation, it follows that $\ln[W(H; \alpha, \beta, a)] = \frac{1}{N}\sum_{n=1}^{N} \ln[G_3(h_{n\cdot})]$. Because G_3 is a generalized mean, both it and $\ln G_3$ are strictly concave for $\beta < 1$. By part (i) of Lemma B1, it follows that $W(\cdot; \alpha, \beta, a)$ satisfies ICS.

Case 4. *Case 4.* Finally, we suppose that $\alpha = 0$ and $\beta = 0$. Then,

$$W(H;\alpha,\beta,a) = \left(\prod_{n=1}^{N}\prod_{d=1}^{D} h_{nd}^{a_d}\right)^{1/N}.$$

Equivalently, we have $\ln[W(H;\alpha,\beta,a)] = \frac{1}{N}\sum_{n=1}^{N}\sum_{d=1}^{D} a_d \log h_{nd}$. Hence, from part (i) of Lemma B1 it follows that $W(\cdot;\alpha,\beta,a)$ satisfies ICS.

(b) Next, we establish necessity by showing that ICS is violated when either (i) $\alpha \geq 1$ or (ii) $\beta \geq 1$.

(i) Suppose that $\alpha \geq 1$. For any $N \in \mathbb{N}$, consider any $h \in \mathbb{R}^N_{++}$ and let $H \in \mathcal{H}_N$ be such that $h_{\cdot d} = h\ \forall d$. For any $a \in Int(S^{D-1})$ and any $\beta \in \mathbb{R}$, the overall achievement score vector associated with H is h. Thus, $W(H;\alpha,\beta,a) = \mu^N_\alpha(h;\xi_N)$. Consider any bistochastic matrix B and let $H' = BH$. By construction, $h'_{\cdot d} = h'\ \forall d$. The overall achievement score vector associated with H' is h', where $h' = Bh$. Hence, $W(H;\alpha,\beta,a) = \mu^N_\alpha(h;\xi_N) \geq \mu^N_\alpha(h';\xi_N) = W(H';\alpha,\beta,a)$ because $\alpha \geq 1$, violating ICS.

(ii) Suppose that $\beta \geq 1$. For any $a \in Int(S^{D-1})$, let $H \in \mathcal{H}_2$ be such that $h_{1\cdot} \neq h_{2\cdot}$ but $\mu^D_\beta(h_{1\cdot};a) = \mu^D_\beta(h_{2\cdot};a) =: \bar{x}$. Thus, $W(H;\alpha,\beta,a) = \bar{x}$. Let $H' = \bar{B}H$, where $\bar{B} = \frac{1}{2}\mathbf{1}_{22}$. It follows that $\mu^D_\beta(h'_{1\cdot};a) = \mu^D_\beta(h'_{2\cdot};a) =: \bar{y}$ and $W(H';\alpha,\beta,a) = \bar{y}$. Because μ^D_β is strictly convex for $\beta > 1$, by part (ii) of Lemma B1, we have $\mu^D_\beta(h_{1\cdot};a) + \mu^D_\beta(h_{2\cdot};a) = 2\bar{x} > \mu^D_\beta(h'_{1\cdot};a) + \mu^D_\beta(h'_{2\cdot};a) = 2\bar{y}$. This implies that $W(H;\alpha,\beta,a) > W(H';\alpha,\beta,a)$. Furthermore, $W(H;\alpha,1,a) = W(H';\alpha,1,a)$ because, by construction, $\bar{x} = \bar{y}$ when $\beta = 1$. Hence, ICS is violated for any $\beta \geq 1$. ■

C. Proof of Theorem 3

For the purpose of the proof, for every $\mathbf{N} \subset \mathbb{N}$ and every $H \in \mathcal{H}_N$, we express $W(H; \alpha, \beta, a)$ as

$$W(H; \alpha, \beta, a) = \mathcal{F}(F(G(h_{1\cdot}), \dots, G(h_{N\cdot}))), \tag{10}$$

where $G : \mathbb{R}^D_{++} \rightarrow \mathbb{R}_{++}$, $F : \mathbb{R}^N_{++} \rightarrow \mathbb{R}_{++}$, and $\mathcal{F} : \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$. The functional forms of G, F, and $\mathcal{F}$ are conditional on α and β, as shown in Table 1.

Table 1: Functional Forms of G, F, and $\mathcal{F}$

	$\mathcal{F}(\cdot)$	$F(\cdot)$	$G(h_{n\cdot})$
$\alpha \neq 0,\ \beta \neq 0:$	$\left(\frac{1}{N}F(\cdot)\right)^{1/\alpha}$	$\sum_{n=1}^N G(\cdot)$	$(\mu^D_\beta(h_{n\cdot}; a))^\alpha$
$\alpha \neq 0,\ \beta = 0:$	$\left(\frac{1}{N}F(\cdot)\right)^{1/\alpha}$	$\sum_{n=1}^N G(\cdot)$	$(\mu^D_0(h_{n\cdot}; a))^\alpha$
$\alpha = 0,\ \beta \neq 0:$	$(F(\cdot))^{1/N}$	$\prod_{n=1}^N G(\cdot)$	$\mu^D_\beta(h_{n\cdot}; a)$
$\alpha = 0,\ \beta = 0:$	$(F(\cdot))^{1/N}$	$\prod_{n=1}^N G(\cdot)$	$\mu^D_0(h_{n\cdot}; a)$

To determine how W changes in response to a sequence of association increasing transfers, we first need to determine how F responds to such a sequence, which in turn depends on whether G is strictly L-subadditive, strictly L-superadditive, or a valuation.[13]

From Table 1, we see that F is either additive or multiplicative. Lemmas C1 and C2 summarize how F is sensitive to a sequence of association increasing transfers when F is additive or multiplicative, respectively.

Lemma C1 *For every* $\mathbf{N} \subset \mathbb{N}\backslash\{1\}$*, every* $H', H \in \mathcal{H}_N$ *such that* H' *is obtained from* H *by a finite sequence of association increasing transfers, and for* $F(H) = \sum_{n=1}^N G(h_{n\cdot})$, (i) $F(H') <$

[13] A twice differentiable function $G : R^D_{++} \rightarrow R_+$, is (i) *strictly L-subadditive* if $\partial^2 G(h_{n\cdot})/\partial h_{nd_1}\partial h_{nd_2} < 0\ \forall d_1 \neq d_2$; (ii) *strictly L-superadditive* if $\partial^2 G(h_{n\cdot})/\partial h_{nd_1}\partial h_{nd_2} > 0\ \forall d_1 \neq d_2$; and (iii) *a valuation* if $\partial^2 G(h_{n\cdot})/\partial h_{nd_1}\partial h_{nd_2} = 0\ \forall d_1 \neq d_2$. See Milgrom and Roberts (1990) or Topkis (1998, p. 43).

Table 2: Modularity Properties of G

Strict L-subadditive	**Strict L-superadditive**	**Valuation**
$\alpha > 0$ and $\alpha < \beta$	$\alpha < 0$, $\alpha < \beta$, and $\beta \neq 0$	$\alpha = \beta \neq 0$
$\alpha = 0$ and $\beta > 0$	$\alpha > 0$, $\alpha > \beta$ and $\beta \neq 0$	$\alpha = \beta = 0$
$\alpha < 0$ and $\alpha > \beta$	$\alpha < 0$ and $\beta = 0$	
	$\alpha > 0$ and $\beta = 0$	
	$\alpha = 0$ and $\beta < 0$	

$F(H)$ if and only if G is strictly L-subadditive, (ii) *$F(H') > F(H)$ if and only if G is strictly L-superadditive, and* (iii) *$F(H) = F(H)$ if and only if G is a valuation.*

Proof. See Boland and Proschan (1988, Proposition 2.5 (a)). ■

Lemma C2 *For every $\mathbf{N} \subset \mathbb{N} \backslash \{1\}$, every $H, H' \in \mathcal{H}_N$ such that H' is obtained from H by a finite sequence of association increasing transfers, and for $F(H) = \prod_{n=1}^{N} G(h_{n\cdot})$,* (i) *$F(H') < F(H)$ if and only if $\ln G$ is strictly L-subadditive,* (ii) *$F(H') > F(H)$ if and only if $\ln G$ is strictly L-superadditive, and* (iii) *$F(H) = F(H')$ if and only if $\ln G$ is a valuation.*

Proof. This result immediately follows from Lemma C1 by taking a logarithm on each side of $F(H) = \sum_{n=1}^{N} G(h_{n\cdot})$. ■

Table 2 summarizes the restrictions on α and β under which G, and hence $\ln G$, is strictly L-subadditive, strictly L-superadditive, or a valuation. With these preliminaries in hand, we now prove Theorem 3.

Proof of Theorem 3. For any N, let H' be obtained from $H \in \mathcal{H}_N$ by a sequence of association increasing transfers. We separately consider the cases in which $\alpha < \beta$, $\alpha > \beta$, and $\alpha = \beta$.

First, we show that if $\alpha < \beta$, then the social welfare index W satisfies DIA. There are four cases to consider: (i) $\alpha > 0$ and $\alpha < \beta$, (ii) $\alpha < 0$, $\alpha < \beta$, and $\alpha \neq \beta$, (iii) $\alpha < 0$ and $\beta = 0$, and (iv) $\alpha = 0$ and $\beta > 0$. In cases (i), (ii), and (iii), $F(\cdot) = \sum_{n=1}^{N} G(\cdot)$. In case (i), by Table 2, G is strictly L-subadditive and, hence, by Lemma C1, $F(H') < F(H)$. Because $\mathcal{F}(\cdot) = (\frac{1}{N}F(\cdot))^{1/\alpha}$ and $\alpha > 0$, it follows that $W(H'; \alpha, \beta, a) < W(H; \alpha, \beta, a)$. In case (ii), by Table 2, G is strictly L-superadditive and, hence, by Lemma C1, $F(H') > F(H)$. Because $\mathcal{F}(\cdot) = (\frac{1}{N}F(\cdot))^{1/\alpha}$ and $\alpha < 0$, it follows that $W(H'; \alpha, \beta, a) < W(H; \alpha, \beta, a)$. In case (iii), by Table 2, $\ln G$ is strictly L-superadditive and, hence, by Lemma C2, $F(H') > F(H)$. Because $\mathcal{F}(\cdot) = (\frac{1}{N}F(\cdot))^{1/\alpha}$ and $\alpha < 0$, it follows that $W(H'; \alpha, \beta, a) < W(H; \alpha, \beta, a)$. In case (iv), $F(\cdot) = \prod_{n=1}^{N} G(\cdot)$. By Table 2, G is strictly L-subadditive and, hence, by Lemma C1, $F(H') < F(H)$. Because $\mathcal{F}(\cdot) = (F(\cdot))^{1/N}$, $W(H'; \alpha, \beta, a) < W(H; \alpha, \beta, a)$. Therefore, W satisfies DIA if $\alpha < \beta$.

Next, we show that if $\alpha > \beta$, then the social welfare index W satisfies IIA. Again, there are four cases to consider: (i) $\alpha < 0$ and $\alpha > \beta$, (ii) $\alpha > 0$, $\alpha > \beta$, and $\beta \neq 0$, (iii) $\alpha > 0$ and $\beta = 0$, and (iv) $\alpha = 0$ and $\beta < 0$. In cases (i), (ii), and (iii), $F(\cdot) = \sum_{n=1}^{N} G(\cdot)$. In case (i), by Table 2, G is strictly L-subadditive and, hence, by Lemma C1, $F(H') < F(H)$. Because $\mathcal{F}(\cdot) = (\frac{1}{N}F(\cdot))^{1/\alpha}$ and $\alpha < 0$, it follows that $W(H'; \alpha, \beta, a) > W(H; \alpha, \beta, a)$. In case (ii), by Table 2, G is strictly L-superadditive and, hence, by Lemma C1, $F(H') > F(H)$. Because $\mathcal{F}(\cdot) = (\frac{1}{N}F(\cdot))^{1/\alpha}$ and $\alpha > 0$, it follows that $W(H'; \alpha, \beta, a) > W(H; \alpha, \beta, a)$. In case (iii), by Table 2, $\ln G$ is strictly L-superadditive and, hence, by Lemma C2, $F(H') > F(H)$. Because $\mathcal{F}(\cdot) = (\frac{1}{N}F(\cdot))^{1/\alpha}$ and $\alpha > 0$, it follows that $W(H'; \alpha, \beta, a) > W(H; \alpha, \beta, a)$. In case (iv), $F(\cdot) = \prod_{n=1}^{N} G(\cdot)$. By Table 2, G is strictly L-superadditive and, hence, by Lemma C1, $F(H') > F(H)$. Because $\mathcal{F}(\cdot) = (F(\cdot))^{1/N}$, $W(H'; \alpha, \beta, a) > W(H; \alpha, \beta, a)$. Therefore, W

satisfies IIA if $\alpha > \beta$.

It remains to be shown that if $\alpha = \beta$, then W satisfies neither DIA nor IIA. If $\alpha = \beta \neq 0$ (resp. $\alpha = \beta = 0$), then by Table 2, G (resp. $\ln G$) is a valuation. Thus, $F(H') = F(H)$ by Lemma C1 (resp. Lemma C2). It then follows that $W(H'; \alpha, \beta, a) = W(H; \alpha, \beta, a)$. Hence, W satisfies neither DIA nor IIA. ■

Chapter III

Rank Robustness of Composite Indices: Dominance and Ambiguity (with James Foster and Mark McGillivray)

Introduction

Composite indices are frequently used by economists and other social scientists to assess the performance of a society when the assessment is based on achievements in more than one dimension. In this chapter, a society may range from a country or a state to an academic department. Recent decades have seen increased use of composite indices, which, by their very nature, combine in various ways *indicators* of achievement in various dimensions. Remarkable attention is given to rankings arising from these indices and this is especially true of country rankings. People are naturally curious as to how their country compares to others, national pride is often at stake, and national governments are often quick to claim credit for a high or higher than expected ranking if it can be linked, dubiously or otherwise, to public policy. More generally, the media, business groups, civil society, sections of the research community, and international organisations regularly monitor and report on country rankings of indices assessing a variety of phenomena such as sustainability, corruption, rule of law, national income, economic policy efficacy, institutional performance, happiness, human well-being, transparency, globalisation, human freedom, peace or vulnerability. Besides country rankings, the other type of ranking that is often of high interest is the ranking of the US graduate school departments by their academic effectiveness. It is widely recognised that most of the preceding phenomena are multidimensional in nature.

The interest of national governments and others in rankings arising from composite indices is, however, blind to long held concerns regarding their construction. A central concern is the *weighting* of indicators. In a perfect world, the weight vectors would be based on the information obtained from a meta production function for the phenomenon in question. An absence of accepted information on these functions has resulted in one of three weighting schemes. The first and the most common is to select weights arbitrarily, typically by taking the simple arithmetic mean of the indicators in question. Using this mean is interpreted as assigning equal weights to each dimension. The proponents of this equal weight approach acknowledge that the approach is deficient, as in reality the dimensions will almost certainly have differential importance, but argue that there is no accepted basis or guidance for doing otherwise. In this sense, the equal weight approach is seen as the least deficient available weighing scheme, one that is likely to attract the least disagreement.[14] The second is the normative approach that involves setting weights either in accordance with individual or societal norms, with the individuals often being those of the designers of the index in question. The third scheme is statistical, being purely data-driven. Many different such approaches have been proposed, the most popular being the principal components analysis, with the first principal component extracted from the dimensional indicators serving as the composite index. A weight vector arising from both the second and the third approaches is also

[14] For example, the proponents of the Environmental Sustainability Index (ESI) argued for equal weights on the grounds that "... no objective mechanism exists to determine the relative importance of the different aspects of environmental sustainability" (Esty et al., 2005, p. 66). Mayer and Jencks (1989) argue that "...we have no reliable basis for doing [otherwise]". Other composite indices, used in environmental, well-being and related fields that employ equal weights include the Child Well-being Index, Commitment to Development Index, Economic Resilience Index, Economic Vulnerability Index, Environmental Performance Index, Environmental Sustainability Index, Gender Empowerment Measure, Gender-related Development Index, Genuine Progress Measure, Global Peace Index, Human Development Index, Human Poverty Index, Index of Economic Freedom, Global Peace Index and the Physical Quality of Life Index. In most of the above cases, the index is formed by taking the simple arithmetic mean of the component indicators.

difficult to defend: the second because of a lack of guidance as to whose norms should be used and the third because of difficulties in interpretation.

No matter whichever of the above three schemes is applied to determine the weights, it is hard to have universal agreement over the choice. A selection of different weighting vectors other than the one initially selected often alters the ranking and thus creates an ambiguity in the comparison, which in turn may result in the policy recommendation based on these indices being indecisive. One may thus cast doubt on the ranking arising from these indices. Specifically, one can ask to what extent these rankings are conditional upon the initial weighting vector.

Such is the focus of this chapter. An ordering based merely on the initially selected weight vector is complete and one is always able to compare any two geographical units unambiguously, but as discussed above, the chosen weigh vector itself is subject to wide disagreement. We, on the other hand, use a framework based on a strict partial ordering instead. We view a comparison as completely robust or unambiguous if rankings are not reversed using any weight vector within a given set. The comparison is incomplete or is stated to be ambiguous, otherwise. Our strict partial ordering framework is analogous to that of Sen (1970) and Foster and Shorrocks (1988a,b). It is also closely related to the model of Knightian uncertainty (Bewley, 2002) and the multiple prior model of Gilboa and Schmeidler (1989). The condition that one can rank any two units only if the ranking is not reversed using any weight vector within a given set is analogous to the situation where one places zero confidence on her initial choice. On the other hand, if one bases the rankings only on the initial choice so as to obtain a complete ordering, then this is analogous to the situation where one is absolutely confident about her selection. Both situations are, under natural circumstances, too stringent. We thus pursue an intermediate approach, where one is partially confident about the choice of initial weight vector. In this framework, one believes that her initial choice is correct with some probability only, but with the rest of the probability, she is uncertain about what should be the correct choice. This framework is similar to that of the multiple prior model of epsilon-contamination applied in Bayesian Statistics and Decision Theory, where an agent is only certain with some probability about her decision, but she is worried that, with some chance, her decision may be completely wrong. The two extreme situations of zero confidence and absolute confidence are special cases of our framework.

We, further, propose a measure by which the robustness of a given comparison may be gauged and illustrate its usefulness using data from the Human Development Index (HDI). The HDI is a very well known and widely used measure of well-being at the national level and the rankings it provides are the subject of intense international interest. This chapter shows how some country rankings are fully robust to changes in weights, while others are quite fragile. It further investigates the prevalence of the different levels of robustness in theory and practice. From the outset it should be emphasised that the fundamental purpose of this chapter is not to discourage the reporting or use of these indices and the rankings they provide. Rather, it is to facilitate more incisive interpretation of these rankings.

It is shown later that the approach we develop in this chapter can easily be extended to situations where composite indices are constructed by taking generalized means of indicators. Also, the approach is applicable to situations where composite indices are constructed by taking the average of the dimension-specific ranks.

The remainder of chapter is structured as follows. The second section provides a description of the mathematical concepts, notation, and definitions used throughout the chapter. A formal treatment of the notion of dominance and its relation to rank robustness is provided in the third section, which also defines and characterizes a strict partial ordering that facilitates the construction of a measure of robustness. The fourth section constructs a rank robustness measure and provides an application of inter-country comparisons in terms of the HDI. The fifth section looks at the prevalence of robust comparisons, highlighting how the number of ambiguous comparisons across an entire sample of observations varies with the critical level of the measure of robustness. The HDI is used in this section to illustrate key points. The final section concludes this chapter.

Notation and Definitions

Let $D \geq 2$ be the number of dimensions under consideration. For $a, b \in \mathbb{R}^D$, the expression $a \geq b$ indicates that $a_d \geq b_d$ for $d = 1, \ldots, D$; this is the vector dominance relation. If $a \geq b$ with $a \neq b$, this situation is denoted by $a > b$; while $a >> b$ indicates that $a_d > b_d$ for $d = 1, \ldots, D$. Let $X \subseteq \mathbb{R}^D$ denote the nonempty set of indicator vectors and let S^D $= \{s \in \mathbb{R}^{D+1} : s \geq 0 \text{ and } \sum_{d=1}^{D+1} = 1\}$ be the simplex of associated weighting vectors. A *composite index* $C : X \times S^{D-1} \rightarrow \mathbb{R}$ combines the dimensional indicators in $x \in X$ using a weighting vector $w \in S^{D-1}$ to obtain an aggregate level $C(x; w) = \sum_{d=1}^{D} w_d x_d$. In what follows, it is assumed that an *initial weighting vector* $w^0 \in S^{D-1}$ satisfying $w^0 >> 0$ has already been chosen; this fixes the specific composite index $C_0 : X \rightarrow \mathbb{R}$ defined as $C_0(x) =$ $C(x; w^0)$ for all $x \in X$. The associated strict ordering of indicator vectors will be denoted by $\mathbf{C}_0$, so that $x \; \mathbf{C}_0 \; y$ holds if and only if $C_0(x) > C_0(y)$. For every $d \in \{1, \ldots, D\}$, we denote the D-dimensional basis vector by v_d, whose d^{th} element is equal to one and the rest of the elements are zero. For example, $v_1 = (1, 0, \ldots, 0)$, $v_2 = (0, 1, 0, \ldots, 0)$, and $v_D = (0, 0, \ldots, 0, 1)$.

Robust Comparisons

Our aim is to construct a general criterion for determining when a given comparison x $\mathbf{C}_0$ y is robust. The motivation is similar in spirit to the use of the Lorenz criterion as a robustness check in inequality evaluations, or stochastic dominance tests for comparisons involving risk or poverty. Let $W \in S^{D-1}$ be a nonempty set of weighting vectors. Define the weak robustness relation $\mathbf{R}_W$ on X by x $\mathbf{R}_W$ y if and only if $C(x, w) \geq C(y, w)$ for all $w \in W$. If both x $\mathbf{C}_0$ y and x $\mathbf{R}_W$ y hold for $w^0 \in W$, then we say that x robustly dominates y (given w^0 and W), and denote this by x $\mathbf{C}_W$ y. In words, the level of the composite index is higher for x than y at w^0, and this ranking is not reversed using any other weighting vector in W. If instead x $\mathbf{C}_0$ y holds, but x $\mathbf{R}_W$ y does not, then this indicates that the ranking $C(x, w^0) > C(y, w^0)$ is not robust (relative to the given W) since the initial inequality is reversed using another weighting vector, say, $C(x, w^1) < C(y, w^1)$ for $w^1 \in W$.

The relations $\mathbf{R}_W$ and $\mathbf{C}_W$ are closely linked with other dominance criteria, including Sen's (1970) approach to partial comparability in social choice and Bewley's (2002) multiple prior model of Knightian uncertainty. Bewley's presentation, in particular, suggests a natural characterization of $\mathbf{R}_W$ among all binary relations $\mathbf{R}$ on X. Consider the following properties of a binary relation $\mathbf{R}$ on X, each of which is satisfied by $\mathbf{R}_W$.

Quasiordering (Q) $\mathbf{R}$ is transitive and reflexive.

Monotonicity (M) (i) If $x > y$ then x $\mathbf{R}$ y; (ii) if $x >> y$ then y $\mathbf{R}$ x cannot hold.

Independence (I) Let $x, y, z, y', z' \in X$ where $y' = \alpha x + (1-\alpha)y$ and $z' = \alpha x + (1-\alpha)z$ for $0 < \alpha < 1$. Then y $\mathbf{R}$ z if and only if y' $\mathbf{R}$ z'.

Continuity (C) The set $\mathbf{R}(z) = \{x \in X \mid x \mathbf{R} z\}$ is closed for all $z \in X$.

Axiom Q allows $\mathbf{R}$ to be incomplete. Axiom M ensures that $\mathbf{R}$ follows vector dominance when it applies, and rules out the converse ranking when vector dominance is strict. Axiom I is a standard independence axiom, which requires the ranking between y and z to be consistent with the ranking of y' and z' obtained by a convex combination with another vector x. Finally, Axiom C ensures that the upper contour sets of $\mathbf{R}$ contain all their limit points. We have the following characterization.

Theorem 6 *Suppose that X is closed, convex, and has a nonempty interior. Then a binary relation $\mathbf{R}$ on X satisfies axioms Q, M, I, and C if and only if there exist a non-empty, closed, and convex set $W \subseteq S^{D-1}$ such that $\mathbf{R} = \mathbf{R}_W$.*

Proof. See Appendix D. ■

Thus any robustness relation satisfying the four axioms is generated by pairwise comparisons of the composite index over some fixed set W of weighting vectors.

The ranking $\mathbf{R}_W$ has an interesting interpretation in terms of the well-known maxmin criterion of Gilboa and Schmeidler (1989) for multiple priors. Suppose we know that x $\mathbf{R}_W$ y for some nonempty, closed set $W \subseteq S^{D-1}$. By linearity of the composite index, this can be expressed as $C(x-y, w) \geq 0$ for all $w \in W$, or as $\min_{w \in W} C(x-y, w) \geq 0$. The Gilboa-Schmeidler evaluation function $G_W(z) = \min_{w \in W} C(z, w)$ represents the maxmin criterion, which ranks a pair of options x and y by comparing $G_W(x)$ and $G_W(y)$, or the respective minimum values of the composite index on the set W. Our robustness ranking $x\ \mathbf{R}_W\ y$ is obtained by applying G_W to the net vector $(x - y)$ and checking whether the resulting value is nonnegative. Indeed, $x\ \mathbf{R}_W\ y$ if and only if $G_W(x - y) \geq 0$.[15]

Theorem 6 shows that under the given axioms, the selection of a robustness criterion reduces to the choice of an appropriate set W of multiple weighting vectors used in $\mathbf{R}_W$. But which W should be used? As we argue below, the answer depends in part on the confidence one places in the initial weighting vector w^0. If one has confidence that w^0 is the most appropriate weighting vector, then this would be reflected in the selection of a smaller set W containing w^0. The limiting case of $W = \{w^0\}$ indicates utmost confidence in w^0 and hence entails no robustness test at all: $x\ \mathbf{C}_0\ y$ is equivalent to $x\ \mathbf{C}_W\ y$. On the other hand, a larger W would suggest less confidence in w^0, a more demanding robustness test $\mathbf{R}_W$, and correspondingly fewer robust comparisons according to $\mathbf{C}_W$. Clearly $\mathbf{C}_{W'}$ is a subrelation of $\mathbf{C}_W$ whenever $W \subseteq W'$. We now investigate the robustness relations for some natural specifications of the set W of allowable weighting vectors.[16]

Full Robustness

We begin with the limiting case where W is the set S^{D-1} of all possible weighting vectors, and denote the associated robustness relations by $\mathbf{R}_1$ and $\mathbf{C}_1$. When $x\ \mathbf{C}_1\ y$ holds we say the comparison $x\ \mathbf{C}_0\ y$ is fully robust since it is never reversed at any configuration of weights. Of course, requiring unanimity over all of S^{D-1} is quite demanding and consequently $\mathbf{C}_1$ is

---

[15] The maxmin criterion applies when $G_W(x) - G_W(y) \geq 0$, while our robustness criterion holds when $G_W(x - y) \geq 0$. The maxmin criterion generates a complete relation, but requires comparisons of $C(x, w)$ with $C(y, w')$ for some $w \neq w'$, which is not easily interpreted in the present context. Note that the maxmin criterion is implied by our robustness criterion

[16] Since $\mathbf{C}_W$ is the intersection of $\mathbf{C}_0$ and $\mathbf{R}_W$, it is a strict partial order (transitive and irreflexive) satisfying conditions I and M.

the least complete among all such relations; however, when it applies the associated ranking of indicator vectors is maximally robust.

Consider the vertices of S^{D-1}, given by $v_d = e_d$ for $d = 1, \ldots, D$, where e_d is the usual basis element that places full weight on the single indicator d. Clearly $C(x, v_d) = x_d$, which suggests a link between the robustness relations and vector dominance. Indeed, we have the following characterizations of $\mathbf{R}_1$ and $\mathbf{C}_1$.

Theorem 7 *Let $x, y \in X$. Then* (i) *x $\mathbf{R}_1$ y if and only if $x \geq y$ and* (ii) *x $\mathbf{C}_1$ y if and only if $x > y$.*

Proof. Suppose that x $\mathbf{C}_0$ y is true. If $x \geq y$ holds, then clearly $C(x; w) = w \cdot x \geq w \cdot y = C(y; w)$ for all $w \in S^{D-1}$, and thus x $\mathbf{C}_1$ y. Conversely, if x $\mathbf{C}_1$ y holds, then setting $w = v_d$ in $C(x; w) \geq C(y; w)$ yields $x_d \geq y_d$ for all d, and hence $x \geq y$. ■

In order to check whether a given ranking x $\mathbf{C}_0$ y is fully robust, one need only verify that the indicators in x are at least as high as the respective levels in y.

One interesting implication of Theorem 7 is that judgments made by $\mathbf{C}_1$ are "meaningful" even when variables are ordinal and no basis of comparison between them has been fixed.[17] Suppose that each variable x_d in x is independently altered by its own monotonically increasing transformation $f_d(x_d)$ and let the resulting transformed indicator vector be $x' = (f_1(x_1), \ldots, f_D(x_D))$.[18] It is clear that $x > y$ if and only if $x' > y'$, and consequently, by Theorem 7 we have x $\mathbf{C}_1$ y if and only if x' $\mathbf{C}_1$ y'. In other words, if $\mathbf{C}_1$ holds for any given cardinalization of the ordinal variables, it holds for all cardinalizations. Note that while $\mathbf{C}_0$

[17] For a technical discussion of "meaningful statements" using a measurement theory approach, see Roberts (1979).

[18] The resulting function $f : X \to \mathbb{R}^D$ defined by $f(x) = (f_1(x_1), \ldots, f_D(x_D))$ is called a *monotonically increasing transformation.*

on its own is not meaningful in this context (as y' $\mathbf{C}_0$ x' is entirely consistent with x $\mathbf{C}_0$ y), the fully robust relation $\mathbf{C}_1$ is preserved and hence is appropriate for use with ordinal variables.

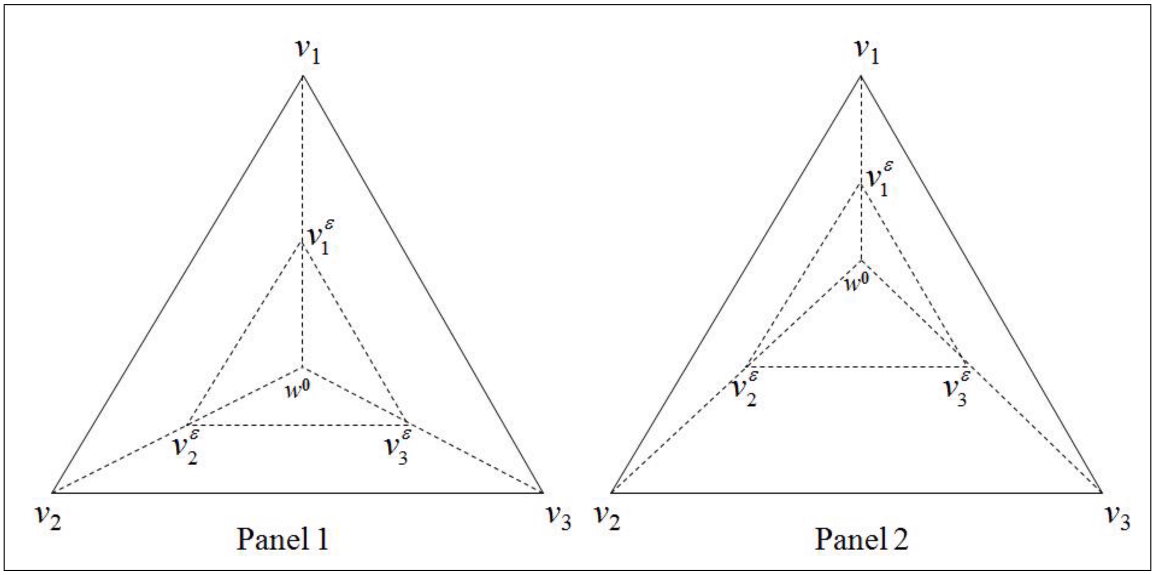
Epsilon Robustness

Now consider $S_\varepsilon^{D-1} \subseteq S^{D-1}$ defined by $S_\varepsilon^{D-1} = (1-\varepsilon)\{w^0\} + \varepsilon S^{D-1}$ for $0 \leq \varepsilon \leq 1$, which is made up of vectors of the form $(1-\varepsilon)w^0 + \varepsilon w$, where $w \in S^{D-1}$. Parameter value $\varepsilon = 0$ yields $S_0^{D-1} = \{w^0\}$ and hence the "no robustness" case, while $\varepsilon = 1$ yields $S_1^{D-1} = S^{D-1}$ or full robustness. Each S_ε^{D-1} with $0 < \varepsilon < 1$ is a scaled down version of S^{D-1} located so that w^0 is in the same relative position in S_ε^{D-1} as it is in S^{D-1}. Figure 1 provides examples of S_ε^{D-1} for the case of $D = 3$ and $\varepsilon = 1/4$, where Panel 1 has $w^0 = (1/3, 1/3, 1/3)$ and Panel 2 has $w^0 = (3/5, 1/5, 1/5)$. As noted in the Figure, ε is a measure of the relative size of S_ε^{D-1}. Moreover, for a given w^0, the sets are nested in such a way that $S_\varepsilon^{D-1} \subset S_{\varepsilon'}^{D-1}$ whenever ε' $> \varepsilon$.

The set S_ε^{D-1} of weighting vectors can be motivated using the well-known epsilon contamination model of multiple priors commonly applied in statistics and decision theory.[19] In that context, w^0 corresponds to an initial subjective distribution and S_ε^{D-1} contains probability distributions that are convex combinations of w^0 and the set of all objectively possible distributions, where $(1-\varepsilon)$ represents the decision maker's level of confidence in w^0 and ε is the extent of the "perturbation" from w^0. The Gilboa-Schmeidler evaluation function G_W then reduces to a form invoked by Ellsberg (1961), namely $G_\varepsilon(z) = (1-\varepsilon)C(z, w^0) + \varepsilon \min_{w \varepsilon S^{D-1}} C(z, w)$ using our notation.

[19] See for example, Carlier, Dana, and Shahidi (2003); Chateauneuf, Eichberger, and Grant (2006); Nishimura and Ozaki (2006); Carlier and Dana (2008); Asano (2008); and Kopylov (2009).

Figure 1: Multiple Weighting Vectors: The epsilon-Robustness Set



Substituting S_ε^{D-1} in the definitions of $\mathbf{R}_W$ and $\mathbf{C}_W$ yields the *ε-robustness relations* $\mathbf{R}_\varepsilon$ and $\mathbf{C}_\varepsilon$. Since the sets S_ε^{D-1} are nested for a given w^0, it follows that $x\ \mathbf{C}_\varepsilon\ y$ implies $x\ \mathbf{C}_{\varepsilon'}\ y$ whenever $\varepsilon > \varepsilon'$. The rankings clearly require $C(x,w) \geq C(y,w)$ for all w in S_ε^{D-1} and hence at each of its vertices $v_d^\varepsilon = (1-\varepsilon)w^0 + \varepsilon v_d$. Define $x^\varepsilon = (x_1^\varepsilon, \ldots, x_D^\varepsilon)$ where $x_d^\varepsilon = C(x, v_d^\varepsilon) = v_d^\varepsilon \cdot x$, and let y^ε be the analogous vector derived from y. The following result characterizes $\mathbf{R}_\varepsilon$ and $\mathbf{C}_\varepsilon$.

Theorem 8 *Let $x, y \in X$. Then* (i) *$x\ \mathbf{R}_\varepsilon\ y$ if and only if $x^\varepsilon \geq y^\varepsilon$ and* (ii) *$x\ \mathbf{C}_\varepsilon\ y$ if and only if $x^\varepsilon > y^\varepsilon$.*

Proof. We need only verify that $x\ \mathbf{C}_0\ y$ and $x^\varepsilon \geq y^\varepsilon$ imply $x\ \mathbf{C}_\varepsilon\ y$. Pick any $w \in S_\varepsilon^{D-1}$, and note that since S_ε^{D-1} is the convex hull of its vertices, w can be expressed as a convex combination of $v_1^\varepsilon, \ldots, v_D^\varepsilon$, say $w = \alpha_1 v_1^\varepsilon + \ldots + \alpha_D v_D^\varepsilon$ where $\alpha_1 + \cdots + \alpha_D = 1$ and $\alpha_d \geq 0$ for $d = 1, \ldots, D$. But then $C(x;w) = w \cdot x = \alpha_1 v_1^\varepsilon \cdot x + \cdots + \alpha_D v_D^\varepsilon \cdot x = \alpha_1 x_1^\varepsilon + \cdots + \alpha_D x_D^\varepsilon$, and similarly $C(y; w) = \alpha_1 y_1^\varepsilon + \cdots + \alpha_D y_D^\varepsilon$; therefore $x^\varepsilon \geq y^\varepsilon$ implies $C(x; w) \geq C(y; w)$. Since w was an arbitrary element of S_ε^{D-1}, it follows that $x \; C_\varepsilon \; y$. ■

Theorem 8 shows that to evaluate whether a given comparison $x \; \mathbf{C}_0 \; y$ is ε-robust, one need only compare the associated vectors x^ε and y^ε. If each component of x^ε is at least as large as the respective component of y^ε, then the comparison is ε-robust; if any component is larger for y^ε than x^ε, then the comparison is not. Checking whether the x^ε vector dominates y^ε is equivalent to requiring the inequality $C(x, w) \geq C(y, w)$ to hold for each vertex $w = v_d^\varepsilon$ of the set S_ε^{D-1}. Note further that x^ε is a convex combination of the vectors $(C_0(x), \ldots, C_0(x))$ and x, namely, $x^\varepsilon = (1 - \varepsilon)(C_0(x), \ldots, C_0(x)) + \varepsilon x$, so that when $\varepsilon = 1$ we obtain the condition $x \geq y$ in Theorem 7, while when $\varepsilon = 0$, the condition reduces to a simple comparison of $C_0(x)$ and $C_0(y)$.

Our approach differs from the existing approach for robustness testing proposed by Cherchye, Ooghe, and Puyenbroeck (2008) in various ways.[20] First, the later approach is applicable to only a particular type of scaling, where indicators are divided by the dimension specific medians. Our approach, however, does not assume any particular form of scaling. Second, Cherchye et al. (2008) assume that the weights on each dimension depend on the dispersion of that dimension. Thus, unlike in our approach, the maximum and minimum possible weights on each dimension are not always one and zero, respectively. Furthermore, a dimension with smaller dispersion has a relatively smaller weight variation compared to a dimension with larger dispersion. Our approach, on the other hand, assumes the variation

[20] There is another literatures that uses sensitivity analysis to verify the strength of comparisons. Sensitivity analysis is different from robustness testing in the sense that it estimates confidence intervals around each composite index depending on different scenarios. If the confidence intervals of two composite indicators do not intersect, an unambiguous comparison is possible. See for example Saisana et al. (2005) and Cherchye et al. (2008).

in weights to be independent of the dispersion of dimensions; a weight equal to one implies that the dimension is the only one that matters and a weight equal to zero implies that the dimension does not matter at all. The third difference between these two approaches is that Cherchye et al. (2008) assume that one region is better than a second region if and only if the former generalized Lorenz dominates the latter. They assume that the different dimensions are interchangeable. Our approach, on the contrary, is interested in dimension specific comparisons. To us, we compare income with income, or education with education, and not income with education.

Before we conclude the section, we should mention that the concept of full robustness and epsilon robustness can easily be extended to the situation where composite indices are constructed as a general mean of dimension-specific indices. The composite index, in this situation, can be defined as $\mathcal{C}(x; w) = (w \cdot x^{\alpha})^{1/\alpha}$, where x^{α} represents the vector with each element of x raised to the power $\alpha \in \mathbb{R}$. The strict ordering of indicator vectors requires x $\mathbf{C}_0$ y holds if and only if $\mathcal{C}_0(x) > \mathcal{C}_0(y)$, or, $(w \cdot x^{\alpha})^{1/\alpha} > (w \cdot y^{\alpha})^{1/\alpha}$. This requires comparing only $w \cdot x^{\alpha}$ and $w \cdot y^{\alpha}$; the comparison is linear in w.[21] Analogous to Theorem 7 and Theorem 8, it can be easily shown that (i) x $\mathbf{C}_1$ y if and only if $x > y$ and (ii) x $\mathbf{C}_{\varepsilon}$ y if and only if $x^{\varepsilon} > y^{\varepsilon}$. Thus, the same approach can be applied to composite indices such as the human poverty index and the inequality adjusted human development index proposed by Foster et al. (2005).

Measuring Robustness

Our method of evaluating the robustness of the comparison x $\mathbf{C}_0$ y fixes a set S_{ε}^{D-1}

[21] This is analogous to a class of human development indices proposed by Chakravarty (2003).

of weighting vectors and confirms that the ranking at w^0 is not reversed using any other $w \in S_\varepsilon^{D-1}$, in which case the associated ε-robustness relation applies. Theorem 8 provides simple conditions for checking whether $x \ \mathbf{C}_\varepsilon \ y$ holds. The present section augments this approach by formulating a robustness measure that associates with any comparison $x \ \mathbf{C}_0 \ y$ a number $r \in [0, 1]$ that indicates its level of robustness.

We construct r using two statistics — one that might be expected to move in line with robustness and another that is likely to work against it. The first of these is $A = C(x; w^0) - C(y; w^0) > 0$, or the difference between the composite value of x and the composite value of y at the initial weighting vector w^0. Intuitively, A is an indicator of the strength of the dominance of x over y at the initial weighting vector. The second is $B = \max_{w \in S^{D-1}}[C(y; w) - C(x; w), 0]$, or the maximal "contrary" difference between the composite values of y and x. Note that when the original comparison is fully robust, then $C(y; w) - C(x; w) \leq 0$ for all $w \in S^{D-1}$ and there is no contrary difference. Consequently $B = 0$. On the other hand, when the comparison is not fully robust, then $C(y; w) - C(x; w) > 0$ for some $w \in S^{D-1}$, and hence $B = \max_{w \in S^{D-1}}[C(y; w) - C(x; w)] > 0$. B is the worst-case estimate of how far the original difference at w^0 could be reversed at some other weighting vector.

We propose $r = A/(A + B)$ as a measure of robustness.[22] Notice that when the initial comparison $x \ \mathbf{C}_0 \ y$ is fully robust, then $B = 0$ and hence $r = 1$. Alternatively, when the initial comparison is not fully robust and $B > 0$, the measure r is strictly increasing in the magnitude of the initial comparison A, and strictly decreasing in the magnitude of the contrary worst-case evaluation B. In addition, if A tends to 0 while B remains fixed,

[22] Permanyer (2007) proposes a different method for measuring robustness. He suggests the radius of the largest ball around the initial weight for which the initial comparison is not reversed as the measure of robustness. His measure has a number of short comings. One limitation of this measure is that it is arbitrarily chosen and no justification is provided.

the measure of robustness r will also tend to 0. These characteristics accord well with an intuitive understanding of how A and B might affect robustness.

Practical applications of r may be hampered by the fact that it requires a maximization problem to be solved, namely, $\max_{w \in S^{D-1}}[C(y;w) - C(x;w)]$. However, by the linearity of $C(y;w) - C(x;w) = (y - x) \cdot w$ in w, the problem has a solution at some vertex v_d where the difference $C(y;w) - C(x;w)$ becomes $y_d - x_d$. Consequently, $B = \max_d(y_d - x_d)$, or the maximum coordinate-wise difference between y and x. The measure r can be readily derived using this equivalent definition.

Now what is the relationship between the robustness measure r and the relation $\mathbf{C}_\varepsilon$ developed in the previous section? The following theorem provides the answer.

Theorem 9 *Suppose that x $\mathbf{C}_0$ y for $x, y \in X$ and let r be the robustness level associated with this comparison. Then the ε-robustness relation x $\mathbf{C}_\varepsilon$ y holds if and only if $\varepsilon \leq r$.*

Proof. Let x $\mathbf{C}_0$ y and suppose that $0 < \varepsilon \leq r$. By the definition of r, we have $\varepsilon \leq A/(A+B)$ and hence $\varepsilon B \leq (1-\varepsilon)A$. Pick any $d = 1, \ldots, D$. Then using the definitions of A and B, we see that $\varepsilon(y_d - x_d) \leq (1-\varepsilon)(w^0 \cdot x - w^0 \cdot y)$ and hence $\varepsilon v_d \cdot y + (1-\varepsilon)w^0 \cdot y \leq \varepsilon v_d \cdot x + (1-\varepsilon)w^0 \cdot x$. Consequently, $v_d^\varepsilon \cdot y \leq v_d^\varepsilon \cdot x$, and since this is true for all d, it follows that $x^\varepsilon \geq y^\varepsilon$ and hence x $\mathbf{C}_\varepsilon$ y by Theorem 8.

Conversely, suppose that x $\mathbf{C}_0$ y and $r < \varepsilon \leq 1$. Then $(1-\varepsilon)A < \varepsilon B$ so that $(1-\varepsilon)(w^0 \cdot x - w^0 \cdot y) < \varepsilon(y_d - x_d)$ for some d, and hence $v_d^\varepsilon \cdot y > v_d^\varepsilon \cdot x$ or $y_d^\varepsilon > x_d^\varepsilon$ for this same d. It follows, then, that $x^r \geq y^r$ cannot hold, and neither can x $\mathbf{C}_\varepsilon$ y by Theorem 8. ■

Raising ε leads to a more demanding robustness criterion and a more incomplete relation $\mathbf{C}_\varepsilon$. Theorem 9 identifies r as the *maximal* ε for which x $\mathbf{C}_\varepsilon$ y holds, and hence the largest

Table 3: Human Development Index: The Top 10 Countries in 2004

Country	**HDI**	**Rank**
Norway	0.965	1
Iceland	0.960	2
Australia	0.957	3
Ireland	0.956	4
Sweden	0.951	5
Canada	0.950	6
Japan	0.949	7
United States	0.948	8
Switzerland	0.947	9
Netherlands	0.947	10

set S_ε^{D-1} for which the original comparison is not reversed. Alternatively, it corresponds to the lowest level of confidence $(1-\varepsilon)$ for which the Gilboa-Schmeidler (or Ellsburg) evaluation function of the net indicator vector $(x-y)$ is always nonnegative; i.e., the largest ε for which $G_\varepsilon(x-y) = (1-\varepsilon)C(x-y,w^0) + \varepsilon \min_{w \in S^{D-1}} C(x-y,w) \geq 0$.

Illustrative Example

We illustrate our methods using data from the 2004 Human Development Index (HDI) dataset as published in the 2006 Human Development Report.[23] The HDI is a composite index $C(x;w^0)$ constructed by taking the simple average of three dimension-specific indicators (of education, health and income) and hence $w^0 = (1/3,1/3,1/3)$ is the initial weighting vector. Table 3 provides information on the top ten countries according to the HDI, includ-

[23] Our underlying dataset was obtained directly from the UNDP and is less severely rounded off than the published data. Thanks to Alison Kennedy for making these data available for our use.

ing their rankings and HDI values.[24] This yields the C_0 relation over these 10 countries, but says nothing about the robustness of any given judgment.

Table 4: Robustness of Three HDI Comparisons

Rank	**Country**	**HDI**	**Health** x_1	**Educ.** x_2	**Income** x_3	**Health** $x_1^{0.25}$	**Educ.** $x_2^{0.25}$	**Income** $x_3^{0.25}$
3	Australia	0.957	0.925	0.993	0.954	0.949	0.966	0.956
5	Sweden	0.951	0.922	0.982	0.949	0.944	0.959	0.951
2	Iceland	0.960	0.931	0.981	0.968	0.953	0.965	0.962
8	USA	0.948	0.875	0.971	0.999	0.93	0.954	0.961
4	Ireland	0.956	0.882	0.99	0.995	0.937	0.964	0.966
6	Canada	0.950	0.919	0.97	0.959	0.942	0.955	0.952

Table 4 focuses on three specific comparisons; the middle columns provide the dimensional indicators x_1, x_2, and x_3 needed to ascertain whether full robustness C_1 obtains. The indicator vector for Australia dominates the indicator vector for Sweden, and hence by Theorem 7, this comparison is fully robust. However, the comparison for Iceland and USA reverses in the income dimension, while the Ireland and Canada comparison has a reversal in health, and so neither of these rankings is fully robust. Observe that the HDI margin between Australia and Sweden (0.006) is identical to the margin for Ireland and Canada, and yet the robustness characteristics of the two comparisons are quite different. The HDI margin between Iceland and USA is twice as large (0.012) and yet it too is not fully robust.

The final columns of Table 4 give the entries of the associated x^{ε} vectors for $\varepsilon = 0.25$ in order to ascertain ε-robustness of the comparisons. A quick evaluation in terms of vector dominance reveals that both the Australia/Sweden and the Iceland/USA comparisons are

[24] Due to rounding, the HDI levels of Switzerland and Netherlands appear to be equal; in fact, Switzerland has a slightly higher HDI than Netherlands.

ε-robust, but the reversal in the Ireland/Canada comparison implies that ε-robustness does not hold for this ranking when $\varepsilon = 0.25$. By Theorem 8 we know that there are weighting vectors in S_ε^{D-1} for which Canada has a higher composite index level than Ireland.

Figure 2: Graphical Analysis of Three HDI Comparisons

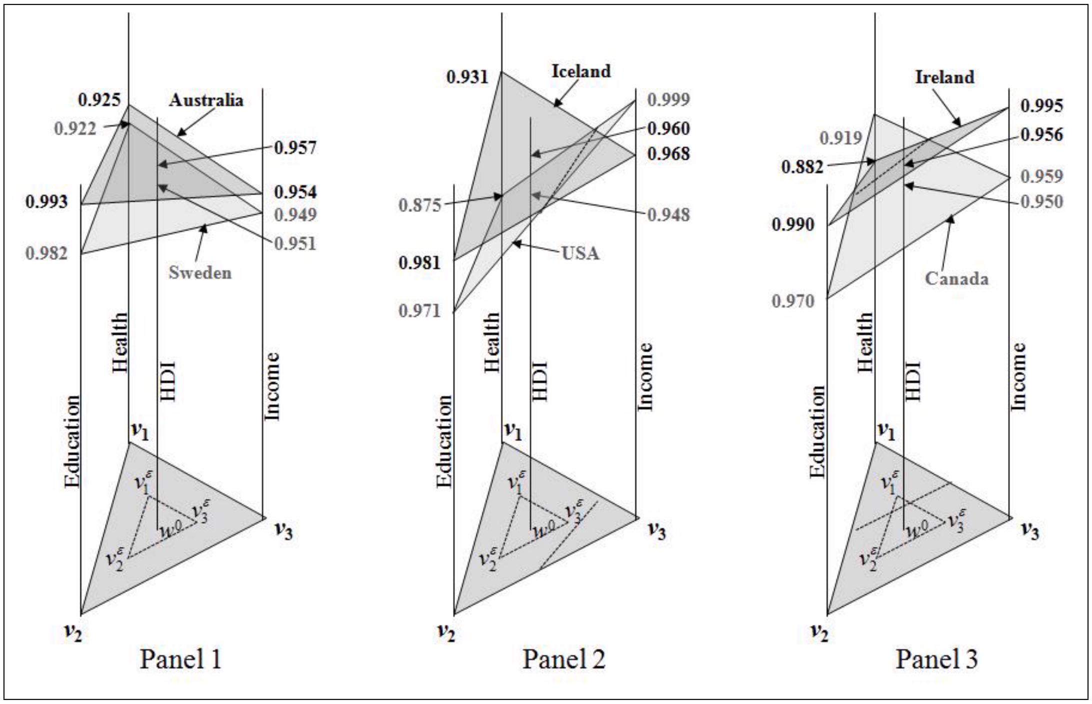


Figure 2 provides a graphical representation of these comparisons. At the base of each diagram is the simplex S^{D-1} of all weighting vectors, including the three vectors v_1, v_2, and v_3 at its vertices and the initial weighting vector w^0 at its center. Also depicted is the smaller set S_ε^{D-1} of weighting vectors and its vertices v_d^ε for $d = 1, 2, 3$, where $\varepsilon = 0.25$. Now suppose that a given country with indicator vector x has been selected. For any weighting vector w in the simplex, the level of the composite index $C(x; w)$ can be graphed as the height above w. The heights above v_1, v_2, v_3, and v^0 are, respectively, the dimensional indicators x_1, x_2, x_3, and the HDI. The linearity of C in w ensures that these points and the remaining $C(x; w)$ values form a tilted "indicator simplex" with vertices as high as the dimensional indicators and a center as high as the country's HDI level. The height at v_d^ε is x_d^ε for $d = 1, 2, 3$, which together are the coordinates of the vector x^ε.

Panel 1 of Figure 2 shows that Australia has a higher HDI level than Sweden. Each of its dimensional indicators is higher, so that vector dominance applies. Consequently, the indicator simplex of Australia is everywhere above the indicator simplex for Sweden, reflective of the fact that full robustness, or $\mathbf{C}_1$, holds. The second panel depicts the rather different scenario for Iceland and USA. The HDI margin is twice as large as in Panel 1, but the indicator simplexes intersect and $\mathbf{C}_1$ does not hold. Iceland performs better than USA in terms of education and health, but has lower indicator value in terms of income. More weight on income can make the USA's composite index level higher than Iceland's. However if we restrict consideration to the smaller set S_ε^{D-1}, no reversals are possible. The intersection of the two indicator simplexes (where composite index levels are equal) projects down to weighting vectors that are outside of S_ε^{D-1}, and $\mathbf{C}_\varepsilon$ holds for $\varepsilon = 0.25$.

The final panel depicts the case of Ireland and Canada, which has the same HDI margin as Panel 1 and intersecting indicator simplexes as in Panel 2, but has different robustness characteristics than both. While Ireland's education and income variables are higher than Canada's, the health index has the opposite orientation, and $\mathbf{C}_1$ cannot hold. If we project the intersection of the indicator simplexes onto S^{D-1}, we obtain a dashed line that cuts through S_ε^{D-1}, implying that $\mathbf{C}_\varepsilon$ does not hold and the Ireland-Canada comparison is not robust for $\varepsilon = 0.25$. This is also evident from Table 4 since Ireland has higher levels of the composite index at two of the vertices of S_ε^{D-1} (namely, v_2^ε and v_3^ε) and a lower level at the remaining one (v_1^ε).

The levels of robustness can also be calculated for each of these comparisons. The Aus-

Table 5: Measure of Robustness

Country		Nor	Ice	Aus	Ire	Swe	Can	Jap	USA	Swit	Neth
	Rank	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Norway	*1*	—									
Iceland	*2*	20	—								
Australia	*3*	35	19	—							
Ireland	*4*	86	14	4	—						
Sweden	*5*	53	94	100	11	—					
Canada	*6*	61	100	60	14	14	—				
Japan	*7*	28	34	23	9	7	2	—			
USA	*8*	77	28	17	67	5	3	1	—		
Switzerland	*9*	49	100	41	16	17	20	6	2	—	
Netherland	*10*	100	68	57	47	25	13	4	7	1	—

tralia/Sweden comparison is fully robust, with $A = 0.006$ and $B = 0$, and hence $r = 100\%$. The Iceland/USA comparison has $A = 0.012$ and $B = 0.031$, and hence $r = 28\%$. In contrast, the Ireland/Canada ranking has $A = 0.006$ and $B = 0.037$, and therefore $r = 14\%$. Table 5 presents the level of robustness of pair-wise comparisons for the top ten countries in the HDI ranking. For every cell below the diagonal the "column country" of the cell has a higher ranking according to $\mathbf{C}_0$ than the "row country". The number in the cell indicates the level of robustness of the associated comparison, expressed in percentage terms. Out of the 45 pair-wise comparisons, four are fully robust as denoted by $r = 100\%$, while 20 of them, or 44.4 percent, are robust at $r = 25\%$. For the entire dataset of 177 countries for the same year, 69.7 percent of the comparisons are fully robust and about 92 percent are robust for $r = 25\%$.

The Prevalence of Robust Comparisons

The focus now shifts from individual comparisons to the entire collection of comparisons associated with a given dataset $\hat{X}$ and an initial weighting vector w^0. The first question is how to judge the overall robustness of the dataset. One option would be to use an aggregate measure (such as the mean) that is strictly increasing in each comparison's robustness level. However, rather than settling on a specific measure we use a "prevalence function" based on the entire cumulative distribution of robustness levels, and employ a criterion analogous to first order stochastic dominance to indicate greater robustness.

Suppose the initial weighting vector is w^0 and there is a dataset $\hat{X}$ containing n observations. Without loss of generality, we enumerate the elements of $\hat{X}$ as $x^1, x^2, \ldots, x^n$ where $C_0(x^1) \geq C_0(x^2) \geq \ldots \geq C_0(x^n)$. The analysis can be simplified by assuming that no two observations in $\hat{X}$ have the same initial composite value, so that $C_0(x^1) > C_0(x^2) > \ldots > C_0(x^n)$.[25] There are $k = n(n-1)/2$ ordered pairs of observations x^i and x^j with $i < j$, and each comparison $x^i \; \mathbf{C}_0 \; x^j$ has an associated robustness level r_{ij}. Let $P = [r_{ij}]$ represent the *robustness profile* of $\hat{X}$ (given w^0), which lists the level of robustness r_{ij} for every ordered pair in a manner similar to Table 5.

The mean robustness level in profile P is given by $\bar{r} = \sum_i \sum_{j>i} r_{ij}/\hat{k}$; it is the average level of robustness of the k comparisons. Of course, a higher mean level $\bar{r}$ does not necessarily tell us anything about the prevalence of $\mathbf{C}_r$ comparisons (or comparisons whose robustness levels are at least r) for any specific r. An alternate approach is to summarize robustness levels in P in a way that reflects the entire distribution, and not just the average. For any given dataset and initial weighting vector w^0, define the prevalence function $p : [0,1] \to [0,1]$

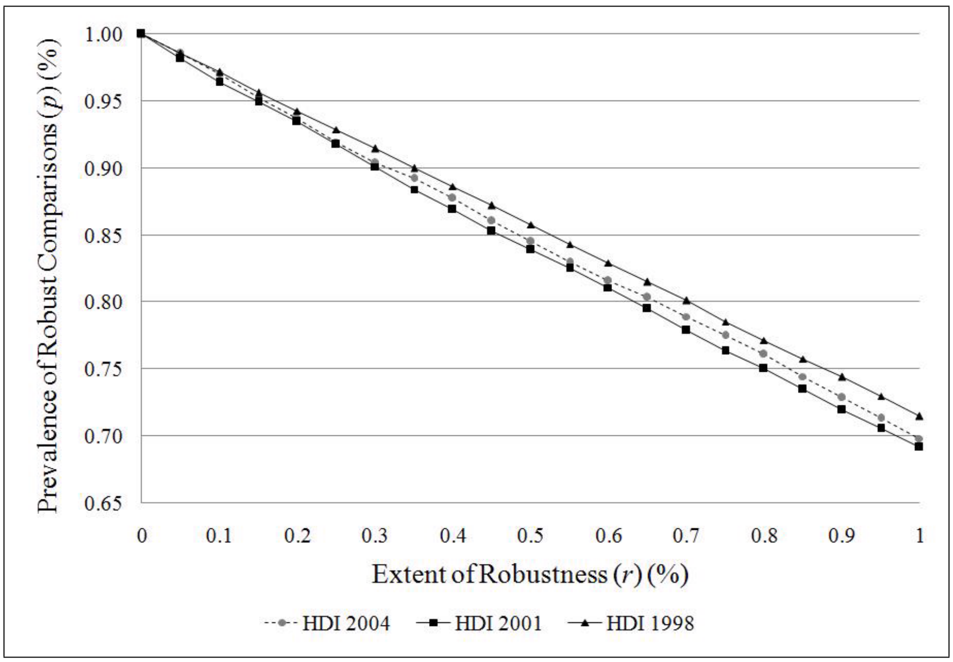
[25] This is true for each of the examples presented below.

to be the function which associates with each $r \in [0,1]$ the share $p(r) \in [0,1]$ of the k comparisons whose robustness levels are at least r. In other words, $p(r)$ is the proportion of comparisons for which the $\mathbf{C}_r$ relation applies.[26]

Illustrative Example

Figure 3 depicts the prevalence functions obtained from HDI datasets for three different years, which uses equal weights across three dimensions to rank 177, 175, and 174 countries, respectively.[27]

Figure 3: Prevalence Functions of HDI for Various Years



Several initial observations can be made from the prevalence functions given in Figure 3. Each graph is downward sloping; reflecting the fact that as r rises, the number of comparisons that can be made by $\mathbf{C}_r$ is lower (or no higher). As r falls to 0, all functions achieve the

[26] At $r = 0$ the complete relation $\mathbf{C}_0$ is used and hence $p(0) = 1$.

[27] Note that the Human Development Indices for the years 1998 and 2004 are obtained from UNDP (2000 and 2006), respectively.

100% comparability arising from $\mathbf{C}_0$; in the other direction, the value of $p(r)$ at $r = 1$ is the percentage of the comparisons that can be compared using $\mathbf{C}_1$ and hence is fully robust. There is a wide variation in $p(1)$ across datasets. It is reasonably large for all the HDI examples, with $p(1)$ being about 69.8% in 2004, 69.2% in 2001, and 71.5% in 1998. The shapes of the $p(r)$ functions are essentially linear for all three HDI dataset. These regularities of prevalence functions are worth examining from a more theoretical perspective. If we set a target of 25 percent robustness, then on an average 92 percent to 93 percent of the HDI comparisons are robust.

Conclusion

Rankings arising from composite indices receive remarkable attention. Yet they are dependent upon an initial weighting vector, and any given judgment could, in principle, be reversed if an alternative weighting vector was employed. This leads one to question rankings provided by composite indices, especially if there is ambiguity over the numerical values of the weights they employ. Many well known and widely used indices are characterised in this way.

Using an analytical framework based on partial ordering and the model of epsilon-contamination used in Bayesian Statistics and Decision Theory, this chapter examined a variable-weight robustness criterion for composite indices that views a comparison as robust if the ranking is not reversed at any weight vector within a given set. It characterized the resulting robustness relations for various sets of weighting vectors. These robustness relationships moderate the complete ordering generated by the composite index. A measure by which the robustness of a given comparison may be gauged was then proposed, and illustrated using the Human Development Index (HDI). The chapter also demonstrated how some rankings are fully robust to changes in weights while others are quite fragile. Finally, the chapter investigated the prevalence of the different levels of robustness in theory and practice and offer insight as to why certain datasets tend to have more robust comparisons.

It was emphasised at the outset of the chapter that its intention was not to discredit or discourage the use of composite indices, but to facilitate better use of them. The chapter helps in this regard by reducing the undue emphasis placed on ranking that are not robust to the choice of weight vector, hopefully placing greater emphasis on those rankings that have higher robustness. It promotes this outcome by allowing end users of composite indices to discern between robust and non-robust comparisons, thereby making the HDI and other composite indices more useful and less misleading.

Two findings of the chapter are worth highlighting further. Both are suggestive of additional research. The first relates to the interesting empirical observation that the prevalence functions associated with the HDI datasets are nearly linear. In other words, increasing r by a given amount decreases the prevalence of robust rankings by a fixed amount, independent of the initial level of r. This means that in the case of the HDI, the entire shape of $p(r)$ is determined by the percentage of fully robust comparisons, $p(1)$. Hence if one were to remove from consideration all fully robust comparisons, the conditional prevalence functions would be virtually identical. Put differently, among all comparisons that are not fully robust, the percentage of comparisons having robustness level r or less is r itself; so, for example, only 5% of these comparisons have robustness level of 0.05 or less (or, equivalently, 0.95 or more). It would be interesting to explore this regularity further.

Further directions for future research ought to be emphasised. The first is to develop and integrate statistical robustness into the analysis. The second involves a link with a theoretical literature that addresses uncertainty. The structure of the general robustness relation defined above for a given set W of weighting vectors is closely related to discussions of "Knightian uncertainty" (Bewley, 1986) and ambiguity in which an individual decision maker has a set of prior probability distributions instead of a single prior. One way in which the approach of this chapter differs from this literature is that it privileges the initial weighting structure instead of treating it as just one of many. The criterion discussed by Bewely requires a strict improvement for possible probability vectors. This chapter's approach allows the comparison to be weak using the other non-distinguished vectors in the set. Nonetheless, there is a fundamental link between the two approaches that would be interesting to explore. Thirdly, the interpretation given in the chapter is that each dimension of the indicator vector is the measured amount of a given indicator. One could instead view the dimensions as being obtained from the underlying indicators by some transformation based on, for example, the utility or welfare from the specific dimension. Such an approach might well be adapted to deal with this case and with other departures from the linearity inherent in composite indices. Finally, there is clearly a link between the fully robust criterion outlined in this paper and first order stochastic dominance in the multidimensional setting. How does the r^{th} degree robustness relate to multidimensional stochastic dominance? Is there an analogue in the framework employed by this chapter to second order multidimensional stochastic dominance? It would be interesting to pursue this direction as well.

D. Proof of Theorem 6

Proof. Let $\mathbf{R}$ be a binary relation on a set X that is closed, convex, and has some z in its interior.

If $\mathbf{R} = \mathbf{R}_W$ for some non-empty, closed, and convex $W \subseteq S^{D-1}$, then it is immediate that $\mathbf{R}_W$ satisfies Q, M, I, and C.

Conversely, suppose that $\mathbf{R}$ satisfies Q, M, I, and C. Define $U = \{x \in X : x \; R \; z\}$ be the upper contour set of $\mathbf{R}$ at z. We know that $z \in U$ by Q and U is closed by C. Moreover, we can show that U is convex. Pick any $x, y \in U$. Let $x' = \alpha x + (1-\alpha)y$ for some α with $0 < \alpha < 1$. Then, where $z' = \alpha z + (1-\alpha)y$, we have $x', z' \in X$ and by axiom I it follows that x' $\mathbf{R}$ z'. Moreover, by a second application of I, it follows from y $\mathbf{R}$ z that z' $\mathbf{R}$ z. Therefore, by Q we have x' $\mathbf{R}$ z and so U is convex.

Since, z is in the interior of X, there exists $\lambda > 0$ such that $N_\lambda = \{x \in \mathbb{R}^D : \|x - z\| \leq \lambda\} \subseteq X$. Define $U_\lambda = U \cap N_\lambda$ and note that it is compact, convex, and contains z, so that the set $K_\lambda = \{z\} - U_\lambda$ is compact, convex, and contains 0. Let $K = \text{Cone } K_\lambda$ be the cone generated by K_λ. It is immediate that K is closed, compact, and contains 0. We can state that K has the property that for $x, y \in X$ we have x $\mathbf{R}$ y if and only if $y - x \in K$. To see this, let $x, y \in X$ and select $\alpha > 0$ small enough that z' satisfying $z = \alpha y + (1-\alpha)z'$ lies in N_λ and $x' = \alpha x + (1-\alpha)z'$ is also in N_λ. Clearly, $z - x' = \alpha(y - x)$ for $\alpha > 0$. So if x $\mathbf{R}$ y, we know that x' $\mathbf{R}$ z by I, and hence $z - x' \in K$ which implies $y - x \in K$. On the other hand, if $y - x \in K$, then since $z - x' \in K$, we have x' $\mathbf{R}$ z so that x $\mathbf{R}$ y by I, establishing the result.

Now let $P = \{p \in \mathbb{R}^D : p \cdot k \leq 0 \text{ for all } k \in K\}$ be the polar cone of K, so that by standard results on polar cones, P is closed and convex. It is clear that $P \subseteq \mathbb{R}^D_+$, since by monotonicity, we have $-v_d \in K$ and so $p \cdot (-v_d) \leq 0$ and $p_d \geq 0$, where v_d is the D-dimensional usual basis vector for co-ordinate d. In addition, we can show that P contains at least one element $p \neq 0$. Indeed, it is clear from M that K contains no $k >> 0$ (otherwise, we would have $x << z$ with x $\mathbf{R}$ z). Then, $K \cap \mathbb{R}^D_{++} = \varnothing$ and since both sets are convex, we can apply the Minkowski separation theorem to find $p^0 \neq 0$ in P. Let $W = S^{D-1} \cap P$, so that $\operatorname{Cone} W = P$. Clearly, K is the polar cone of both P and W, hence, $K = \{t \in \mathbb{R}^D : w \cdot t \leq 0 \text{ for all } w \in W\}$.

We now show that $\mathbf{R} = \mathbf{R}_W$. If x $\mathbf{R}$ y, then $y - x \in K$ and so $w(y - x) \leq 0$ for all $w \in W$, hence x $\mathbf{R}_W$ y. Conversely, if x $\mathbf{R}_W$ y, then by definition we have $w(y - x) \leq 0$ for all $w \in W$, hence $x - y \in K$ or x $\mathbf{R}$ y. ■

Chapter IV

Composite Indices: Rank Robustness, Statistical Association, and Redundancy

(with James Foster and Mark McGillivray)

Introduction

It is noted in Chapter III that the rankings yielded by composite indices can sometimes be reversed by a plausible change to the initial vector of weights, while in other cases, the rankings yielded are preserved when the vector of weights is changed. In Chapter III, we define and characterize a general rank robustness criterion that discerns between these situations for a given initial weighting vector and provide necessary and sufficient conditions for rankings to exhibit full robustness (where weights can range the entire simplex) or a weaker form of robustness (where weights are restricted to a smaller simplex around the initial weighting vector). We propose a practical measure to evaluate the level of robustness of given comparison (say, between Norway and Denmark). This provides a useful toolkit for judging the robustness of the rankings generated by composite indices.

Empirical applications of these methods reveal that there is a wide variation in the prevalence of robust comparisons as one evaluates different composite indices over their respective datasets. Why do some composite indices appear to have more robust comparisons than others? What characteristics of a given dataset are related to the prevalence of robust comparisons? These are the questions addressed in the present chapter.

We begin by analyzing the prevalence of robustness for several well-known composite indices on their respective datasets, and show that some of them have much greater robustness than others. We examine various ways of transforming datasets, including transformations that leave the prevalence of robust comparisons unchanged and others that increase robustness. We explore a key determinant of the robustness – the statistical association between component variables – and establish a key relationship between the prevalence of full robustness and the well-known Kendall tau rank correlation coefficient. These results shed new light on the role of positive association in multidimensional measurement. Previous research argued that high associations between component variables are undesirable as they are indicative of redundancy, which occurs when one component provides largely the same information as the index as a whole. The present results reveal a favorable aspect of positive association: its impact on the robustness of the associated rankings.

The chapter consists of five additional sections. The second section briefly presents the robustness approach of Chapter III. The third section examines the prevalence of robustness for three well-known composite indices. The fourth section provides several theorems on the prevalence of robustness and in particular investigates how the statistical association between components affects robustness. The fifth section looks at the issue of redundancy and its relationship to rank robustness. The final section concludes the chapter.

Robustness

We first outline the notation and definitions used in our analysis of robustness. Let $D \geq 2$ be the number of dimensions under consideration. For the two D-dimensional vectors a and b, the expression $a \geq b$ means that $a_d \geq b_d$ for all $d = 1, \ldots, D$, which is the *vector dominance* relation. If $a \geq b$ and $a \neq b$, then this situation is indicated by $a > b$; whereas $a >> b$ denotes $a_d > b_d$ for all $d = 1, \ldots, D$. The *least upper bound* of a and b, denoted by

$a \vee b$, is the vector having $\max\{a_d, b_d\}$ as its d^{th} coordinate; the *greatest lower bound* of a and b, denoted by $a \wedge b$, is the vector having $\min\{a_d, b_d\}$ as its d^{th} coordinate.

Let $X \subset \mathbb{R}^D$ be the set of achievement vectors and let $S^D = \{s \in \mathbb{R}^{D+1} : s \geq 0$ and $\sum_{d=1}^{D+1} s_d = 1\}$ denote the simplex of associated weighting vectors. A *composite index* $C : X \times S^{D-1} \rightarrow \mathbb{R}$ combines the dimensional achievements in $x \in X$ using a weighting vector $w \in S^{D-1}$ to obtain an aggregate level $C(x; w) = w \cdot x$. We assume that an *initial weighting vector* $w^0 \in S^{D-1}$ has already been selected and this fixes the specific composite index $C_0 : X \rightarrow \mathbb{R}$ defined as $C_0(x) = C(x; w^0)$ for all $x \in X$. The associated strict ordering of achievement vectors will be denoted by $\mathbf{C}_0$, so that x $\mathbf{C}_0$ y holds if and only if $C_0(x; w^0) > C_0(y; w^0)$. We let $\hat{X} \in \mathbb{R}^{ND}$ denote a *dataset* of achievement vectors with the nd^{th} element being the d^{th} achievement of the n^{th} achievement vector. The n^{th} row of the dataset is denoted by $x^n \in \mathbb{R}^D$, which is the n^{th} achievement vector for $n = 1, \ldots, N$. Without loss of generality, we assume that $C_0(x^1) \geq C_0(x^2) \geq \ldots \geq C_0(x^N)$.

Our treatment of robustness is normative, being based on an epsilon-contamination model of ambiguity, is closely related to the theory of Knightian uncertainty (Bewley, 2002). In this treatment, if $C(x; w^0) > C(y; w^0)$, the comparison between a pair of achievement vectors, x and y in $\mathbb{R}^D$, is considered to be *fully robust*, denoted by x $\mathbf{C}_1$ y, if $C(x; w) \geq C(y; w)$ no matter what weighting vector $w \in S^{D-1}$ is used. It can be shown that full robustness is equivalent to $x > y$. When vector dominance does not hold, and the comparison is not fully robust, in Chapter III, we propose using intermediate partial orderings $\mathbf{C}_r$ for $0 < r < 1$, which requires agreement over a smaller set of weighting vectors $S_r^{D-1} = \{w \in S^{D-1} : w = rw^0 + (1-r)w'\}$ for some $w' \in S^{D-1}\}$. Clearly, $S_1^{D-1} = S^{D-1}$ and S_r^{D-1} contracts to $\{w^0\}$ when r tends to 0, so that C_r becomes less stringent, and more complete, as r falls to 0. We have construct vectors x^r and y^r having the property that $x \mathbf{C}_r y$ if and only if $x^r > y^r$; this provides a straightforward test for checking whether $\mathbf{C}_r$ holds.[28]

Chapter III also presents a related method for measuring the level of robustness of a given comparison. Suppose that $x \mathbf{C}_0 y$ and hence $C(x; w^0) > C(y; w^0)$ for a given pair $x, y \in X$. Let $A = C(x; w^0) - C(y; w^0)$ be the difference in aggregate achievements using the initial weighting vector w^0. In the context of the Human Development Index, this is analogous to the difference in HDI values for two countries, and represents the margin by which x dominates y. Let $B = \max_{w \in S^{D-1}}[C(y; w) - C(x; w), 0]$ be the maximum contrary difference between aggregate achievements as w ranges across S^{D-1}. This represents the maximum margin by which y could dominate x if the weights were allowed to vary. The measure of robustness is defined by $r = A/(A + B)$. Intuitively, when $B = 0$ so that full robustness $x \mathbf{C}_1 y$ holds, then $r = 1$; when B becomes large relative to A, then the measure of robustness r falls towards 0. It turns out that the maximum possible contrary difference B is also the maximum coordinate-wise difference between y and x, and hence r is straightforward to calculate. Moreover, it can be shown that the intermediate robustness ordering $x \mathbf{C}_{r'} y$ holds for all r' below or equal to the robustness level r, but fails to hold for $r' > r$; so the robustness measure gives the highest (or most stringent) robustness ordering that is applicable to the given pair.

Chapter III goes on to define a prevalence function to analyze how the share of robust comparisons varies with the specific level of robustness. To illustrate this analysis, we assume that there is a dataset $\hat{X}$ with N observations and that overall achievement is calculated using an initial weighting vector w^0. The analysis is simplified by assuming that no two

[28] Specifically, $x^r = r(C_0(x)\mathbf{1}) + (1 - r)x$ and $y^r = r(C_0(y)\mathbf{1}) + (1 - r)y$, where $\mathbf{1} = (1, 1, \ldots, 1)$ is the unit vector.

observations in $\hat{X}$ have the same initial composite value, so that $C_0(x^1) > C_0(x^2) > \ldots > C_0(x^N)$.[29] There are $k = N(N-1)/2$ ordered pairs of observations x^i and x^j with $i < j$, and each comparison $x^i\ \mathbf{C}_0\ x^j$ has an associated robustness level r_{ij}. Let $P = [r_{ij}]$ represent the robustness profile of $\hat{X}$ (given w^0). For any given dataset $\hat{X}$ and initial weighting vector w^0, the prevalence function $p : [0,1] \in [0,1]$ associates with each $r \in [0,1]$, the share $p(r) \in [0,1]$ of the k comparisons whose robustness levels are at least r. Equivalently, $p(r)$ is the share of the comparisons for which the r^{th} robustness ordering $\mathbf{C}_r$ holds.

Suppose that p and q are the prevalence functions for dataset $\hat{X}$ (given w^0) and dataset $\hat{Y}$ (given u^0), respectively. We say that $\hat{X}$ *has greater robustness than* $\hat{Y}$ if $p(r) \geq q(r)$ for all $r \in [0,1]$, with $p(r) > q(r)$ for some $r \in [0,1]$. In words, no matter the target level of robustness r, the share of all comparisons in $\hat{X}$ with robustness level r or more is at least as high as the respective share in $\hat{Y}$, and for some r it is higher. The two are said to *have the same robustness* if their prevalence functions are the same. In the next section we apply this approach to several composite indices and compare their associated prevalence functions.

Prevalence of Robustness: Empirical Applications

The prevalence function is now constructed for several composite indices and datasets having the country as their unit of analysis. The first is the Human Development Index or HDI for the years 1998 and 2004, as obtained from the Human Development Reports (United Nations Development Programme, 2000, 2006). The HDI contains three components, capturing achievements in per capita income, education, and health, respectively, and is simply the arithmetic (or equal weighted) mean of the three components. Each component has been normalized to range between zero and one, and hence the HDI take values in the

[29] This is true for each of the composite indices presented below.

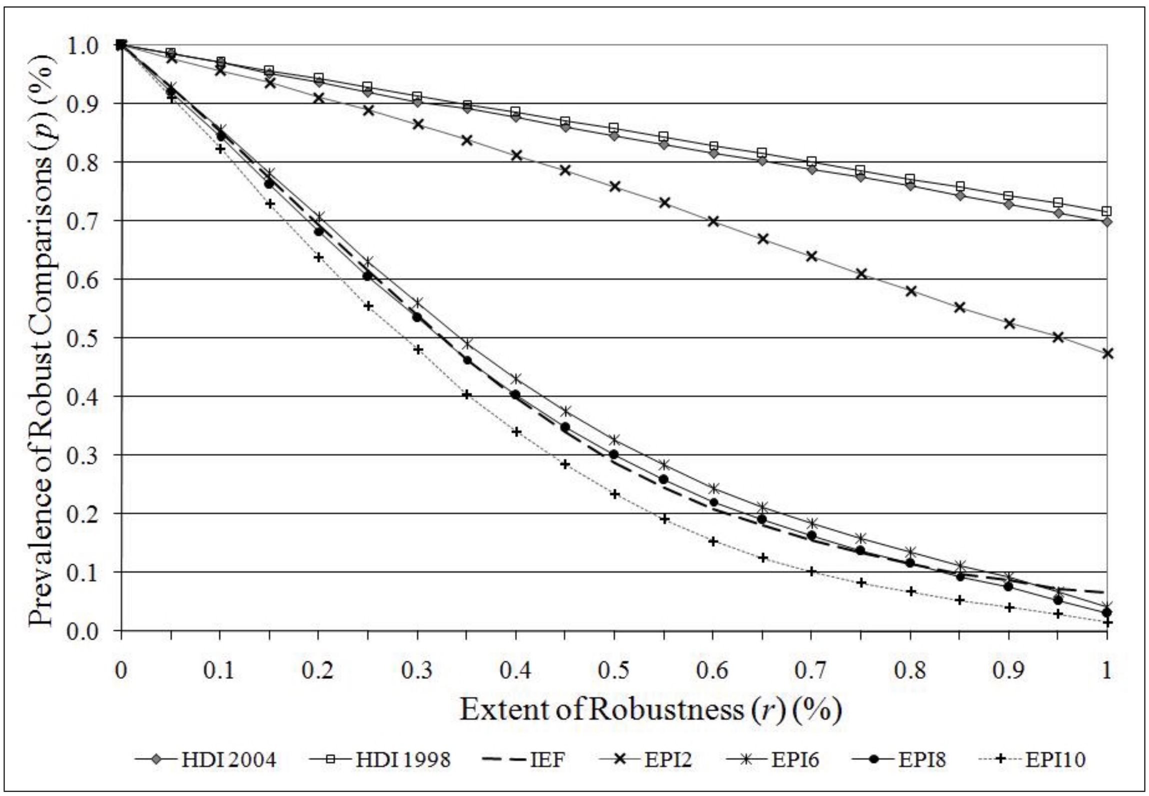
same range. The HDI provides a ranking of 177 countries for each of the above-mentioned years. The second composite index is the Index of Economic Freedom (IEF). The IEF is based on achievement in ten dimensions relevant to economic freedom.[30] Each component index has been normalized to range between zero and one hundred, and the IEF is formed by taking the arithmetic mean of the ten components. We examine the IEF for 2007, which ranks 157 countries. These data were obtained from the Heritage Foundation (2008). The third composite index is the Environmental Performance Index (EPI). The EPI is based on 25 component indices. A number of versions of the EPI exist, each differentiated by the level of aggregation of the components. We examine four versions: EPI2, EPI6, EPI8 and EPI10. EPI2 is based on two equally weighted summary measures of environmental health and ecosystem vitality, respectively. EPI6, EPI8, and EPI10 are based on a mix of summary and individual indices of environmental health, air pollution, the impact of water, biodiversity and habitat, productive natural resources and climate, and are obtained by aggregating six, eight, and ten of these component indices, respectively. Full descriptions of the EPI can be found in Esty et al. (2008). The EPIs under consideration in this chapter rank 149 countries for the year 2007.

Prevalence functions for the above-mentioned composite indices are shown in Figure 4 with $p(r)$ presented in percentage terms. Each function is downward-sloping, reflecting the fact that as r rises, the number of comparisons that can be made by $\mathbf{C}_r$ is lower (or no higher). As r falls to zero, all functions achieve the 100% comparability arising from C_0; in the other direction, the value of $p(r)$ at $r = 1$ is the percentage of the comparisons involving vector dominance, and hence are fully robust. There is, interestingly, a wide variation in

[30] The ten dimensions business freedom, trade freedom, monetary freedom, government size, fiscal freedom, property rights, investment freedom, financial freedom, freedom from corruption, labor freedom.

$p(1)$ across each composite index under consideration. It is clearly highest for the HDI, with $p(1)$ being 69.8% for the 1998 HDI rankings and 73.2% for those of 2004. Put differently, 69.8% and 73.2% of pair-wise HDI comparisons are fully or 100% robust in 2004 and 1998, respectively. The value of $p(1)$ for EPI2 rankings is 47.4%. It is much lower for the remaining indices, being 4.2%, 3.0%, 1.5% and 6.5% the EPI6, EPI8, EPI10 and IEF, respectively.

Figure 4: Prevalence Functions $p(r)$



For all r between zero and one, it is clear from Figure 4 that the robustness is greater for the HDI than for the EPI and the EFI. The 1998 HDI prevalence function is also higher than the 2004 HDI prevalence function. The EPI10 exhibits the lowest prevalence of robust comparisons. An additional feature of Figure 4 is that shapes of the $p(r)$ functions are different, with those associated with the HDI being essentially linear, and the others exhibiting pronounced curvatures. Drawing on these observations, we now examine the prevalence of robustness from a more theoretical perspective and consider transformations that allow the robustness of different composite indices to be compared.

Prevalence, Transformation, and Positive Association

We begin with some basic transformations that leave the prevalence functions fixed. We then consider changes in the dataset that increase the prevalence of fully robust comparisons. In particular, we show the key role played by association among dimensions.

Fixed Robustness and Transformations

Our first transformations yield pairs of datasets that have similar robustness properties. A *monotonically increasing transformation* of X is a function $f : X \to \mathbb{R}^D$ that can be written as $f(x) = (f_1(x_1), \ldots, f_D(x_D))$ where each function $f_d(x_d)$ is monotonically increasing; a *common-slope affine transformation* of X has the additional property that each function $f_d(x_d)$ can be written as $f_d(x_d) = \alpha x_d + \beta_d$ for some $\alpha > 0$ and β_d in $\mathbb{R}$. We say that $\hat{Y}$ *is obtained from* $\hat{X}$ *by a common-slope affine transformation* (respectively, by *a monotonically increasing transformation*) if $\hat{Y} = \{f(x) : x \in \hat{X}\}$ for some transformation f having the appropriate property.

Applying a monotonically increasing transformation to a dataset preserves the orderings of achievements within each dimension, but can disrupt the weighted averages across dimensions. In particular, it is possible that $C(x'; w^0) > C(x; w^0)$ and $C(y'; w^0) < C(y; w^0)$ where y' and y are transformations of x' and x, respectively, which implies that the robustness profiles of $\hat{X}$ and $\hat{Y}$ can be rather different for the same w^0. On the other hand, if we restrict consideration to common-slope affine transformations, we see that $C(y;w) = w \cdot y = \alpha w \cdot x + w \cdot \beta$ where $\beta = (\beta_1, \ldots, \beta_D)$, and hence $C(x';w) \geq C(x;w)$ if and only if $C(y';w) \geq C(y;w)$, where y' and y are the respective transformations of x' and x. In this case, $\hat{X}$ and $\hat{Y}$ have the same robustness profile and hence the same prevalence function $p(r)$ given w^0. So, for example, if every dimension is scaled up or down in the same proportion, this will leave $p(r)$ unchanged, as will simply adding a different constant to each dimension. On the other hand, multiplying each dimension by a different positive constant alters the implicit weighting across dimensions, potentially changing the rankings of transformed observations. Using an arbitrary monotonic increasing transformation, likewise, can alter rankings and lead to different prevalence functions for the transformed dataset. Note, though, that fully robust comparisons are preserved under a monotonic transformation, and hence the prevalence $p(1)$ of full robustness does not change. These results are summarized in the following theorem.[31]

Theorem 10 *Suppose that the initial weighting vector is fixed. If $\hat{Y}$ is obtained from $\hat{X}$ by a monotonically increasing transformation, then $\hat{Y}$ and $\hat{X}$ share the same prevalence value $p(1)$. If $\hat{Y}$ is obtained from $\hat{X}$ by a common-slope affine transformation, then they share the same prevalence function $p(r)$.*[32]

In the example of the HDI, the normalized income, education, and health variables used to construct index values are actually monotonic transformations of underlying variables involving a nonlinear function in the case of income, and affine transformations with different

[31] The result on monotonic transformations would be true even if the initial weighting vectors were different. The role played by common-slope affine transformations is similar to assumptions used in social choice theory. See, for example, Blackorby et al. (1984).

[32] The first part of Theorem 10 will generate same prevalence function $p(1)$ even if we use the dominance criterion proposed by Cherchye et al. (2008).

slopes across the three variables. Consequently, the specific shapes of the transformations can influence HDI comparisons as well as their measured robustness levels. However, as indicated in Theorem 10, these transformations do not influence fully robust comparisons and $p(1)$. If one restricts consideration to $\mathbf{C}_1$ comparisons, there would be no need to select the "right" transformations or even to transform variables at all: one could use the original income, education, and health variables directly.

A second form of transformation replaces each variable in the achievement vector with one or more copies of that variable. A *replicating transformation* of X is a function $f : X \to \mathbb{R}^{D'}$ for some $D' > D$ such that $f(x) = (f_1(x_1), \ldots, f_D(x_D))$, where each $f_d(x_d)$ is the k_d-fold replication $(x_d, \ldots, x_d) \in \mathbb{R}^{dk_d}$ for some integer $k_d \geq 1$. We say that $\hat{Y}$ is obtained from $\hat{X}$ by a *replicating transformation* if $\hat{Y} = \{f(x) : x \in \hat{X}\}$ for some transformation f of this type. Transformed achievement vectors have higher dimension D' and, consequently, the associated weighting vectors must be adjusted to account for this. Now, which initial weighting vector u^0 for $\hat{Y}$ would correspond to the original w^0 for $\hat{X}$? One option is to divide the weight equally among the associated dimensions in u^0; however, it turns out that any allocation of the weight across its associated dimensions will do. We say u^0 is consistent with w^0 if, for each $d = 1, \ldots, D$, the weight on x_d is equal to the sum of the k_d entries in u^0 associated with $f_d(x_d) = (x_d, \ldots, x_d)$. So for example, if $D = 2$ and f replicates each entry two times, then $w^0 = (1/2, 1/2)$ is consistent with $u^0 = (1/6, 2/6, 1/4, 1/4)$. We have the following result.

Theorem 11 *If $\hat{Y}$ is obtained from $\hat{X}$ by a replicating transformation, and u^0 is consistent with w^0, then $\hat{Y}$ and $\hat{X}$ have the same prevalence function $p(r)$.*

Proof. See Appendix E. ■

In other words, according to Theorem 11, appending copies of one or more existing variables leaves the comparisons and the robustness properties of a dataset unaffected, as long as the effective weight on each variable is unchanged. As an example, consider what would happen if the education variable in an HDI dataset were replicated to obtain a four variable dataset. Using equal weights of 1/4 for the four dimensional dataset would likely alter rankings since this would, in effect, increase the aggregate weight on education. However, if the total weight on the two education variables is maintained at 1/3, say where each variable receives a weight of 1/6, then all comparisons and robustness levels would be the same as before.

One implication of this is that the number of variables *per se* does not have an independent impact on a dataset's robustness. In contrast, the empirical evidence provided by Figure 4 *does* might suggest that a greater number of variables is associated with lower robustness. The evidence is particularly striking for the three EPI examples, where the aggregation of variables, and hence the decrease in the number of variables, clearly leads to increased robustness — even though they use the same underlying data. Is this due to the decreased number of variables?

Let us examine how EPI6 is constructed from EPI10. The first and fifth variables in EPI6 are each obtained by combining three distinct variables in EPI10 (namely, variables 1-3 and variables 7-9), while the remaining variables are unchanged. Weights from the initial weighting vector u^0 for EPI10 are used to construct each new variable in EPI6 as a weighted average of the source variables from EPI10, and the weight on the new variable is the sum of the corresponding weights in u^0. The new w^0 is thus consistent with u^0. Now consider a ten variable replication of EPI6 that repeats variable 1 three times and variable 5 three times and let the initial weighting vector be u^0. By Theorem 11, this intermediate dataset has precisely the same robustness profile and prevalence function as EPI6. It is not the number of variables that is driving the observed decrease in robustness. Instead, its source is found in the transformation from the intermediate dataset to EPI10, by which the perfectly correlated triplets are converted to variables that are less positively associated. The fall in robustness is due to disagreements among the new variables, rather than the higher number of variables *per se*. Association among variables is likely a key driver of robustness, which is explored further in the next section.

Increased Robustness and Positive Association

What factors generally lead to greater robustness? At an intuitive level, the possibility of fully robust comparisons is related to the degree of correlation or association among the dimensional variables. For example, if two of the achievements are perfectly negatively correlated, so that when one rises, the second falls, then it is impossible for vector dominance and hence **C**1 to hold. On the other hand, if there is complete positive association between all variables, so that when any variable rises, all rise, then every achievement vector is comparable by vector dominance, and **C**1 is universally applicable.[33] We saw in Figure 4 that both HDI datasets have high levels of robustness, and that the prevalence function is higher for 1998 than for 2004. Kendall's tau correlation coefficients for 2004 are 0.55 for health and education, 0.66 for health and income, and 0.58 for income and education, which indicates strong, positive association among variables; the respective values for 1998 are

[33] Note that if there are more than two dimensions, then it is impossible for all pairs of variables to be perfectly negatively correlated; in other words, there is no analogous notion of perfect negative association in higher dimensions.

even higher, at 0.59, 0.70, and 0.60.[34] Both intuition and empirical evidence suggest a link between positive association and robustness. We now turn to the theoretical justification for such a link.

For simplicity, assume that the dataset $\hat{X}$ has the property that within each dimension, all observed values of the variable are distinct.[35] Given any two dimensions c and t, let E_{ct} be the number of *concordant* pairs of observations in which one of the two observations has higher values in both dimensions c and t. Let G_{ct} be the number of *discordant* pairs in which one observation is higher in one dimension and the second is higher in the other. Then Kendall's tau correlation coefficient for dimensions c and t is defined as $\tau_{ct} = (E_{ct} - G_{ct})/(E_{ct} + G_{ct})$. Note that the denominator of this expression is $k = N(N-1)/2$ while $G_{ct} = k - E_{ct}$, so that $\tau_{ct} = 2E_{ct}/k - 1$.

Now consider the special case where there are only two variables, and so there is a single coefficient $\tau = \tau_{12}$ and number $E = E_{12}$ of concordant pairs. In this special case, the number of concordant pairs is precisely the number of fully robust pairs, so the share of fully robust comparisons is $p(1) = E/k$. Therefore, $\tau = 2p(1) - 1$ and we have the following result.

Theorem 12 *Suppose that $D = 2$ for a given dataset $\hat{X}$. Then the share $p(1)$ of fully robust comparisons is determined by Kendall's tau correlation coefficient τ according to the formula $p(1) = (\tau + 1)/2$.*

In the case of two variables, there is a direct relationship between $p(1)$ and the level of association as measured by Kendall's tau. Whenever $\tau = 1$ so that the variables have perfect positive association, we must have $p(1) = 1$. If $\tau = -1$, and perfect negative association

[34] Kendall's tau correlation coefficient is a measure of association or correlation based on ranks of the variables concerned. See Kendall and Dickinson (1990).

[35] This rules out ties and simplifies the definition of Kendall's tau correlation coefficient.

obtains, then $p(1) = 0$. The independence case of $\tau = 0$ implies $p(1) = 1/2$, so that half the comparisons would be fully robust in this case. The example of EPI2 has $\tau = -0.053$ and hence $p(1) = 0.474$ by Theorem 12.

Now consider the general case of $D \geq 2$. Full agreement across all dimensions entails concordance in any two dimensions, hence $p(1) \leq E_{ct}/k = (\tau_{ct} + 1)/2$ for any pair c and t. We have the following result.

Theorem 13 *Let $\tau_{\min} = \min_{c,t} \tau_{ct}$ be the minimum value of Kendall's tau correlation coefficient across all pairs of variables c and t in dataset $\hat{X}$. Then the share $p(1)$ of fully robust comparisons is bounded as follows: $p(1) \leq (\tau_{\min} + 1)/2$.*

This result shows that the smallest Kendall tau coefficient, appropriately transformed, provides us with an upper bound for the proportion of comparisons that are fully robust. If $\tau_{\min} = 1$, so that all pairs of variables move together in full accord, then $p(1) = 1$ and the bound is tight. If $\tau_{\min} = -1$, say, when a pair of variables exhibits a perfect negative association, then no comparison is robust and $p(1) = 0$ is equal to this bounding value. For $0 < \tau_{\min} < 1$, the actual value of $p(1)$ can be equal to or below the bound. For example, for the 2004 HDI dataset, $\tau_{\min} = 0.55$, and thus according to Theorem 13, we have $p(1) \leq 0.78$. As noted above, the actual prevalence of fully robust comparisons is $p(1) = 0.698$. For EPI6, EPI10, and EFI, the respective values of $\tau_{\min}$ are -0.147, -0.237, and -0.3395, yielding upper bounds on $p(1)$ of 0.43, 0.38, and 0.33 respectively. The true values for $p(1)$ are 0.042, 0.015, and 0.065, respectively.

When there are several dimensions, pair-wise associations can provide only partial information on the magnitude of $p(1)$. An interesting alternative is to adjust the definition of Kendall's tau itself to obtain a multidimensional measure of association that corresponds exactly to $p(1)$. Let E be the number of pairs of observations in which one of the two observations has higher values in all dimensions, and G be the number of pairs for which the two observations disagree in at least one dimension. Given any dataset $\hat{X}$ having an arbitrary number of dimensions $D > 0$, we define Kendall's coefficient of positive association by $\tau = (E - G)/(E + G)$, or the number of fully robust comparisons minus the number that are not fully robust, over the total number of comparisons. With dimensional ties ruled out, the total number of comparisons is once again $k = N(N-1)/2$, while $G = k - E$, so that $\tau = 2E/k - 1 = 2p(1) - 1$ and $p(1) = (\tau + 1)/2$.

In the two-dimensional case, the coefficient τ reduces to the standard Kendall's tau; for more dimensions, it requires agreement across all dimensions before counting the comparison as increasing positive association. So for example, the positive association measures for the HDI datasets in 1998 and 2004 are, respectively, $\tau = 0.464$ and $\tau = 0.396$, while for the EFI it drops to $\tau = -0.87$. The coefficient for the EPI dataset rises from -0.97, to -0.916, to -0.053 as we move from largest to smallest number of dimensions. This is a useful way of restating a robustness property of datasets using more familiar terminology, while emphasizing the fundamental link between positive association and robustness.

An alternative route makes use of the general notion of "increasing association" found in Boland and Proschan (1988), among other sources.[36] We say that dataset $\hat{Y}$ is obtained from dataset $\hat{X}$ by an association increasing rearrangement if for some $x \neq x'$, we have: (a) neither $x \geq x'$ nor $x' \geq x$ holds; (b) $y = x \vee x'$ and $y' = x \wedge x'$; and (c) $y'' = x''$ for all $x'' \neq x, x'$. In other words, the datasets are identical apart from a pair of non-comparable

[36] In the literature on multidimensional inequality and poverty, increasing association was first introduced by Atkinson and Bourguignon (1982). Tsui (1999, 2002) based the notion of correlation increasing majorization on the 'basic rearrangement' used by Boland and Proschan (1988).

observations in $\hat{X}$ that were made comparable in $\hat{Y}$ by placing all the higher values in one observation (the least upper bound) and all the lower values in another (the greatest lower bound). We have the following result.

Theorem 14 *Suppose that the initial weighting vector is fixed. If dataset $\hat{Y}$ is obtained from dataset $\hat{X}$ by a series of association increasing rearrangements, then the share $p(1)$ of fully robust comparisons is higher for $\hat{Y}$ than for $\hat{X}$.*

Proof. See Appendix F. ∎

One natural implication of the theorem is that an association increasing rearrangement must lead to a higher value for Kendall's coefficient of positive association τ. It is also easy to see that none of the pair-wise coefficients τ_{ct} will fall, and that at least one will rise. Consequently, this form of transformation is especially useful for illustrating the connection between full robustness and positive association.

Theorem 14 provides information on the share $p(1)$ of fully robust comparisons, but not on $p(r)$ for $r < 1$. The following example shows how greater association across variables need not translate to increased overall prevalence. Suppose that $\hat{X}$ is made up of the four vectors $x^1 = (30, 80)$, $x^2 = (100, 30)$, $x^3 = (90, 100)$, and $x^4 = (80, 120)$. With equal initial weights, we see that $C_0(x^1) = 55$, $C_0(x^2) = 65$, $C_0(x^3) = 95$ and $C_0(x^4) = 100$, and yet only two comparisons x^3 $\mathbf{C}_0$ x^1 and x^4 $\mathbf{C}_0$ x^1 are fully robust. Let $\hat{Y}$ be made up of the four vectors $y^1 = (30, 30)$, $y^2 = (100, 80)$, $y^3 = y^3$, and $y^4 = y^4$, so that is obtained from by an association increasing rearrangement. Then the number of fully robust comparisons rises to three, since now y^2 $\mathbf{C}_0$ y^1, y^3 $\mathbf{C}_0$ y^1, and y^4 $\mathbf{C}_0$ y^1 hold. Clearly, $p(1)$ rises as a result of the association increasing rearrangement.

What about the prevalence $p(r)$ at other values of r? For example, let $r = 0.40$, and note that the respective x^r vectors used in evaluating $\mathbf{C}_r$ are $(45, 65)$, $(79, 51)$, $(93, 97)$, and $(92, 108)$ for $\hat{X}$ and $(30, 30)$, $(94, 86)$, $(93, 97)$, and $(92, 108)$ for $\hat{Y}$. Checking each collection for vector dominance, we find that the number of $\mathbf{C}_r$ comparisons in $\hat{X}$ is four, while only three $\mathbf{C}_r$ comparisons are possible in $\hat{Y}$, and hence $p(r)$ is *negatively* affected by the association increasing rearrangement. Note that the rearrangement results in a vector y^2 that is not comparable to the other two unchanged vectors, y^3 and y^4, and this is preserved in $\mathbf{C}_r$; whereas, the non-comparability of x^2 with x^3 and x^4 does not survive the averaging underlying $\mathbf{C}_r$. Since this example has two dimensions, it also follows that Theorem 12 applies, and Kendall's tau coefficient is higher in $\hat{Y}$ than $\hat{X}$. Consequently, $p(r)$ can strictly fall when there is greater association, or when the tau coefficient between the two dimensions rises. While it is clear that $p(1)$ is linked to positive association among variables, the specific mix of factors that determine the placement and shape of $p(r)$ for $r \in (0, 1)$ has yet to be determined.

Robustness and Redundancy

The results of the previous section show that greater positive association increases the prevalence of fully robust comparisons and, in this sense, is a desirable attribute of a multi-dimensional dataset. There is an alternative literature that takes a rather different view of positive association, and we will now briefly examine these arguments in light of our findings.

A number of previous studies have critiqued the HDI based on the statistical association between the three components used to construct the composite index (McGillivray, 1991, 2005; McGillivray and White, 1993; Cahill, 2005). McGillivray (1991), in particular, provided an argument based on a notion of "redundancy of composition," which arises when there is a strong positive association between a composite index and one of its components. High redundancy of composition is considered to be an undesirable property on the grounds of parsimony: if a single component provides basically the same ranking as the composite index, why not use the former instead of the latter? A second argument invokes the notion of "multidimensionality" of the index: if each pair of component variables is highly correlated, then the index could hardly be characterized as multidimensional, and once again, a single dimension may be all that is needed.

The force of these arguments is mitigated somewhat by our robustness results. To be sure, when the variables are highly correlated in a given dataset, the index may well be tracked by a single component and may act like a unidimensional measure; but the comparisons it makes will tend to be robust.[37] Note that this favorable conclusion (like the critiques) is contingent on the actual dataset employed. At a different point in time, or over a specific subset of observations, the associations may be dramatically different and the conclusions could be reversed.[38] So the terms "redundant", "multidimensional" and "robust" should not be associated with a given composite index, but rather jointly to the index and a specific dataset. In addition, once a robustness perspective is adopted, the parsimony or multidimensional arguments carry less force: if we replace the original variables with a single one, we lose all information on robustness, since a single variable always generates an unambiguous ranking.

There remains an interesting and unresolved tension between the need for a composite

[37] It is easy to demonstrate formally that the higher the associations between components on a composite index, the higher will be the correlation between the index itself and any one of its components.

[38] Suppose we are interested in the group of thirty least developed countries according to the HDI. The Kendall's tau rank correlation coefficients between the 2004 HDI and its three components are 0.18, 0.41, and 0.38, respectively, and the Kendall's tau coefficients between each pair of the three components are merely −0.31, −0.01, and 0.08. A similar pattern is found in other groups of interest.

index to improve upon unidimensional alternatives and the desire for the comparisons it makes to be robust. This question has implications for the choice of a specific variable to represent a given dimension in practice. Is it preferable to select a component variable that has low association with the other variables (to improve the multidimensional integrity of the index)? Or might it be better to seek out a variable that has high association with the others (to ensure more robust comparisons)? Further guidance on how to address this tension lies beyond the scope of the present chapter.

Conclusion

This chapter has analyzed the robustness of rankings obtained from composite indices – the multidimensional indices that combine information on two or more component indices using a weighted average. It examined the empirical prevalence of robust comparisons for three well-known and widely used indices: the Human Development Index, the Index of Economic Freedom and the Environmental Performance Index. The rank robustness of the Human Development Index was found to be the highest, with 73% of pair-wise 1998 country rankings of this index being fully robust. The Environmental Performance Index was the least robust, with no more than 6.5% of its pair-wise rankings being fully robust. The chapter then examined the link between various characteristics of the dataset and the prevalence of robust comparisons. One characteristic found to be relevant was the statistical association among index components, and many results were proved linking robustness and association. In particular, maximal robustness is obtained when components are perfectly positively associated. The chapter briefly touched upon a dilemma concerning the design of composite indices. According to the above results, highly positive associations among component variables are desirable as they can enhance rank robustness. But according to previous research, such associations are to be avoided on the grounds of redundancy. Should the design of a composite index be focused on rank robustness or on the avoidance of redundancy, or should we try to attain an optimal balance between the two? This question has been left to future research.

One further question raised by this chapter concerns the shape of prevalence functions and the implied empirical distribution of robust comparisons. It is evident that for both years, the HDI prevalence functions are approximately linear (more precisely, affine), as is the function associated with EPI2. The other prevalence paths have a strictly convex shape. A question is: what is it about the former composite indices and their datasets that produce a linear form? Linearity ensures that, if consideration is restricted to comparisons that are not fully robust, the empirical distribution of robustness levels is approximately uniform. In other words, the robustness level r is also the share of these comparisons having a robustness of r or below, and the share of comparisons having, say, $r = 0.95$ or above is $1 - r = 0.05$. This is certainly a notable regularity, and it would be useful to identify its source. Additional structure on the nature of this association, such as is available with a copula, may be helpful in this regard.

Appendix

E. Proof of Theorem 11

Proof. Suppose that y is a replicated achievement vector associated with x, so that $y = f(x)$ for a replicating transformation f. Given the initial weighting vector w^0 and a consistent weighting vector u^0, it is clear that $C(y; u^0) = u^0 \cdot f(x) = w^0 \cdot x = C(x; w^0)$. Now, let $r \in (0, 1]$ and select any $d = 1, \dots, D$ along with an index value d' of one of its copies. Let v_d^r denote the dimension d vertex of the simplex S_r^{D-1} in $\mathbb{R}^D$ and let $v_{d'}^r$ denote the dimension d' vertex of the simplex $S_r^{D'-1}$ in $\mathbb{R}^{D'}$. It is clear that $C(x; v_{d'}^r) = v_{d'}^r \cdot x = (1-r)C(x; w^0) + rx_d = (1-r)C(y; u^0) + ry_{d'} = v_{d'}^r \cdot y = C(y; v_{d'}^r)$. Hence, where y' and y are the respective transformations of x' and x, we have (i) $C(x'; w^0) \geq C(x; w^0)$ if and only if $C(y'; u^0) \geq C(y; u^0)$, and (ii) $C(x'; v_d^r) \geq C(x; v_d^r)$ if and only if $C(y'; v_{d'}^r) \geq C(y; v_{d'}^r)$. Since (ii) holds for each d and every associated d', it follows from Theorem 8 in Chapter III that $x' \ \mathbf{C}_r \ x$ if and only if $y' \ \mathbf{C}_r \ y$, and $p(r)$ is the same for both. ■

F. Proof of Theorem 14

Proof. Fix the initial vector w^0 and let $\hat{Y}$ be obtained from $\hat{X}$ by a single association increasing rearrangement involving x, x', y, and y' as defined in (a)-(c) above. If we can show that $p(1)$ rises, then we are done. To do this, we need only focus on comparisons involving at least one of the vectors x and x' in $\hat{X}$, since the remaining vectors are unchanged. Consider first the comparison involving both x and x'. By (b) we know that neither $x \geq x'$ nor $x' \geq x$ holds, and hence by Theorem 7, neither $x \ \mathbf{C}_1 \ x'$ nor $x' \ \mathbf{C}_1 \ x$ can be true. However, by construction $y > y'$ and since, by assumption, no achievements in any given dimension of x

and x' can be equal, we must have $y >> y'$. Again, by Theorem 7, it follows that y $\mathbf{C}_1$ y' holds, which represents a gain of one comparison for $\hat{Y}$ as compared to $\hat{X}$.

Now consider a case-by-case analysis of comparisons involving vectors x and x' and any given unchanged vector x''. (i) Suppose that x'' can be compared to both x and x' using $\mathbf{C}_1$. The case where x $\mathbf{C}_1$ x'' and x'' $\mathbf{C}_1$ x' simultaneously hold is impossible, since it implies $x \geq x'$ in contradiction to (a). Similarly the case where x' $\mathbf{C}_1$ x'' and x'' $\mathbf{C}_1$ x both apply contradicts $x' \geq x$, and is likewise impossible. On the other hand, if x'' $\mathbf{C}_1$ x and x'' $\mathbf{C}_1$ x' hold, then $x'' \geq x$ and $x'' \geq x'$ must both be true, and hence $x'' >> x$ and $x'' >> x'$ since no two vectors in $\hat{X}$ can have equal entries in a given dimension. By construction, then, $y'' >> y$ and $y'' >> y'$, which yields y'' $\mathbf{C}_1$ y and y'' $\mathbf{C}_1$ y', by Theorem 7. Similarly, x' $\mathbf{C}_1$ x'' and x $\mathbf{C}_1$ x'' yields y' $\mathbf{C}_1$ y'' and y $\mathbf{C}_1$ y'', and so in all possible cases, y'' can be compared to both of y and y' using $\mathbf{C}_1$. Clearly, $\hat{X}$ and $\hat{Y}$ have the same number of fully robust comparisons of this type. (ii) Suppose that x'' can be compared to exactly one of x and x' using $\mathbf{C}_1$. If the comparison is x $\mathbf{C}_1$ x'', then $x >> x''$ and hence by construction $y >> y''$, which implies y $\mathbf{C}_1$ y''. In a similar fashion, if the comparison is x' $\mathbf{C}_1$ x'', then we also conclude that y $\mathbf{C}_1$ y''. Alternatively, if the comparison is x'' $\mathbf{C}_1$ x, then $x'' >> x$ and hence by construction $y'' >> y'$, which implies y'' $\mathbf{C}_1$ y'. By the same argument, if the comparison is x'' $\mathbf{C}_1$ x', then we conclude y'' $\mathbf{C}_1$ y' once again. So in each circumstance, y'' can be compared to at least one of y and y' using $\mathbf{C}_1$ and hence $\hat{Y}$ has at least as many fully robust comparisons of this type as $\hat{X}$. (iii) Suppose that x'' can be compared to neither of x and x' using $\mathbf{C}_1$. Then, trivially, $\hat{Y}$ has at least as many fully robust comparisons of this type as $\hat{X}$. Consequently, the number of fully robust comparisons across cases (i) to (iii) is at least as high for $\hat{Y}$ as for $\hat{X}$; and given the original single comparison gain by $\hat{Y}$ over $\hat{X}$, it follows that $p(1)$ must be strictly higher for $\hat{Y}$ than for $\hat{X}$. ■

Chapter V

Multidimensional Welfare: An Indian Experience

Introduction

In Chapter II, a new class of multidimensional welfare indices has been introduced. The class is based on generalized means and has two parameters. Under appropriate restrictions of parameters, subclasses of this class are sensitive to two forms multidimensional inequality. One form of inequality is concerned with the spread of the marginal distributions of attributes and the other is concerned with the association across the attributes. It is discussed how a multidimensional social welfare index that is sensitive to these two forms of inequality can influence policy recommendations. Because the indices in this class are based on general means, they are amenable to empirical applications and statistical tests can be easily developed.

This chapter has two-fold objectives. The first is to show how this newly developed multidimensional index in Chapter II can be applied for evaluating social welfare in the context of developing countries, where social welfare is mostly gauged by a single dimension — usually income. However, an increase in income in these countries may not necessarily result in improvements in other attributes of well-being due to different forms of market failures and the absense of highly competent governance. The second objective is to show how the welfare evaluations are altered when the welfare indices are subjected to sensitivity to inequality.

For our purpose in this chapter, we choose India showing how the welfare evaluations across Indian states are altered when both multidimensionality and inter-personal inequality is incorporated into the social welfare evaluation. Like other developing countries, in India, the social welfare has been predominantly measured by per-capita income. There has been almost a three-fold increase in the national per-capita gross domestic product between 1990-91 and 2007-08. At the same time, however, the national family health survey and the human development report reveal that more than fifty percent of the rural women are illiterate, fifty seven infants do not survive out of every thousand newborns, nearly ninety percent of the rural households use solid biomass fuel for cooking purposes, and sixty seven percent of the population live without improved sanitation facilities as mandated in the millennium development goals (MDG) by the UNDP. Clearly, an increase in one attribute, such as income, does not necessarily lead to improvement in other attributes of well-being.

The second reason why India is found to be appropriate is that the inequality of achievement in different attributes remain high across the population and across various population sub-groups, such as, across regions, across religions, and across castes and tribes. Although the attribute-specific averages explain the story partly, they ignore the existing inter-personal inequality.

Because this class of indices requires the attributes to be continuous, three attributes of well-being are carefully constructed using several indicators. Seth (2009) has applied this class to the Mexican context showing how the state rankings are altered when different forms of inequalities are considered. However, the health variable is not available at the household level and, therefore, Seth (2009) does not appropriately capture the inequality in health. This chapter, on the other hand, selects all three attributes in such a way that each of them captures well-being at the household level. Moreover, we use the demographic and health survey data set, which collects internationally comparable data for many other developing countries. This chapter also develops confidence intervals relevant to the survey to verify the statistical significance of the evaluations generated by the indices.

The rest of this chapter is organized as follows. The second section describes the class of multidimensional welfare indices and provide an outline of the data set used in this chapter. In the third section, the attributes and indicators are introduced based on which social welfare is evaluated. The fourth section is devoted towards developing the statistical properties of the indices. The fifth section discusses the results reflecting how the rankings across different population subgroups are altered as an inequality sensitive index is used as opposed to an index that is not sensitive to inequality at all. The final section concludes this chapter.

The Welfare Index and the Data

To begin with, we spend some time recapitulating the multidimensional social welfare index introduced in Chapter II and then describe the data set.

The Class of Welfare Indices

The new class of multidimensional index has two parameters. An index in this class first aggregates the achievements of each person to obtain an overal achievement score. Then in the second stage, these achievement scores are aggregated to obtain the social welfare index. Let the achievements of a society consisting of N individuals and D attributes of well-being be summarized by the matrix H, where the nd^{th} element of h_{nd} denotes the achievement of individual n in attribute d, for all $d = 1, \ldots, D$ and all $n = 1, \ldots, N$. Our social welfare index can be defined as:

$$W(H;\alpha,\beta,a) = \left(\frac{1}{N}\sum_{n=1}^{N}\left(\sum_{d=1}^{D} a_d h_{nd}^{\beta}\right)^{\alpha/\beta}\right)^{1/\alpha}, \quad (11)$$

where the parameter α measures society's aversion towards inter-personal inequality in these achievements, the parameter β measures the degree of substitutability across the attributes of well-being of any individual, and a is a D-dimensional weight vector such that $a_d > 0$ and $\sum_{d=1}^{D} a_d = 1$.[39] The d^{th} element in the weight vector a signifies the importance that is attached to the d^{th} attribute when measuring the overall achievement score for each individual. Note that in each stage of aggregation, an index in our class uses a generalized mean and thus the index is a generalized mean of generalized means.

It is shown in Chapter II that the class of indices defined in (11) is sensitive to two distinct forms of multidimensional inequality. Sensitivity to the first form of inequality requires that if the average achievement of each attribute remains unchanged but the distribution of each attribute becomes more dispersed, then the social welfare index should register a fall. The second form of sensitivity requires that if the marginal distribution of each attribute remains unaltered but the association across attributes increases, then the level of social welfare should fall, provided the attributes are substitutes to each other. By the attributes being substitutes, we mean that if an individual has lower achievement in one attribute and higher achievement in another, then the person can compensate for her lower achievement in the former attribute by her higher achievement in the latter. It is shown in Chapter II that the index registers a fall due to an increase in association among attributes if $\alpha < \beta$ and an increase due to an increase in association if $\alpha > \beta$. Similarly, the value of the index falls as

[39] For $\alpha = 0$ and $\beta = 0$, the corresponding geometric mean forms should be used.

the spread of the attributes increases if $\alpha < 1$ and $\beta < 1$.

The Data and the Unit of Analysis

For the analysis, we select three attributes that are often considered important while measuring social welfare: *material well-being*, *educational well-being*, and *quality of health.*[40] We use the third National Family Health Survey (NFHS-3) 2005-06 data set for the following two reasons. This dateset is chosen for its sound quality. Furthermore, it is a part of the Demographic Health Survey (DHS), which collects comparable data for many developing countries. It is possible to make cross-country welfare comparison in future applying the techniques introduced in Chapter II. The NFHS-3 collects information on a nationally representative sample of 109,041 households and covers 99 percent of the Indian population from twenty-eight states and the national capital territory of Delhi.[41]

In this survey, the entire country is divided into fifty-eight regions in terms of rural and urban areas, and then eight major cities are further divided into slum and non-slum areas. Overall, there are seventy-three regional divisions. Each region is assigned a nationally representative population weight equal to the projected population of households in that region divided by the total number of sample households drawn from that region. In our analysis, we assume that samples drawn from each region are independently and identically distributed. However, samples drawn from any two different regions are indeed independent but are not necessarily drawn from an identical distribution.

Like the previous two rounds (1992 and 1997), this survey does not collect any information

[40] Variations of these three attributes are commonly used when constructing the human development index (UNDP) and various physical quality of life indices (e.g. see Morris 1979).

[41] All three rounds of National Family Health Survey are coordinated by International Institute for Population Sciences(IIPS), Mumbai; ORC Macro; and the Ministry of Health and Family Welfare (MOHFW), Government of India.

on household income or consumption expenditure. Instead, it collects detailed information on household asset ownerships. The availability of information only at the household level does not allow an individual level analysis. It is implicitly assumed that all members in a household share the same level of material well-being. Hence, our unit of analysis in this chapter is the household.

Dimensions and Indicators

The choice of appropriate indicators for constructing the three chosen attributes is crucial. We now discuss the choice of indicators and their strengths and weaknesses.

Material Well-Being

The choice of a reasonable measure of material well-being is highly debatable. There are three possible alternatives to measure material well-being at the household level: income, consumption expenditure, and wealth or asset ownership summarized by an asset index score. An asset index score is a composite index of the set of assets owned by a household. Among these three indicators, income data are not often reliable as respondents tend to under-report their earnings and, for developing countries, there are not many reasonable sources to verify the response. Moreover, it is difficult to measure the income for self-employed and agricultural workers owing to non-accountability and seasonal issues. The consumption expenditure data, on the other hand, are free from these problems and thus are superior to income data. However, the asset index score has certain crucial advantages over the consumption expenditure. First, ownership of asset is supposed to be a better indicator of long run material well-being because it does not fluctuate frequently. Second, there is likely to be less recall bias or mis-measurement for assets.[42] Third, the time required to collect asset information is much shorter because the list of assets is generally much shorter than the list of commodities used for consumption (McKenzie, 2005). Nevertheless, an asset index may not be the best method for measuring short run material well-being, since it does not fluctuate much in the short run. Indeed, there are other challenges. The first is the selection of an appropriate set of assets while constructing the asset index score. The second challenge is to attach a reasonable weight to each asset. Third, there is a possibility of clumping, where a large proportion of households has the same score (Howe et al., 2008).

There are numerous ways of constructing an asset index score (Filmer and Scott, 2008). The most reasonable approach is to calculate the monetary value of the set of assets, but the set of prices and the depreciation costs are not readily available. The second reasonable approach is the regression based method (Stifel and Christiaensen, 2007), which calculates the weight on each asset by regressing the consumption expenditure on the set of assets. We can not pursue this approach because it requires availability of household level consumption expenditure data. There are also different statistical procedures such as using principal component analysis, factor analysis, and multiple correspondence analysis to determine the weights on assets. However, the choice of weights generated by these latter methods are not intuitive and are thus often difficult to justify. We, therefore, choose to pursue an approach that is simple and intuitive.

There are two possible alternatives. One is simply counting the number of assets a household owns, but this is difficult to justify: for example, a watch is certainly not as expensive as a car. The second simple approach is to weight the assets so that the weights

[42] Recall bias occur when a survey respondent's answer is affected by the memory of the respondent.

reflect some value adjustments. For this purpose, we assume that if an asset is more valuable then it is less likely to be held by the population. Thus, the weight attached to the asset should be proportional to the share of the population not holding it (Morris et al., 2000). This way we meet the second challenge while constructing an asset index.

Let us denote the set of the K assets by $X = (X_1, \ldots, X_K)$. The set of assets owned by household $n \in \mathbf{N}$ is denoted by $X^n = (X_1^n, \ldots, X_K^n)$, where $X_k^n = 1$ if household n owns asset k and $X_k^n = 0$ otherwise, for all $k = 1, \ldots, K$. Let us denote the share of the population that does not own asset k by g_k. Then the asset index score for household n can be represented by the following formula:

$$A_n = \frac{\sum_{k=1}^{K} \phi(g_k) X_k^n}{\sum_{k=1}^{K} \phi(g_k)}, \tag{12}$$

where $\phi' > 0$. If $\phi(g_k) = g_k$, then the weight on asset k is linear in g_k. Note that we are interested in the comparisons across population subgroups at the national level, and so the weights are based on the national level coverage of the assets.

We choose a set of twenty assets from the NFHS data set. When choosing the assets, we take into account the following two aspects. First, the set of assets should be comparable across various population sub-groups. Second, our approach in (12) requires every asset to be dichotomous. The chosen set of assets consists of electricity connection, refrigerator, bicycle, motor cycle, car, phone, cell phone, watch, bank account, mattress, pressure cooker, chair, cot or bed, table, electric fan, color TV, black & white TV or radio or transistor, sewing machine, computer, and water pump.[43]

We report the national non-coverage rate of the assets and the two sets of weights in

[43] We do not include any information on housing because this indicator is not dichotomous. Several other asset indicators such as a thresher, a tractor, or an animal-drawn cart are not included because they are rural based indicators in general, and thus can lead to erroneous conclusions during rural versus urban comparison.

Table 6: Weights Attached to the Selected Assets

Assets	**Non-Coverage Rate**	**Linear Weight** $\phi(g_k) = g_k$	**Squared Weight** $\phi(g_k) = g_k^2$
Electrification	0.32	0.03	0.01
Refrigerator	0.85	0.07	0.08
Bicycle	0.49	0.04	0.03
Motorcycle or scooter	0.83	0.07	0.07
Car	0.97	0.08	0.10
Telephone	0.86	0.07	0.08
Mobile telephone	0.83	0.07	0.08
Watch or clock	0.22	0.02	0.01
Own a bank or post-office account	0.59	0.05	0.04
Mattress	0.43	0.03	0.02
Pressure cooker	0.62	0.05	0.04
Chair	0.46	0.04	0.02
Cot/bed	0.17	0.01	0.00
Table	0.57	0.04	0.04
Electric fan	0.46	0.04	0.02
Color television	0.75	0.06	0.06
Sewing machine	0.82	0.06	0.07
Computer	0.97	0.08	0.10
Water pump	0.90	0.07	0.09
Black and White TV/Radio/transistor	0.56	0.04	0.03

Table 6. For the first set of weights, $\phi(g_k) = g_k$, whereas, for the second, $\phi(g_k) = g_k^2$. Morris et al. (2000) used the first set of weights. However, in our context, the first set of weights seems bit unrealistic as it does not allow enough variation in the weights across the assets. The second set of weights attaches comparatively more weight on the rarer assets. We have found that the Spearman's rank correlation coefficient between both asset index scores at the individual level is almost one. We have also calculated the Spearman's rank correlation coefficient between the asset index scores for the second set of weights and the wealth index in the NFHS-3 data set that uses the principal component analysis. The resulting coefficient is 0.92. Thus, our choice of the second set of weights is robust to the choice of the other weights. The clumping problem is automatically taken care off as we choose a large number

of assets.

Educational Well-Being

The second attribute for evaluating welfare is the educational well-being of the households. We assume that educational well-being increases with the presence of household members who have completed more years of education. Other members in the households indeed benefit from their presence (Basu and Foster, 1998). The following choices are available in the survey as an indicator for this attribute: the first obvious choice is to use the years of education completed by the head of the household; the second is to use the maximum level of education completed by any member in the household; and a third potential choice is to use the average years of education completed by the adult members in the household. Each of these choices has its pros and cons. Given that nearly forty percent of the household heads in India is fifty years or older, the first indicator would under-estimate the recent improvement in the knowledge base of the country as a whole. The problem with the second indicator is that it does not distinguish between the level of knowledge of two households with the same number of members. It provides an identical score to a five-member household with only one person having completed fifteen years of education, to another household in which all five members have completed fifteen years of education.

It is apparent that the third indicator is the best choice among these three indicators, and we measure the material well-being of a household by the average years of education completed by the adult members in that household.[44] For household n, let us denote the set of $M_n \in \mathbb{N}$ members by $\mathbf{M}_n = \{1, \ldots, M_n\}$. If the years of education completed by member

[44] Indeed, the indicator suffers from two limitations. The first is that it ignores the level of education completed by the non-adult members. The second is that it does not consider the qualitative differences among education standards. The former limitation requires further research, but the latter is a data limitation.

m is denoted by Y_m, then the educational well-being of the household is denoted by

$$E_n = \frac{1}{22 \times M_n} \sum_{m=1}^{M_n} Y_m. \tag{13}$$

We divide the right hand side by 22, which is the highest year of education possible, so that the index lies between zero and one, like the asset index.

Quality of Health

The third attribute chosen is the quality of health of the household. It is difficult to find an appropriate indicator for measuring the health quality of an entire household. The survey contains few direct health indicators such as the body mass index (BMI) and level of anaemia for the respondents, and stunting and wasting for the children. These indicators are, however, not sufficient to capture the health of an entire household. Moreover, they are not monotonically increasing with the quality of health. In other words, a higher BMI of a person does not necessarily imply a better quality of health for that person. The alternative is to construct a proxy index that captures the risk to the health of the entire household. The higher is the value of the index, the better should be the quality of health of a household. We construct a health risk index combining four sub-attributes that consist of the following six indicators: (i) the access to safe drinking water, (ii) the availability of an improved toilet facility, (iii) the number of people the toilet is shared with, (iv) the type of fuel used for cooking, (v) the availability of a separate cooking room that is not used for sleeping, and (vi) the number of persons per sleeping room. Each of these indicators contribute in different ways towards the quality of health of the members in the household. For further discussion, see Mishra et al. (1999), World Health Organization (2000), Beggs and Siciliano (2001), Bartram et al. (2005), and Rehfuess et al. (2009). Note that these sub-attributes and indicators are also included among the Millennium development goals (United Nations, 2003).

The first sub-attribute is safe drinking water. This indicator identifies whether a household has access to safe drinking water and is denoted by $l_1^n = 1$ if household n has access, and $l_1^n = 0$, otherwise. The second sub-attribute is an improved sanitation facility, which is measured using two indicators: access to an improved toilet facility (l_2) and the number of other households the toilet is shared with (l_3). It is argued that shared toilet facilities can be less hygienic and can deter household members from using it (World Health Organization and UNICEF, 2006). For household n, $l_2^n = 1$ if the household has access to an improved toilet facility, and $l_2^n = 0$, otherwise; and $l_3^n = 1$ if the toilet is not shared with anyone, and if the toilet is shared with c other households, then $l_3^n = 1/c$. The third sub-attribute is indoor pollution, which is also measured using two indicators: the type of cooking fuel used as a source of energy for daily cooking purposes (l_4) and whether the household cooks in the same room used for sleeping (l_5). For household n, $l_4^n = 1$ if the household uses MDG mandated non-biomass fuel for cooking and $l_4^n = 0$, otherwise. Similarly, $l_5^n = 1$ if the household does not cook in the room used for sleeping and $l_5^n = 0$, otherwise. The final sub-attribute is crowding, which is measured by the number of persons per bedroom (l_6). This indicator is believed to be an important health risk indicator because transmission of communicable and respiratory diseases are higher in situations of crowding. For household n, $l_6^n = M_n/R_n$, where R_n is the number of rooms used for sleeping. Because it is difficult to justify normatively which sub-attribute is of greater importance in reducing health risks, each of the sub-attributes is weighted equally. Thus, the health risk score for household n can be written as

$$L_n = \frac{1}{4} l_1^n + \frac{1}{4} \left(\frac{l_2^n + l_3^n}{2} \right) + \frac{1}{4} \left(\frac{l_4^n + l_5^n}{2} \right) + \frac{1}{4} l_6^n. \tag{14}$$

For all $n \in \mathbf{N}$, the overall achievement score for household n is calculated as $Q(h_{n \cdot}) = \mu_\beta^3 (A_n, E_n, L_n; \xi_3)$.

Statistical Properties of the Attributes

Before we apply these three attributes to construct multidimensional welfare indices, we analyze the distribution of each dimensional achievements across the households by estimating a Gaussian kernel density for each attribute. For any $N \in \mathbb{N}$, let $y = (y_1, \ldots, y_N) \in \mathbb{R}_{++}^N$ denote a distribution of achievements. Following Gisbert (2003), the kernel density estimate is defined as

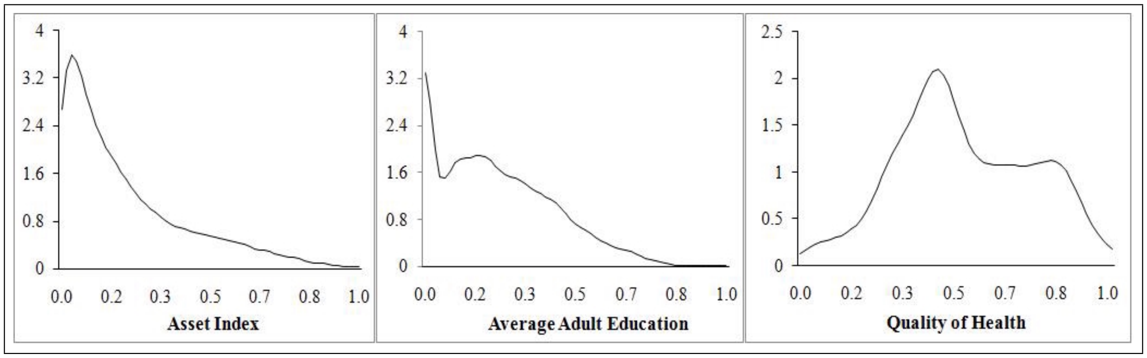
$$\hat{f}(\hat{y}) = \frac{1}{\mathcal{B}} \sum_{n=1}^{N} w_n K \left(\frac{y_n - \hat{y}}{\mathcal{B}} \right)$$

for any $\hat{y}$ in $\mathbb{R}$, where $w_n \geq 0$ and $\sum_{n=1}^{N} w_n = 1$, K is the density of the standard normal distribution. The bandwidth $\mathcal{B}$ is calculated as $\mathcal{B} = 0.9 N^{1/\kappa} \times \min \{\mathrm{sd}(y), \mathrm{iqr}(y)\}$, where κ < 0, and $\mathrm{sd}(y)$ and $\mathrm{iqr}(y)$ denote the standard deviation and the interquartile range of y, respectively.

In Figure 5, we depict the kernel density estimates of the three selected attributes.[45] Although the kernel density for the asset index scores is unimodal, the two other kernel densities are not. There is a large number of Indian households in which none of the adult members has finished even one year of education. That explains the existence of the left

[45] To ensure sufficient smoothing of the kernel density estimates, the values of κ are assumed to be -6, -6, and -9 for the asset index, average adult education index, and the index of health risk, respectively.

Figure 5: Kernel Density Estimate of the Three Attributes



mode of the kernel density for average adult education, in addition to the usual mode on the further right. The health risk score also has two modes. Note that the index of health risk is a composite index of both continuous and dichotomous variables and, unlike the asset index, there is not a large number of indicators. Thus, clumping is a possibility. Note that the distribution of both the asset index scores and the average education scores is right skewed, as expected. However, the distribution of the health risk score is not. The possible reason is that the majority of the Indian households have access to safe drinking water and this indicator receives 25 percent weight. Possibly, there is a group of households with access to safe water and with high performance in other sub-attributes; and another group of households with access to safe drinking water who do not perform well in other sub-attributes.

Stastistical Tests

In this section, the asymptotic properties of the class of two-parameter generalized mean social welfare indices are analyzed and relevant statistical tests are developed in order to verify the statistical significance of the estimates. In many sample surveys, data are collected in two stages. First, the entire region of interest is divided into a fixed number of strata and then random samples are drawn from each of these stratum. Each stratum corresponds to a particular joint distribution of achievements across the population. Thus, samples drawn from the same stratum are independently and identically distributed. However, random draws from any two different strata are independent but are not identically distributed.

Suppose, there are M such strata and the multivariate distribution corresponding to stratum m is denoted by F_m for all $m = 1, \ldots, M$. Let a total number of N_m samples are drawn from stratum m and $N = \sum_{m=1}^{M} N_m$ be the total sample size. Subscript n_m represents the n^{th} sample drawn from the m^{th} strata. Samples $\{h_{n_m 1}, \ldots, h_{n_m D}\}_{n_m=1}^{N_m} \sim F_m$ receive a population representation weight of $w_m = P_m/P$, where P_m stands for the population size in stratum m and $P = \sum_{m=1}^{M} P_M$ is the total population size. We assume that as $N \to \infty$, $N_m \to \infty$ but $N_m/N \to w_m$ for all m. This is a restriction on the sampling design. We assume that as the total sample size increases, then the sample size in each strata also increases. The population version of the social welfare index can be written as

$$\theta = \left(\sum_{m=1}^{M} w_m \left(\int \left(\mu_\beta^D \left(h_{n_m \cdot}; a\right)\right)^\alpha dF_m\right)\right)^{\frac{1}{\alpha}} = \left(\sum_{m=1}^{M} w_m \pi_m\right)^{\frac{1}{\alpha}} = \pi^{\frac{1}{\alpha}}, \tag{15}$$

where $\pi_m = \int (\mu_\beta^D \left(h_{n_m \cdot}; a\right))^\alpha dF_m$ for all m, and $\pi = \sum_{m=1}^{M} w_m \pi_m$.[46]

Therefore, for a given achievement matrix $H \in \mathcal{H}$, the statistic of interest can be written

[46] In terms of the discrete variables, the population version of the social welfare index can be written as $\theta = (\frac{1}{P} \sum_{m=1}^{M} \sum_{n_m=1}^{P_m} (\mu_\beta^D (h_{n_m \cdot}; a))^\alpha)^{1/\alpha}$.

as

$$\hat{\theta} = \left(\sum\nolimits_{m=1}^{M} \frac{w_m}{N_m} \sum\nolimits_{n_m=1}^{N_m} \left(\mu_\beta^D \left(h_{n_m}.; a\right)\right)^\alpha\right)^{\frac{1}{\alpha}} = \left(\sum\nolimits_{m=1}^{M} w_m \left(\frac{1}{N_m} \sum\nolimits_{n_m=1}^{N_m} \eta_{n_m}\right)\right)^{\frac{1}{\alpha}},$$

where $\eta_{n_m} = (\mu_\beta^D (h_{n_m}.; a))^\alpha$ for all m and $n_m = 1, \ldots, N_m$. Then $\hat{\pi}_m = \frac{1}{N_m} \sum_{n_m=1}^{N_m} \eta_{n_m}$ for all m, and $\hat{\pi} = \sum_{m=1}^{M} w_m \hat{\pi}_m$.[47]

By the Weak Law of Large Number, as $N \to \infty$, then $\hat{\pi}_m \xrightarrow{p} E(\eta_{n_m})$ for all m and also $\hat{\pi} \xrightarrow{p} \sum_{m=1}^{M} w_m E(\eta_{n_m})$. Thus, by the continuous mapping theorem, $\hat{\theta} \xrightarrow{p} (\sum_{m=1}^{M} w_m (E\eta_{n_m}))^{1/\alpha}$.

We assume that the variance of η_{n_m} be finite and from the restrictions on weights, it clearly follows that $w_m/w_{m'} < \infty$ for all $m \neq m'$. Then, applying the theorem in Eremin (1999, p. 1012), we have

$$\sqrt{N} \left(\hat{\pi} - \pi\right) \xrightarrow{D} \text{Normal} \left(0, \sum\nolimits_{m=1}^{M} \frac{N}{N_m} w_m^2 \sigma_{\eta_m}^2\right),$$

where $\sigma_{\eta_m}^2$ is the variance of η_{n_m} for all m.

By applying the Delta method, and replacing $\pi = \theta^\alpha$ from (15), we have

$$\sqrt{N}(\hat{\theta} - \theta) \xrightarrow{D} \text{Normal} \left(0, \frac{\theta^{2(1-\alpha)}}{\alpha^2} \sum\nolimits_{m=1}^{M} \frac{N}{N_m} w_m^2 \sigma_{\eta_m}^2\right).$$

We first consistently estimate $\sigma_{\eta_m}^2$ for all m as

$$\hat{\sigma}_{\eta_m}^2 = \widehat{\text{var}} \left(\eta_{n_m}\right) = \frac{1}{N_m - 1} \sum\nolimits_{n_m=1}^{N_m} \left(\eta_{n_m} - \hat{\pi}_m\right)^2.$$

[47] This framework is analogous to the framework in Bickel and Freedman (1984) and Eremin (1999).

Therefore, a consistent estimator of σ_θ^2 is

$$\hat{\sigma}_\theta^2 = \frac{\hat{\theta}^{2(1-\alpha)}}{\alpha^2} \sum\nolimits_{m=1}^{M} \left(\frac{w_m^2 N}{N_m (N_m - 1)} \sum\nolimits_{n_m=1}^{N_m} \left(\eta_{n_m} - \hat{\pi}_m\right)^2 \right).$$

Hence, the *standard error* of the estimate is

$$SE(\hat{\theta}) = \frac{\hat{\sigma}_\theta}{\sqrt{N}} = \frac{\hat{\theta}^{1-\alpha}}{|\alpha|} \sqrt{\sum\nolimits_{m=1}^{M} \frac{w_m^2}{N_m (N_m - 1)} \left(\sum\nolimits_{n_m=1}^{N_m} \left(\eta_{n_m} - \hat{\pi}_m\right)^2 \right)}. \tag{16}$$

Finally, we calculate the *confidence interval* for the statistic $\hat{\theta}$. In this situation, both the mean and the variance are unknown and so the test statistic is equal to

$$T = \frac{\hat{\theta} - \theta}{SE(\hat{\theta})} = \frac{\sqrt{N}(\hat{\theta} - \theta)}{\hat{\sigma}_\theta}.$$

We know that if $\sqrt{N}(\hat{\theta} - \theta) \xrightarrow{D} \text{Normal}\,(0, \sigma_\theta^2)$ and $\hat{\sigma}_\theta^2 \xrightarrow{p} \sigma_\theta^2$, then $T \xrightarrow{D} \text{Normal}\,(0, 1)$ (Bierens, 2004, Theorem 6.21). Hence, the confidence interval of θ is given by

$$\hat{\theta} - z_\delta SE(\hat{\theta}) \leq \theta \leq \hat{\theta} + z_\delta SE(\hat{\theta}), \tag{17}$$

where z is a standard normal distribution, z_δ is the critical value with confidence level of $(1 - \delta)\,\%$, and the formulation of $SE(\hat{\theta})$ follows from (16). [48]

[48] Under certain circumstances, it is assumed that the entire sample is drawn independently from an identical distribution. Our method applies here as well, because it is identical to the situation for $M = 1$ and thus $\hat{\theta} = (\frac{1}{N} \sum_{n=1}^{N} (\mu_\beta^D(h_{n\cdot}; a))^\alpha)^{1/\alpha}$, where N is the sample size. The corresponding standard error can be estimated as $SE(\hat{\theta}) = \frac{1}{|\alpha|}(\hat{\theta})^{1-\alpha} \sqrt{(\sum_{n=1}^{N} (\eta_n - \hat{\pi})^2 / N(N-1))}$, and the confidence interval can be estimated using (17).

Results

Table 7 summarizes the average achievement scores in three attributes across various population subgroups. The entire sample is divided into four different types of sub-groups: by states, by religions, by castes and tribes, and by rural and urban areas.[49] The population of the eight cities are further divided into slum and non-slum areas.[50]

Among the twenty-eight Indian states, the people of Goa enjoy the highest level of material well-being; whereas, the people of Kerala enjoy both the highest level of educational well-being and the best quality of health. The citizens of the capital territory of Delhi have higher average achievement compared to the citizens of any other Indian state in all three attributes. On the other hand, the people of Bihar and Jharkhand, the two neighboring states, share the lowest level of well-being in all three attributes. Note that a state that has higher well-being in one attribute does not necessarily have higher well-being in other attributes. For example, consider Rajasthan, which ranks 17th in terms of material well-being but ranks 26th in terms of educational well-being. The people of West Bengal are not as well-off in terms of material well-being as they are in terms of health risk.

In terms of religion, the entire sample is divided into six major groups: Hindu, Muslim, Christian, Sikh, Buddhist/neo-Buddhist, and others.[51] The dominant religious group in India is Hindu, covering almost eighty percent of the Indian population. It is evident from Table 7 that the Sikhs, who are primarily residents of Punjab, score highest in terms of

[49] The sample size for the analyses across states, religions, and rural and urban areas is 106,674. The sample size for comparison across castes and tribes is 105,818. Because the social welfare index is based on general means, we replace the score of zero by a marginal positive number.

[50] The eight large cities for which the slum/nonslum data are available are Chennai, Delhi, Hyderabad, Indore, Kolkata, Meerut, Mumbai, and Nagpur. The corresponding sample size that we use for our analysis is 18,168.

[51] The 'others' group consists of people from various religions such as Jain, Parsi/Zoroastrian, Jewish, Donyi polo, etc. and of course the group of people with no religion.

both asset ownership and quality of health; whereas, the educational well-being is highest for Christians. Although the difference in material well-being and risk to health is marginal between Hindus and Muslims, Hindus enjoy a substantially higher level of well-being in terms of education.

The caste system is highly prominent in Indian society and we also decompose the Indian population in terms of Scheduled Castes, Scheduled Tribes, Other Backward Classes (OBCs), none of these three classes (the upper-caste people in particular), and people belonging to neither any caste nor any tribe. It can be seen from Table 7 that the group of people who are classified as Scheduled Tribes have the lowest achievements in all three attributes. This picture contrasts to the group of people belonging the 'none of the above' class. Unlike the other two population sub-groups above, the dimensional rankings are mostly robust in this case. A further subgroup decomposition in terms of rural and urban areas shows that the people living in urban areas enjoy twice as much educational and material well-being as their rural counterparts. The rural population also suffer from much higher health risks. Note that the rural area consists of more than two-third of the total Indian households

In Table 8, we report two welfare indices for every population sub-groups. In the second column, we report the social welfare index $W(\cdot; 1, 1, a)$, which is the simple average of average dimensional achievements. By construction, this index is not sensitive to either of the two forms of inter-personal inequality. The ranks of the population subgroups are provided in the parentheses. The third column reports the confidence interval of the reported level of social welfare using the statistical tests developed in the previous section. The fourth column reports the value of a social welfare index that is sensitive to both forms of multidimensional inequality. We assume the level of inequality aversion to be $\alpha = -0.5$ and the degree

Table 7: Achievements across Various Population Subgroups

States	**Pop Share**	**Asset Index** $\phi(g_k) = g_k$	**Rank**	**Average Adult Educ.**	**Rank**	**Quality of Health**	**Rank**
Andhra Pradesh	8.7%	0.20	19	4.4	24	0.56	18
Arunachal Pradesh	0.1%	0.20	18	4.8	21	0.59	11
Assam	2.7%	0.19	23	5.4	18	0.57	16
Bihar	7.0%	0.14	29	3.4	29	0.48	23
Chhattisgarh	2.1%	0.15	26	4.1	27	0.44	26
Delhi	1.2%	0.47	1	8.8	1	0.73	1
Goa	0.2%	0.45	2	7.8	3	0.68	4
Gujarat	4.9%	0.28	10	5.8	12	0.59	12
Haryana	1.8%	0.32	7	5.8	13	0.58	15
Himachal Pradesh	0.6%	0.34	5	7.1	5	0.60	9
Jammu and Kashmir	0.7%	0.30	9	5.8	11	0.56	20
Jharkhand	2.5%	0.15	27	4.0	28	0.36	29
Karnataka	5.8%	0.23	14	5.7	15	0.55	21
Kerala	2.8%	0.38	4	8.3	2	0.69	2
Madhya Pradesh	6.2%	0.18	24	4.5	22	0.42	28
Maharashtra	9.6%	0.27	11	6.6	7	0.58	14
Manipur	0.2%	0.26	12	7.1	6	0.55	22
Meghalaya	0.3%	0.19	22	5.1	19	0.56	19
Mizoram	0.1%	0.32	6	7.2	4	0.67	5
Nagaland	0.2%	0.21	16	5.7	14	0.60	10
Orissa	3.8%	0.15	28	4.3	25	0.43	27
Punjab	2.3%	0.41	3	6.1	9	0.66	6
Rajasthan	5.4%	0.21	17	4.1	26	0.47	25
Sikkim	0.1%	0.25	13	5.6	16	0.69	3
Tamil Nadu	7.6%	0.21	15	6.0	10	0.57	17
Tripura	0.4%	0.19	21	5.5	17	0.63	7
Uttar Pradesh	13.8%	0.19	20	4.4	23	0.48	24
Uttarakhand	0.8%	0.30	8	6.5	8	0.61	8
West Bengal	8.3%	0.18	25	4.9	20	0.59	13
Religion							
Hindu	81.7%	0.21	4	5.2	5	0.53	6
Muslim	12.4%	0.20	5	4.0	6	0.55	4
Christian	2.7%	0.28	3	7.1	1	0.61	2
Sikh	1.6%	0.46	1	6.3	3	0.68	1
Buddhist/neo-buddh.	0.8%	0.19	6	5.6	4	0.53	5
Others	0.7%	0.33	2	6.8	2	0.57	3
Caste/Tribe							
Scheduled castes	19.4%	0.15	4	3.9	4	0.48	4
Scheduled tribes	8.6%	0.11	5	3.0	5	0.39	5
OBCs	40.0%	0.20	3	4.8	3	0.51	3
None of above	29.8%	0.32	1	7.0	1	0.64	1
No Caste or Tribe	2.2%	0.21	2	5.1	2	0.60	2
Rural/Urban							
Urban	32.5%	0.35	-	7.6	-	0.69	-
Rural	67.5%	0.15	-	3.9	-	0.46	-
Slum\Non Slum							
Slum	37.4%	0.31	-	7.0	-	0.65	-
Non Slum	62.6%	0.48	-	9.6	-	0.77	-
India	**100%**	**0.23**	**-**	**5.1**	**-**	**0.51**	**-**

Table 8: Association Sensitive Welfare Indices with $a = (1/3, 1/3, 1/3)$

States	$W(\cdot,1,1,a)$		**Confidence Interval** (99%)	$W(\cdot,\text{-}0.5,0.1,a)$		**Confidence Interval** (99%)	$I(\cdot,\text{-}0.5,0.1,a)$
Andhra Pradesh	0.320	(22)	(0.314, 0.325)	0.135	(22)	(0.130, 0.141)	0.52
Arunachal Pradesh	0.337	(18)	(0.326, 0.348)	0.139	(20)	(0.129, 0.150)	0.54
Assam	0.333	(19)	(0.326, 0.340)	0.153	(18)	(0.145, 0.161)	0.49
Bihar	0.259	(27)	(0.252, 0.265)	0.099	(27)	(0.094, 0.103)	0.56
Chhattisgarh	0.259	(26)	(0.253, 0.265)	0.100	(26)	(0.095, 0.106)	0.57
Delhi	0.533	(1)	(0.525, 0.541)	0.369	(2)	(0.355, 0.383)	0.29
Goa	0.494	(2)	(0.485, 0.502)	0.324	(3)	(0.309, 0.339)	0.32
Gujarat	0.377	(12)	(0.370, 0.384)	0.201	(12)	(0.191, 0.211)	0.43
Haryana	0.386	(9)	(0.378, 0.394)	0.214	(11)	(0.202, 0.226)	0.42
Himachal Pradesh	0.424	(6)	(0.415, 0.432)	0.282	(5)	(0.267, 0.297)	0.31
Jammu and Kashmir	0.374	(13)	(0.365, 0.382)	0.216	(10)	(0.204, 0.227)	0.39
Jharkhand	0.233	(29)	(0.226, 0.239)	0.074	(29)	(0.070, 0.079)	0.66
Karnataka	0.346	(17)	(0.341, 0.352)	0.159	(17)	(0.153, 0.165)	0.50
Kerala	0.483	(3)	(0.476, 0.490)	0.373	(1)	(0.361, 0.385)	0.20
Madhya Pradesh	0.267	(25)	(0.261, 0.274)	0.098	(28)	(0.094, 0.103)	0.61
Maharashtra	0.383	(10)	(0.378, 0.389)	0.200	(13)	(0.192, 0.208)	0.45
Manipur	0.379	(11)	(0.372, 0.386)	0.249	(7)	(0.239, 0.259)	0.31
Meghalaya	0.326	(21)	(0.317, 0.335)	0.144	(19)	(0.133, 0.154)	0.51
Mizoram	0.438	(5)	(0.429, 0.447)	0.310	(4)	(0.293, 0.327)	0.25
Nagaland	0.354	(15)	(0.347, 0.360)	0.183	(15)	(0.174, 0.191)	0.43
Orissa	0.259	(28)	(0.253, 0.265)	0.102	(25)	(0.097, 0.107)	0.57
Punjab	0.450	(4)	(0.442, 0.457)	0.279	(6)	(0.267, 0.291)	0.34
Rajasthan	0.290	(24)	(0.284, 0.296)	0.102	(24)	(0.096, 0.107)	0.62
Sikkim	0.395	(8)	(0.386, 0.405)	0.232	(8)	(0.219, 0.246)	0.34
Tamil Nadu	0.350	(16)	(0.344, 0.356)	0.170	(16)	(0.163, 0.176)	0.47
Tripura	0.356	(14)	(0.347, 0.364)	0.192	(14)	(0.178, 0.207)	0.39
Uttar Pradesh	0.291	(23)	(0.287, 0.295)	0.122	(23)	(0.119, 0.126)	0.54
Uttarakhand	0.403	(7)	(0.394, 0.412)	0.219	(9)	(0.207, 0.231)	0.43
West Bengal	0.328	(20)	(0.323, 0.334)	0.138	(21)	(0.132, 0.144)	0.52
Religion							
Hindu	0.325	(4)	(0.324, 0.327)	0.140	(4)	(0.138, 0.142)	0.54
Muslim	0.310	(6)	(0.306, 0.313)	0.125	(6)	(0.121, 0.128)	0.54
Christian	0.405	(2)	(0.397, 0.413)	0.202	(2)	(0.190, 0.215)	0.47
Sikh	0.473	(1)	(0.465, 0.482)	0.298	(1)	(0.283, 0.314)	0.33
Buddhist/neo-Buddhist	0.325	(5)	(0.310, 0.340)	0.159	(3)	(0.140, 0.179)	0.47
Others	0.403	(3)	(0.392, 0.414)	0.131	(5)	(0.117, 0.145)	0.66
Caste/Tribe							
Scheduled castes	0.269	(4)	(0.266, 0.272)	0.109	(4)	(0.107, 0.112)	0.54
Scheduled tribes	0.209	(5)	(0.205, 0.212)	0.067	(5)	(0.065, 0.069)	0.63
OBCs	0.311	(3)	(0.308, 0.313)	0.139	(3)	(0.137, 0.141)	0.51
None of above	0.425	(1)	(0.422, 0.428)	0.232	(1)	(0.227, 0.237)	0.43
No Caste or Tribe	0.350	(2)	(0.343, 0.357)	0.156	(2)	(0.146, 0.167)	0.50
Rural/Urban							
Urban	0.464	-	(0.461, 0.467)	0.288	-	(0.283, 0.293)	0.35
Rural	0.263	-	(0.261, 0.265)	0.107	-	(0.106, 0.109)	0.54
Slum\Non Slum							
Slum	0.426	-	(0.421, 0.431)	0.297	-	(0.289, 0.304)	0.26
Non Slum	0.561	-	(0.556, 0.566)	0.436	-	(0.426, 0.445)	0.20
India	**0.329**	**-**	**(0.327, 0.330)**	**0.141**	**-**	**(0.139, 0.142)**	**0.54**

of substitution between attributes to be $\beta = -0.5$. Note that we have $\alpha < \beta$, which reflects our implicit assumption that the three attributes are substitutes to each other at the disaggregated level.

A comparison of ranks generated from these two indices reveals how the ranks alter after considering the existing inequality within the sub-groups. The rank of Madhya Pradesh decreases to 28 from 25. Kerala seizes the first rank from Delhi owing to the lowest level of inequality across her citizens. The gap in well-being between the rural and the urban areas further widenes after accounting for inequality. A similar pattern follows when comparing the level of well-being between slum and non-slum areas in the eight major cities of India. The ranking among the religious groups are mostly robust, except that the social welfare of the people in the 'others' category falls sharply. Note that this sub-group contains various religious groups and therefore the impact of high inter-religion inequality is a possible cause behind this finding. The analysis in terms of the the social welfare indices indicates persistent disparity across different castes and tribes in the Indian context. The group of people who are scheduled tribes not only have lower social welfare but also a massive amount of within group inequality.

To provide some idea about the extent of existing inequality within each group, we can estimate the level of inequality (I) using the formulation:

$$I\left(H;-0.5,0.5,a\right)=1-\frac{W\left(H;-0.5,0.5,a\right)}{W\left(\bar{H};-0.5,0.5,a\right)};$$

where $H, \bar{H} \in \mathcal{H}_N$ such that $\bar{H} = BH$, where $B = \mathbf{1}_{NN}/N$. Intuitively, the ideal situation is believed to be achieved when there is perfect equality across the population. Therefore, inequality is measured as the relative gap between the current welfare level to that of the ideal welfare level. The sixth column of Table 8 reports the level of inequality using the above formulation. The inequality in overall achievements in Kerala is the lowest and it is highest in Jharkhand. Note that the level of inequality in Madhya Pradesh is high and the state decreases by three ranks. Rural India also shows a significantly higher level of within group inequality compared to the urban area. Our study finds that the level of inequality is higher in slums than in non-slum areas.

Conclusion

In this chapter, we have applied the multidimensional social welfare indices developed in Chapter II to the Indian context. For our purpose, we used the National Family Health Survey data for the year 2005/06, which is also the Demographic Health Survey data. The chapter also develops the confidence intervals which are appropriate for the dataset. We carefully contruct three attributes of well-being using several indicators. The attributes are chosen so that they are continuous.

We use two different indices from our class. One is analogous to the Human Development Index, where the index is the simple arithmetic mean of arithmetic means. Thus, the index is not sensitive to either of the two forms of multidimensional inequality. The other index is sensitive to both forms of inequality. The parameter values are chosen in such a manner that the attributes are substitutes to each other. Our results show that the state of Kerala ranks best when the welfare index is sensitive to inequality across the population.

However, note that the dataset does not contain any information on income. As a result, we can not directly compare our multidimensional results directly with the results using an income based approach. A possible future direction of research would be to apply our class of indices using a dataset that contains information on either income or consumption expenditure.

A second possible direction of reseach would be to quantify the separate impact of the two forms of inequality. We calculated the level of inequality for every population subgroup but were not able to say which form of inequality had relatively more importance than the other. Finally, due to the lack of good health indicators, we were not able to construct a sound health indicator. Instead, we constructed a health risk attribute as a proxy. Also, our method of constructing the indicator was crude. Further research is required to construct a more sophisticated health attribute.

Chapter VI

Measuring Multidimensional Poverty in India: A New Proposal (with Sabina Alkire)

Introduction

One of the principal objectives of post-independence Indian development planning has been to eradicate poverty, thus improving the lives of those battered by deprivation and suffering. This goal is important in itself and also in turn strengthens social, political, and economic outcomes. Although this objective has remained constant, the mechanisms for addressing it have evolved. To improve the effectiveness and timeliness of policy, recent attention has focused both on direct deprivations and on income poverty. In some cases, this is because data on deprivations can be gathered more quickly than income data and at a lower cost; in other cases this arises from a direct interest in deprivations for which income poverty is an insufficient proxy. This chapter explores how the measurement of multiple deprivations may be strengthened and made more relevant for policy.

Initially, Indian poverty measures were unidimensional and based on income or expenditure. From 2002, India identified rural households as 'below the poverty line' according to a thirteen-item census questionnaire. The 2002 census process was subsequently accused of corruption and low data quality and coverage; the methodology was subject to criticisms because of the weighting and aggregation processes; and the content of the 13-item survey was challenged.

Informed by such criticisms, this chapter draws on the 2005/6 National Family Health Survey. First, it explores concerns over BPL data quality. Next, we use the NFHS dataset, which is arguably of better quality, to match the dimensions in the rural BPL census, and find 10 plausible matching indicators. We construct a pseudo-BPL score using the current methodology, and compare this with the identification and aggregation methodology proposed by Alkire and Foster (2008). Their identification strategy addresses some weaknesses of the BPL. Also, it goes beyond the BPL because it can be disaggregated, and hence can provide policy guidance at the village, block, or district level as to the components of deprivation. Using a decomposable measure would make much better use of BPL census data at minimal extra cost. For example, poverty in Orissa is driven more by deprivations in the quality of the air the household members breathe and in nutrition, whereas deprivation in assets figures more strongly in Rajasthan. In both states, a lack of women empowerment, a lack of access to sanitation, and a lack of education are widespread. Comparing the BPL methodology and the Alkire and Foster (AF) methodologies lead to different results. If all else were equal, according to the AF method, as many as 33 percent of extremely poor rural Indians would not have received a BPL card using the 2002 BPL method.

To respond to the criticisms regarding data content in the BPL survey, in this chapter we present an illustrative index of multiple deprivation, which employs nine variables, each of which represent policy goals in the 11th plan. Once again, the results are compared with income poverty and with pseudo-BPL status. Finally, the poverty rates are disaggregated by state and broken down by dimension. We demonstrate that an alternative measurement methodology is able to specify the composition of multidimensional poverty in any given state or group and to guide policy concretely and specifically.

This chapter proceeds as follows. The second section provides a brief history of poverty measurement in India and describes how Indian poverty measurement methodologies moved from being single dimensional to multi-dimensional. In the third section, we provide the theoretical framework of the 2002 BPL, which is the key approach implemented by the Indian government, and critically evaluate the process drawing on the existing literature. The fourth section describes an improved multidimensional methodology for identification and measurement proposed by Alkire and Foster (2008). The fifth section describes the NFHS data and our construction of pseudo-BPL measures and of AF measures. The sixth section compares the 2002 BPL approach with the AF methodology. The seventh section develops an index of deprivation, using NFHS data, which responds to criticisms regarding the data content in the BPL. We compare these results with income poverty and with pseudo-BPL status for sample respondents. The final section concludes.

Poverty Measurement Methodologies: Brief Review

This section provides a brief history of poverty measurement mechanisms since independence. Under the first four quinquennial plans, the government of India aimed to reduce income poverty by pursuing a high rate of economic growth measured solely in terms of the per capita gross domestic product. The rate of economic growth, however, was insufficient to cause a sharp fall in income poverty across all states and, consequently, for the first two and a half decades, the income poverty rate hovered between 38 percent and 57 percent without any particular trend. The official measure of poverty for that entire period was based on consumption expenditure (Radhakrishna and Ray, 2005).

In the early seventies, for the first time the basic minimum needs approach gained prominence. The Planning Commission appointed a Task Force on 'Projections of Minimum Needs and Effective Consumption Demand' that defined the rural poverty line as the per capita consumption expenditure level needed for a minimum required calorie intake in rural and urban areas. Thus although poverty measurement remained in income space, the basis of poverty measurement evolved from the income-based approach to the basic-needs-based approach. According to the recommendation of the task force, the minimum basic food intake requirement for the rural and the urban habitants was 2400 calories and 2100 calories, respectively (Government of India, 1979). Based on these minimum calorie requirements, the minimum required subsistence income levels were determined for different regions. These minimum required income levels were used as regional poverty thresholds. Since then the Indian poverty analysis has been based on consumption expenditure (Datt and Ravallion, 2002).

To improve the effectiveness and timeliness of policy, recent attention has focused on specific deprivations besides income poverty. To target services to the most needy, the government developed a measure by which families were categorised as living 'below the poverty line' (BPL). Since 1992, three successive BPL censuses (1992, 1997, 2002) identified rural families that are below the poverty line and thus eligible for government support such as subsidized food or electricity, and schemes to construct housing and encourage self-employment activities. Each BPL census applied a unique identification technique. The first BPL survey in 1992 gathered self-reported income data and used the all-India income poverty line to identify BPL households. This generated very high estimates of rural poverty (52.5%). Moreover, this approach was based on income data, which may be less accurate than consumption data (Atkinson and Micklewright, 1983; Grosh and Glewwe, 2000).

To improve upon the 1992 methodology, the 1997 BPL census used expenditure and multiple criteria rather than income data alone, and excluded the visibly non-poor. It had two parts. The first part was administered to all rural households, and identified as "visibly non-poor" households who met certain requirements. If the household was not registered as visibly non-poor, it was administered a survey, which gathered basic socio-demographic information, as well as household characteristics, and consumption expenditures over the past 30 days. However, critics including a subsequent Expert Review criticised the 1997 methodology for four reasons. First, the exclusion criteria were too stringent (the possession of a single ceiling fan was grounds for exclusion). Secondly, poverty lines for all states and Union Territories (UT) were lacking. Thirdly, the BPL criteria were not uniform across states; hence, interstate comparisons were difficult. Fourthly, there were no procedures available to add new families to the BPL lists for five years (Government of India, 2002; Hirway, 2003; Jalan and Murgai, 2007; Sundaram, 2003). Finally, the non-poor households were identified according to their resources rather than what household members were capable of being and doing. This is the fundamental distinction between the needs-based approach and the capability approach of Amartya Sen.

The next section describes the 2002 BPL methodology in detail, and identifies both its strengths and shortcomings.

Below The Poverty Line (BPL) 2002: Methodology and Critiques

In 2002, rural households were asked a set of non-income questions and the responses were used to identify those households that were qualified to receive BPL cards. No additional analysis was conducted using the census dataset other than the identification of the BPL card holders. How did this proceed?

2002 BPL Methodology

The 2002 rural BPL census consists of thirteen questions for each household, comprising topics such as food, housing, work, land ownership, assets, education, and so on.[52] Depending upon the response category selected, the household is assigned a score (0-4) for each variable. A household's score is then summed to create an aggregate score. A poverty cutoff is fixed at the State level or at lower levels for the aggregate score. Households falling below that area's poverty cut-off are identified as 'BPL'. At the state or UT level, a further limit was fixed: the number of households identified as BPL was limited to ten percent above the BPL figures estimated in 1999-2000.

Like every other poverty measure, the 2002 BPL methodology involves two components: the identification of the poor and the aggregation of the data into a single poverty index (Sen, 1976). Let us introduce the notation that we use to describe the 2002 BPL method. We denote the set of all positive integers by $\mathbb{N}$. Let us assume that there are $N \in \mathbb{N}$ households in the economy and the well-being of each household is gauged using $D \in \mathbb{N}$ dimensions. The achievements of the households in the entire society are summarized by an $N \times D$ dimensional matrix $H \in \mathbb{R}_{+}^{ND}$, where $\mathbb{R}_{+}^{ND}$ is the $N \times D$-dimensional non-negative euclidean space. The set of all $N \times D$ dimensional matrices is denoted by $\mathcal{H}$. The sum of entries in any given vector or matrix a is denoted by $|a|$, while $\mu(a)$ is used to represent the arithmetic mean of a. The achievement of the n^{th} household in the d^{th} dimension is denoted by h_{nd} for all $d = 1, \ldots, D$ and all $n = 1, \ldots, N$.

The first stage of the BPL method identifies which households are multidimensionally poor. Let us designate the set of categories for the d^{th} dimension by $I_d = \{0, 1, \ldots, i_d\}$,

[52] These questions and the response categories are reprinted in Appendix G.

where $i_d \in \mathbb{N}$ is the score attached to the highest category in the d^{th} dimension.

First, an $N \times D$-dimensional matrix X is constructed from the matrix H, where x_{nd} is the nd^{th} element of X such that $x_{nd} \in I_d$ for all d and for all n. For example, suppose the dimension of well-being is acres of land holding. Instead of using the amount of land holding directly, five categories were created. See Appendix G. Thus, the n^{th} element in the d^{th} dimension can take any integer value between zero and i_d such that $0 \leq x_{nd} \leq i_d$. Each household is provided a score in each dimension based on their achievement in that particular dimension. The overall welfare score of the household is calculated by summing the dimensional scores. The welfare score of the n^{th} household is denoted by $\mathcal{D}_n = \sum_{d=1}^{D} x_{nd}$. The minimum possible welfare score is zero and the maximum possible welfare score is $\hat{\mathcal{D}} = \sum_{d=1}^{D} i_d$. Therefore, $0 \leq \mathcal{D}_n \leq \hat{\mathcal{D}}$ for all n. A household is identified as poor if the welfare score of that household lies below a certain threshold, which is called a poverty line or a poverty cut-off and is denoted by z. The n^{th} household is poor and identified as 'below the poverty line' if $\mathcal{D}_n < z$ and non-poor, otherwise.

After identifying the poor, an N-dimensional vector $Y = (y_1, \ldots, y_N)$ is created such that $y_n = 1$ if $\mathcal{D}_n < z$ and $y_n = 0$, otherwise. In other words, Y is a vector containing only zeros and ones: an element is equal to one if the corresponding household is poor and zero if the household is non-poor. Finally, the BPL poverty rate is equal to:

$$P_{BPL} = \frac{1}{N} \sum_{n=1}^{N} y_n.$$

We can think of each BPL question as a dimension of social welfare, i.e., $D = 13$. The response to each question comprises five categories, i.e., $I_d = \{0, 1, 2, 3, 4\}$ for all d. The worst category is assigned a score of zero; whereas the best category is assigned a score of four. In the three intermediate categories a higher value implies a better category. The score for the n^{th} household in the d^{th} dimension is equal to h_{nd}, where $0 \leq h_{nd} \leq 4$ for all d and all n. The minimum possible overall welfare score is zero and the maximum possible overall welfare score is $\hat{\mathcal{D}} = 52$, i.e. $0 \leq \mathcal{D}_n \leq 52$ for all n. Households falling below that area's poverty cut-off (these vary by state or district) are identified as 'BPL'.

Critiques of the BPL Process

The 2002 BPL results have come under fierce criticism from many sides. See Hirway (2003), Jain (2004), Jalan and Murgai (2007), Mukherjee (2005), and Sundaram (2003) among others. The criticisms might be roughly divided into three kinds: methodological drawbacks in identification and aggregation, data quality and corruption, and issues of data content.

Methodological Drawbacks in Identification and Aggregation

The main methodological criticisms of the BPL indicator are as follows:

1. *Cardinalization* – The method by which the response variables are summed into a welfare score $\mathcal{D}_n$ is problematic for the following reasons. First, the raw data are categorical, and their ordering might be disputed. Yet even if one agrees with the ordering of the responses, the distance between the responses for each dimension is not known. There is no justification for assuming the distance between each category to be uniform.

   Furthermore, the inter-dimensional comparison of scores presumes cardinality across dimensions. For example, a household is assigned a score of two if either the household members enjoy one (not two) square meals per day throughout the year, or if the household includes at least one person who has completed secondary schooling. However, these two situations may not appear to reflect the same degree of deprivation. In a country where about 60 percent of students drop out before completing secondary education, a household with a member completing secondary education is reasonably well off. Nevertheless, a household seems less likely to be well nourished if the entire household were to survive only on one square meal a day for the entire year. The cardinalization of ordinal data in this way may not be highly intuitive.

2. *Complete Substitutability across Dimensions* – A second and related problem is that the scores for the thirteen dimensions are aggregated into a single overall score such that $\mathcal{D}_n = \sum_{n=1}^{N} x_{nd}$, and the poor are identified according to a cut-off set across the aggregate score, $\mathcal{D}_n < z$. This simple aggregation is equivalent to treating all dimensions as perfect substitutes. A one-point gain in one dimension can be compensated by an equivalent one-point decrease in any other dimension, at any other level of achievement. Once again, this does not appear to be a convincing argument. The problem can be explained in terms of the poverty focus axiom and the deprivation focus axiom.[53] According to the poverty focus axiom, if there is an increase in any dimension among the non-poor, the poverty rate should remain unchanged. According to the deprivation focus axiom, if there is an increase in any dimension in which a household is not deprived (whether the household is poor or non-poor), the poverty rate remains the same. Although the BPL does not identify deprivation thresholds, intuitively the

[53] For formal definitions of the poverty and deprivation focus axioms, see Alkire and Foster (2008).

BPL method satisfies the poverty focus axiom, but not the deprivation focus axiom.

Consider a marginally poor household in Uttar Pradesh that requires only one point so that it can move above the BPL poverty line. Along with other achievements, the household owns 5 hectares of un-irrigated land but survives normally on one square meal per day but less than one square meal occasionally. The household is deprived in terms of food security but is not deprived in terms of land holding. Note that if the household owned 5.1 hectares of land it would score '4' rather than '3' in that dimension. Further, this change in score in a non-deprived dimension would increase its aggregate score, hence pull it above the BPL poverty line. The total substitutability among the BPL dimensions at all levels appears to be particularly undesirable given their equal weight and the problems in data content.

3. *Equal Weighting of Dimensions* – The thirteen dimensions are combined using equal weights. This implies that each dimension makes an equally important and equally valuable contribution to poverty. But no justification for these weights is provided. Jalan and Murgai (2007) argue that the relative weights on dimensions should also be allowed to vary across states because different indicators do not have the same impact across states. For example, education should be weighted differently in Bihar than in Kerala. Alternatively, in a different context, Atkinson et al. (2002) argues that the dimensions should be chosen explicitly such that they are roughly equal in normative (ethical) importance. Sen and others argue that the weights, being value judgements, must be subject to public discussion (Sen, 1996, 2004).

4. *Varying Poverty Lines* – No national poverty line is set; rather states, and in some cases districts, set their own poverty line across the 52-point scale. Jain (2004) observes that the district poverty line varies from 12-15 in Madhya Pradesh, driven by the need for each district to match "the 'already declared' proportion of poor in that district". In this situation, the fact that the BPL status of a family with fourteen points depends only upon its district level quota for the year 2000 seems rather difficult to defend, particularly when the poverty quotas are controversial, which is our next point. While there is no easy response to this situation, the need for flexibility and state autonomy must be balanced against the need to maintain uniform standards.

5. *Imposed Poverty Quotas* – To ensure that the numbers of BPL households did not exceed fiscal resources, the States' BPL estimates were capped so that they could not exceed the NSSO 1999-2000 estimates by more than 10 percent. This particular cap has been widely disputed, because BPL is not measuring income poverty. Using the 1999-00 and 2004-05 NSS datasets, Jalan and Murgai find that identification of the poor through 2002 BPL method is an inadequate proxy for consumption poverty. "The BPL score misclassifies nearly half (49 percent) of the [consumption] poor as non-poor, and conversely, 49 percent of those identified as BPL poor are actually [consumption] non-poor. Even in the "best" state, Orissa, 32 percent of the poor are misclassified while in the worst state, Andhra Pradesh, three out of every four poor people are misclassified as non-poor based on the BPL indicator" (p 7). As Hirway argues, given the multidimensional approach of the BPL census, "There is no reason why the two estimates should match, and there is no logic in reducing the estimates of poverty of one kind to match the other kind of poverty!" (p. 4804). While clearly there are

needs to impose some limits for reasons of fiscal constraints and accuracy, the use of 1999-2000 NSS data creates errors of inclusion and exclusion in states.

6. *Neglect of Intensity of Poverty Across Households* – The BPL method is not sensitive to the inequality in terms of the range of deprivations BPL households suffer. In a region with high inequality among the poor, the BPL method does not provide the policy maker information on who among the BPL are extremely poor rather than marginally poor. However, the extreme poor might claim special priority either in terms of the targeting or level of provision of government services.

7. *High Cost; Low Policy Impact* – Fielding a rural census of households is costly, and gives rise to potentially powerful data to guide policy even at very local areas. Unfortunately, the BPL measure makes very rudimentary use of the BPL data. The current BPL identification gives rise to an aggregate headcount. But this cannot be decomposed to show the composition of poverty in different villages, blocks and districts or for different cultural groups or kinds of households. Such analysis is extremely important for policy since it allows a policy maker to understand the components of poverty for each group, and thus to craft an effective and efficient response.

Corruption, Data Quality and Data Coverage

"Targeting", Hirway observed, "is not a statistical exercise, but is a major political activity" (p. 4804). Because households identified as BPL access multiple benefits, Hirway observes, there is "a mad rush in our villages to be enrolled as BPL households." Concretely, "The rich and powerful in a village frequently pressurises the talati and the sarpanch to include their names in BPL lists" (p. 4805, see also Khera (2008)); poor households may not be interviewed, their interviews may be distorted, and they may not be able to convince the local elite to include their names on the list. Jain gives particular examples of poor data quality: "Charua Singh was excluded from the BPL list because the enumerator had filled up the form without visiting Charua's house." (Jain, 2004, p. 4982). Jain also argues that pavement dwellers, who have no address for the BPL ration cards, households displaced by riots and communal violence, manual scavengers, and communities involved in caste-based prostitution are systematically excluded from BPL status.

Corruption crowds out the poor from BPL card ownership. Drawing on village level studies in Rajasthan, Khera (2008) reports the striking finding that 44 percent of poor households did not have a BPL card, and 23 percent of those with a BPL card were non-poor. Hirway finds 11-18 percent of the 1997 BPL list members in Gujarat are clearly local elite, and 14 percent of the poor households were excluded from the BPL lists. Further, the truly poor (rather than the mis-classified elite) had greater difficulty in using their BPL status effectively to enjoy all its intended benefits. In participatory social assessments in West Bengal, Mukherjee (2005) found that "in some villages the [BPL] list had been manipulated to the extent of 50 percent with the inclusion of many non-poor households." (p. 12). The manipulation appeared to occur after the survey, through corruption: "Though door-to-door BPL survey was conducted, the final outcomes in terms of the BPL list shocked many genuine poor in terms of not finding their names on the list" (p. 12).

Although some crosschecks were successful in revising BPL lists to correct inaccuracies, others were infiltrated. For example, a triangulation process had been set up to verify the BPL results, in that the BPL list was to be read out in the gram sabha so that inaccuracies could be addressed, and a revised list read out at a second meeting. But in Jain's Madhya Pradesh case study, this cross check rarely functioned. "In Petlawad, the block level panchayat officials declared the first list as final and entertained no grievances; the Kotma block panchayat officer refused to disclose the list in public... As per the study of 100 panchayats, it was found that in 67 panchayats, no second gram sabha meeting was organised for approving the list..." (Jain, 2004, p. 4983).

It is true that the case studies are dispersed and anecdotal. But as the 2002 BPL census did not have explicit mechanisms to correct for distortions in the "situations where the poor are not powerful enough to assert themselves and the administration is not strong enough to identify the poor correctly" (Hirway, 2003, p. 4806), the grave doubts about data quality seem worth exploring further.

Data Content and Periodicity

Even if the thirteen 2002 BPL indicators had been implemented accurately and without corruption, a number of authors argue that the outcomes would still be inaccurate. In the case study from MP, Jain and the Alliance Campaign for Good Governance argue that the BPL 2002 had "inappropriate indicators". They argue that even if the dimensions were justifiable (a separate question), the indicators should have taken into account the quality of land, the size of house, whether clothes were provided as gifts, and the quality as well as number of meals eaten per day.

In addition, the BPL census focuses mainly on resources (land, house, clothing, food, bathroom, consumer goods, loans, 'want from government'), rather than on capabilities – the things that households are able to do and be (be nourished, be healthy). The education questions come closest to approximating capabilities. The difficulty with resources, as Amartya Sen argues, is that the capabilities a physically disabled household or a pregnant mother are able to achieve from a given bundle of resources (2 kgs rice and a bicycle, for example) may be very different than the capabilities others could achieve. Concretely, Jain observes that the BPL systematically excludes certain categories of people such as the disabled, who may score above the poverty line in the space of resources (fan, clothes, or bicycle) but not be able to enjoy basic capabilities. Given the diversity of people's ability to convert resources into capabilities, if development aims at expanding capabilities, the constituent indicators should, when possible, focus directly on capabilities (such as nutritional status) rather than resources (number of meals).

Another striking aspect of the BPL survey, which has not received sufficient critical comment, is the response structures. The response structure on the status of the household labour force will systematically regard female-headed households as more deprived, which is understandable (although it is unclear what score will be given if women and men both work, and why that might be inferior to men alone working). However, if a household is unable to work because of illness, disability, or unemployment, they may respond 'other' and thus be given the least deprived score of 4, which seems aberrant. A similar difficulty is evident in the response structure 'means of livelihood'. Both Sundaram (2003) and Jalan and Murgai (2007) find the ordering of the livelihood category problematic – for example, it assumes that a small business household (e.g. an artisan) is always better off than one employed in agriculture (e.g. a landowner). Also a household who has 'no indebtedness' scores the value of 4, regardless of whether it has no loan because it is socially marginalized (drug addicts), and family and banks will not lend to them, or because it is sufficiently wealthy not to require a loan.

The ranking of the last two dimensions are particularly confusing. In case of reason for migration from the household, the logic of ordering is not transparent. While many poor households are migrants, the more educated, more empowered are also subject to migratory pressures and many rural poor are 'left behind.' Yet according to this response structure, a nuclear Bengali family whose son is a high profile software engineer residing in Bengaluru (earlier well-known as Bangalore) would receive a score of two, whereas a family of bonded laborers, who has not migrated anywhere would receive a score of three.

The final question of the BPL is 'preference of assistance from government'. It is not evident how responses will reveal information regarding the respondent's own socio-economic status. There is no proper justification as why a family seeking assistance on housing would receive higher score than a family seeking assistance for skill upgradation. Moreover, the responses will be influenced by respondents' assessment of government capabilities. This is a discrete variable in which the elements are difficult to order at all; the BPL practice of ascribing a cardinal meaning to the resulting scores merits review. From the discussion of the last few paragraphs, it is evident that some of the response structures are in fact misleading and require the introduction of more useful dimensions of social welfare.

A further and distinct set of criticisms refer to the fact that the BPL surveys are only conducted every five years, but households' economic status can shift rapidly, and transient spells of poverty affect many households. Unless there are ways to update the BPL status between surveys, even if the initial identification of BPL households was accurate, it is certain to become inaccurate over time. The likely magnitude of that inaccuracy could be important to consider.

This section has enumerated in detail the tremendous challenges that were encountered in the 2002 BPL census process. A number of these challenges, relating to corruption and to the census instrument, have been the focus of other accounts and surely will be addressed in the next BPL census. The remainder of this chapter focuses on the above methodological criticisms and suggests an alternative approach.

Multidimensional Poverty: A New Methodology

A Planning Commission Report from the Working Group on Poverty Alleviation (Government of India, 2006) explicitly took a "multi-dimensional view of poverty" (p. 18) which it also calls a 'multiple deprivation' view (p. 24) rather than a norm based on calories or income. It interpreted the 2002 BPL not as a proxy for income or expenditure poverty, but rather as a direct measure of multidimensional poverty that encompasses expenditure poverty and goes beyond it. The Report explicitly stated that "the possibility of conflict between the magnitude of poverty as revealed by the BPL surveys and as estimated on the basis of NSS surveys ... need not be a major issue ..." (p. 25).

This approach is consistent with other empirical work, which has identified the inherent value of multidimensional poverty measures for guiding policy (Laderchi et al., 2003; Laderchi, 2008). Many have argued that human poverty and deprivation go beyond income or ownership of material wealth (Drèze and Sen, 2002). Yet even in this case, direct attention to other variables such as education, health, and nutrition might not be required if income were a sufficient proxy for these outcomes and if policies to reduce income poverty consistently reduced other deprivations. Unfortunately, this is not necessarily the case now any more than in the early periods after independence. Since liberalization, India has enjoyed a strong rate of economic growth. Yet human development indicators remain uneven and weak. The first page of the 11th plan of India states the following concern: "the National Family Health Survey-3 (NFHS-3) shows that almost 46 percent of the children in the 0 to 3 years' age group suffered from malnutrition in 2005–06, and what is even more disturbing is that the estimate shows almost no decline from the level of 47 percent reported in 1998 by NFHS-2" (Government of India, 2008). More generally, across developing countries, Bourguignon et al. (2008) find "little or no correlation between growth and the non-income MDGs". Another reason to use indicators in addition to income is that some families experience multiple deprivations, whereas others are deprived only in one dimension. Clearly, the households with multiple deprivations should be targeted. For these reasons, it is useful to explore measures of human deprivation that can identify households with multiple deprivations. Finally, it is useful to see the leading components of deprivation in different states and districts, as analysis of such data can be used to design the most effective sequence of interventions.

In the previous section, we critically evaluated the BPL approach. In our first criticism of the BPL approach, we pointed out the methodological drawbacks of the identification and aggregation process. This section is devoted towards addressing these methodological weaknesses and proposes adopting a recent methodology for multidimensional poverty measurement developed by Alkire and Foster (2008). The Alkire and Foster (AF) method was selected because it addresses the methodological concerns of the current BPL aggregation method discussed in the previous section in the following ways:

1. *Valid treatment of ordinal data* – The AF measure is suitable for ordinal data. By applying dimension-specific cut-offs, households are classified as either deprived or

non-deprived in that dimension. This has the effect of dichotomising ordinal data and thus avoids the problem of cardinalization.

2. *Poverty and Deprivation Focused* – By applying cut-offs to each dimension, each household is judged to be deprived or not in that dimension independently of its achievements in other dimensions. Thus, we do not have a situation of perfect substitutability where an increase in landholdings from 5 to 5.1 hectares can compensate for a decrease from one square meal per day to complete food insecurity. Rather, multidimensional poverty status only depends on dimensions in which households are deprived.

3. *Equal or general weights* – It is possible to weight the dimensions equally, or, to weight indicators and dimensions differently, or indeed to explore several weighting structures and the robustness of the BPL status according to variable weights.

4. *Poverty lines can be fixed or flexible* – In our example, we have used the same deprivation cut-offs nationally both for each deprivation and across deprivations. However, these could be fixed at district or state levels if that were deemed more appropriate.

5. *Highly informative for policy* – Finally and most importantly, in the current BPL measure, the census data are used solely to designate households as BPL or 'Above the Poverty Line' (APL). However using the AF measure, the BPL population of any state or ethnic group can be scrutinised to see what deprivations are mainly responsible for their multidimensional poverty. This information, taken together with other analyses made possible by the same data (hence at minimal extra cost) can inform policy. Using the AF measure, responses can be tailored to the composition of poverty in different states or districts, making them more efficient and effective.

The Alkire and Foster Methodology

As in the discussion of BPL methodology, consider a society with N households and D dimensions.[54] Let $\mathcal{H}$ denote the set of all $N \times D$ matrices and $H \in \mathcal{H}$ represents an achievement matrix of a society, where h_{nd} is the achievement of the n^{th} household in the d^{th} dimension for all d and all n.[55] The n^{th} row and the d^{th} column of H are denoted by $h_{n\cdot} = (h_{n1}, \ldots, h_{nD})$ and $h_{\cdot d} = (h_{1d}, \ldots, h_{Nd})$. The row vector $h_{n\cdot}$ summarizes the achievements of household n in the D dimensions; whereas, the column vector $h_{\cdot d}$ represents the distribution of achievements in the d^{th} dimension across the N households. We denote the D-dimensional deprivation cut-off vector by $\mathbf{z}$, where the deprivation cut-off for the d^{th} dimension is indicated by z_d.

Corresponding to any $H \in \mathcal{H}$, an $N \times D$ dimensional deprivation matrix g^0 is constructed, where the nd^{th} element is denoted by g^0_{nd}. Any element of g^0 can take only two values as follows:

$$g^0_{nd} = \begin{cases} 1 & \text{if } h_{nd} < z_d \\ 0 & \text{otherwise} \end{cases} .$$

In other words, the nd^{th} entry of the matrix is equal to one when the n^{th} household is deprived in the d^{th} dimension and is equal to zero when the household is not deprived. From the matrix g^0, we construct an N-dimensional column vector C of deprivation counts such that the n^{th} element $c_n = |g^0_n|$ represents the number of deprivations suffered by the n^{th} household. If the dimensions in H are cardinal, then we construct a normalised gap matrix

[54] Although we choose households rather than individuals as the unit of analysis in order to parallel the BPL methodology, it is of course possible to focus instead upon individuals.

[55] A society could be a nation, state, or any geographic region.

g^1, where the nd^{th} element is:

$$g^1_{nd} = \begin{cases} (z_d - h_{nd}) & \text{if } h_{nd} < z_d \\ 0 & \text{otherwise} \end{cases} .$$

By construction, $g^1_{nd} \in [0, 1]$ for all n and all d, and each element gives the extent of deprivation experienced by the n^{th} household in the d^{th} dimension. The generalized gap matrix is denoted by g^α, with $\alpha > 0$. The nd^{th} element of g^α is denoted by g^α_{nd}, which is the normalised poverty gap raised to the power α.

Now, we are in a position to provide an outline of the class of multidimensional poverty measure proposed by Alkire and Foster (2008). The first stage of multidimensional poverty measurement is to identify the poor. Most existing poverty measures identify the poor either by the union approach or by the intersection approach. According to the union approach, a household is identified as poor if the household is deprived in at least one dimension. On the other hand, a household is identified as poor according to the intersection approach if the household is deprived in all dimensions. Note that the 2002 BPL method does not follow either of these approaches. If dimensions are equally weighted, the multidimensional approach proposed by Alkire and Foster identifies a household as poor if the household is deprived in at least k dimensions, where $k \in \{1, \ldots, D\}$.[56] Thus, k can be considered as a second poverty cut-off.

Let us define the identification method ρ_k such that $\rho_k(h_{n\cdot}, z) = 1$ if $c_n \geq k$, and $\rho_k(h_{n\cdot}, z)$ $= 0$ if $c_n < k$. This implies that a household is identified as multidimensionally poor if the household is deprived in at least k dimensions. Note that for $k = 1$, the identification

[56] Equal weights are presented first for simplicity; we discuss general weights below.

criterion is equivalent to the union approach; whereas, the identification criterion is the same as the intersection approach for $k = D$. The set of multidimensional poor, according to this identification criterion, is defined by $Z_k = \{n : \rho_k(h_{n\cdot}, z) = 1\}$. A censored matrix $g^0(k)$ is obtained from g^0 by replacing the n^{th} row with a vector of zeros whenever $\rho_k(h_{n\cdot}, z)$ $= 0$. An analogous matrix $g^\alpha(k)$ is obtained for $\alpha > 0$, with the nd^{th} element $g^\alpha_{nd}(k) = g^\alpha_{nd}$ if $\rho_k(h_{n\cdot}, z) = 1$, while $g^\alpha_{nd}(k) = 0$ if $\rho_k(h_{n\cdot}, z) = 0$.

Based on this identification method, Alkire and Foster define the following poverty measures. The first natural measure is the percentage of individuals that are multidimensionally poor. Analogous to the single-dimensional headcount ratio, the multidimensional Headcount Ratio is defined by $HCR(H; z) = Q/N$, where Q is the number of individuals in the set Z_k. This measure has the advantage of being easily comprehensible and estimable. Moreover, this measure can be applied using ordinal data. Unfortunately, it is completely insensitive to the intensity and distribution of poverty, as first noticed by Watts (1969) and Sen (1976) in the single-dimensional context. It also fails to satisfy the properties of *transfer* and *monotonicity*. In addition, in the multidimensional context, it violates *dimensional monotonicity*. Alkire and Foster describe this problem as follows: if a household already identified as poor becomes deprived in an additional dimension in which the household was not previously deprived, HCR does not change. Finally, this measure is not flexible to dimensional decomposition, which is often useful for policy recommendation.

To overcome the limitations of the multidimensional headcount ratio, Alkire and Foster propose the class of dimension-adjusted Foster-Greer-Thorbecke measures, defined by $M_\alpha(H; z) = \mu(g^\alpha(k))$ for $\alpha \geq 0$. For $\alpha = 0$, the class of measures yields the *Adjusted Headcount Ratio*, defined by $M_0 = \mu(g^0(k))$. The adjusted headcount ratio is the total number of deprivations experienced by all poor households divided by the maximum number of deprivations that could possibly be experienced by all households and is formulated by $|g^0(k)|/ND$. It can also be expressed as a product between the percentage of multidimensional poor (HCR) and the average deprivation share across the poor given by $A = |g^0(k)|/QD$. Thus, $M_0 = HCR \cdot A$. In words, A provides the fraction of possible dimensions D in which the average multidimensionally poor household is deprived. In this way, M_0 summarises information on both the incidence of poverty and the average extent of a multidimensional poor household's deprivation. This measure is as easy to compute as the HCR and can be calculated with ordinal data, but it is indeed superior to the HCR since it satisfies the property of dimensional monotonicity described above.

When some data are cardinal, for $\alpha = 1$, the class of dimension-adjusted FGT measures yields the *Adjusted Poverty Gap*, given by $M_1 = \mu(g^1(k))$, which is the sum of the normalised gaps of the poor $|g^1(k)|$ divided by the highest possible sum of normalised gaps ND. It can also be expressed as the product between the percentage of multidimensional poor households HCR, the average deprivation share across the poor A, and the average poverty gap G, where $G = |g^1(k)|/|g^0(k)|$. Thus, $M_1 = HCR \cdot A \cdot G$. M_1 summarises information on the incidence of poverty, the average range of deprivations, and the average depth of deprivations of the poor. It satisfies not only dimensional monotonicity, but also monotonicity: if an individual becomes more deprived in any dimension in which they are already deprived, M_1 will increase.

Finally, for $\alpha = 2$, this class of measures yields the *Adjusted Squared Poverty Gap* (M_2), defined by $M_2 = \mu(g^2(k))$, which is the sum of the squared normalised gaps of the poor $|g^2(k)|$ divided by the highest possible number of normalised gaps ND. It can also be expressed as the product between the percentage of multidimensionally poor HCR, the average deprivation share across the poor A, and the average severity of deprivations S, which is given by $S = |g^2(k)| \,/\, |g^0(k)|$. Thus, $M_2 = HCR\cdot A\cdot S$. M_2 summarises information on the incidence of poverty, the average range, and severity of deprivation of the poor. If there is a regressive transfer among two poor persons, then M_2 increases, unlike M_1 and M_0. This measure satisfies both types of monotonicity principle, the transfer principle, and is sensitive to the inequality among the poor because it emphasizes the deprivations of the poorest.

All members of the M_α family are decomposable by population subgroups. Given two separate achievement matrices H_1 and H_2, with population size of N_1 and N_2, respectively, the overall poverty level for $N = N_1 + N_2$ individuals is obtained by:

$$M_\alpha\left(H_1, H_2; \mathbf{z}\right) = \frac{N_1}{N} M_\alpha\left(H_1; \mathbf{z}\right) + \frac{N_2}{N} M_\alpha\left(H_2; \mathbf{z}\right).$$

Clearly, this can be extended to any number of subgroups. All members of the $M_\alpha(H; \mathbf{z})$ family can be decomposed into dimensional subgroups as $M_\alpha(H; z) = \sum_{d=1}^{D} \mu\left(g^\alpha_{\cdot d}\left(k\right)\right)/D$, where $g^\alpha_{\cdot d}$ is the d^{th} column of the censored matrix $g^\alpha(k)$. It is a very convenient decomposability property; $\mu\left(g^\alpha_{\cdot d}\left(k\right)\right)/M_\alpha(H; z)$ can be interpreted as the post-identification contribution of the d^{th} dimension to overall multidimensional poverty.

The M_α family of measures are neutral to the association increasing transfers defined in Chapter II (p. 21). If one achievement matrix is obtained from another achievement matrix by an association increasing transfer among the poor, both of them yield the same level of poverty. The additive form enables the family of measures to evaluate the achievement of each household in each dimension independently of the achievements in the other dimensions.

In this sense, the M_α family of measures is analogous to the first group of measures of Bourguignon and Chakravarty (2003).

Weighting

Apart from identification and aggregation, another important challenge in multidimensional poverty measurement is how to weight different dimensions. The weights implicitly indicate the dimensional importance and/or policy priority. In the preceeding analysis, the dimensions were presented as if they were equally weighted. Equal weights is an arbitrary and normative weighting system that is appropriate in some, but not all, situations (Atkinson et al., 2002). In many other cases, some dimensions are believed to be more important than others, and hence should to receive a relatively higher weight. Thus, we move from equal weights to unequal weights. The M_α family can be easily extended to a more generalized form considering unequal weighting structures.

Let w be a D-dimensional row vector with the d^{th} element being equal to w_d, which is the weight associated with the d^{th} dimension such that $|w| = D$. We define the $N \times D$ dimensional matrix $g^\alpha(w_d)$ with the nd^{th} element being equal to g^α_{nd} that takes two values as follows:

$$g^\alpha_{nd}(w_d) = \begin{cases} w_d\left((z_d - h_{nd})/z_d\right)^\alpha & \text{if } h_{nd} < z_d \\ 0 & \text{otherwise} \end{cases}.$$

The weighted column vector C of deprivation counts can be obtained with the n^{th} element being equal to $c_n = |g^0_{n\cdot}|$; c_n varies between 1 and D. In this situation, the dimensional cut-off for the identification step is a real number k, such that $0 < k \leq D$, instead of k being a positive integer. When $k = \min\{w_d\}$, the criterion is nothing but the union approach, whereas, $k = D$ yields the intersection approach. Also note that if $w_d = 1$ for all d, then the weighting structure turns out to be the equal weighting structure. After the multidimensionally poor are identified, the identification method is denoted by ρ_k such that $\rho_k(h_{n\cdot}; \mathbf{z}, w_d) = 1$ when $c_n \geq k$, and $\rho_k(h_{n\cdot}; \mathbf{z}, w_d) = 0$ when $c_n < k$. Finally, a censored matrix $g^0(k, w_d)$ is obtained from $g^0(w_d)$ by replacing the n^{th} row with a vector of zeros whenever $\rho_k(h_{n\cdot}, z) = 0$. An analogous matrix $g^\alpha(k, w_d)$ is obtained for $\alpha > 0$, with the nd^{th} element $g^\alpha_{nd}(k, w_d) = g^\alpha_{nd}(w_d)$ if $\rho_k(h_{n\cdot}; \mathbf{z}, w_d) = 1$, while $g^\alpha_{nd}(k, w_d) = 0$ if $\rho_k(h_{n\cdot}; \mathbf{z}, w_d) = 0$. The class of dimension-adjusted FGT measures is defined by $M_\alpha(H; \mathbf{z}, w_d) = \mu(g^\alpha(k; w_d))$ for $\alpha \geq 0$.

Having introduced the new methodology, we now compare it to the methodology applied in the 2002 BPL process. Our empirical results draw on the National Family Health Survey dataset for the period of 2005-06, which is introduced in the next section.

Data

The National Family Health Survey (NFHS-3) for the year 2005/06 has been collaboratively conducted by the International Institute for Population Sciences (IIPS), Mumbai, India; ORC Macro, Calverton, Maryland, USA; and the East-West Center, Honolulu, Hawaii, USA. The survey interviewed 124,385 women aged 15-49 and 74,369 men aged 15-54 from 109,041 households and from all 29 states of India including Delhi. Unlike the previous two surveys, NFHS-3 interviewed never-married women, never-married men, and ever-married men in addition to ever-married women. Besides collecting information on household characteristics, such as housing structures, access to sanitation, water sources, and assets, the survey collected data on individual characteristics, such as the level of education and the health status of the respondents. Numerous questions in the survey are analogous to the questions asked in the BPL questionnaire. This allows us to make comparisons between the BPL method and the AF method of poverty measurement. We list all the related questions in Table 9.

In order to compare our findings with the rural BPL population, we focus on rural areas. The rural BPL survey is uniform, and distinct from urban BPL methods. The NFHS collects information for men and women in 58,805 rural households. Because the unit of analysis for the BPL method is household instead of individual, we keep the household as our unit of analysis. In this chapter, we weight the households by the nationally representative sample weight provided in the dataset (See Appendix K.)

The 2002 BPL Method versus the AF Method

In this section, we use the NFHS-3 dataset to compare the identification technique of the BPL 2002 method with that of the AF method. First, we select dimensions or variables to match the BPL questionnaire as closely as possible, and report the descriptive statistics. Then we replicate the 2002 BPL score structure using the chosen set of variables and identify the households that are poor using a pseudo-BPL method. The pseudo-BPL method applies the BPL 2002 method to identify the poor, but only using the matched dimensions from the NFHS-3 data set. Finally, we compare the results obtained using the pseudo-BPL method to the results obtained using the AF method for the same set of variables drawn from the NFHS data set.

Matching Dimensions, Indicators, and their Poverty Cut-offs

We select NFHS variables or questions that match, as closely as possible, to those present in the 2002 BPL questionnaire.[57] The match is not perfect, and no proxy is available for three of the questions, thus our comparison is affected by the differences in dimensions. In the first three columns of Table 9, we summarize the questions asked in the BPL questionnaire and the analogous questions asked in the NFHS-3 questionnaire. It is evident from Table 9 that ten out of thirteen questions in the NFHS-3 are analogous to the BPL questions. Out of the ten questions, some are directly matched; the rest are obtained by manipulating several other questions.[58] The 2005/6 NFHS is not able to match BPL questions 3, 12, and 13.[59] The chosen variables restrict the sample size to 42,717 households, which contain 238,179 persons from 28 states of India.

We exclude Delhi from our analysis because Delhi primarily consists of urban areas; whereas our analysis focuses on rural areas. Note that all of our results are corrected for population weights. The fourth column of Table 9 reports the dimension-specific headcount poverty rates, which give us an idea of the deprivation rates in each dimension.[60] It is evident that majority of the rural Indian population is deprived in three dimensions: sanitation, land, and loan.

[57] See Appendix G.

[58] For detailed description of the related NFHS variables and the corresponding poverty cut-offs, please see Appendix H.

[59] The earlier version of NFHS contained information on how many clothes households in the household owned, but the current version of the survey does not ask that question.

[60] Note that the poverty rates are calculated in terms of the proportion of individuals instead of the proportion of households. We first identify the households that are deprived in a particular dimension and assume that all members in those households are deprived in that dimension. Thus, the poverty rate is the proportion of sample population in the deprived households to the total sample population.

[61] There are 68.9% of households with at least one child in the age group of 5-14 and 19.9% of them contain at least one child laborer.

[62] Out of the 68.9% of households with at least one child in the age group of 5-14 years, 9.04% contain at least one child that does not attend school.

Table 9: NFHS Questions Analogous to BPL Questions and Dimensional Headcount Ratios

	BPL Questions	Relevant NFHS-3 Questions	Dimensions	Headcount (NFHS)
1.	Size group of operational holding of land	Acres of irrigated and un-irrigated agricultural land holdings	Land	70
2.	Type of house	Type of House	Housing	18
3.	Average availability of normal wear clothing	N/A	–	–
4.	Food Security	Body mass index of the respondent	Food Security	44
5.	Sanitation	Type of toilet facility	Sanitation	77
6.	Ownership of Consumer durables	Access to different assets	Asset	31
7.	Literacy status of the highest literate adult	Highest education level attained by the family members	Education	26
8.	Status of the Household Labour Force	Number of hours the children worked for household and non-household members (5-14)	Labour	16[61]
9.	Means of livelihood	Occupation of the respondent and her partner	Occupation	29
10.	Status of children (5-14 years) [any child]	The reason why the children do not go to school (5-14)	Child Status	7[62]
11.	Type of indebtedness	Anyone in the household has a Bank or Post Office account	Loan	64
12.	Reason for migration from household	N/A	–	–
13.	Preference of Assistance	N/A	–	–

As the analysis of poverty in this chapter is multidimensional, one might be interested in the breadth of poverty. A household that is deprived in one dimension may not be deprived in any other dimension. In contrast, a household could be deprived in eight out of ten dimensions. Both households are deprived in at least one dimension. Does it mean that they are equally poor? The answer is indeed no. The breadth of deprivation for the latter household seems more intense. Thus, it would be interesting to explore the breadth of poverty among the rural Indian population. In other words, it would be interesting to see how many people are deprived in one dimension, in two dimensions, and so on. In the first column of Table 10, we report the exact number of dimensions in which any particular household is deprived. For example, 10 percent of the sample are deprived in exactly one dimension (it does not matter which one), and not deprived in the other nine dimensions. The second column reports the percentage of people deprived in exactly that many dimensions. In the third column, we provide a pie-chart to diagrammatically visualize the distribution of the breadth of multi-dimensional poverty.

Table 10: Indicators and Cut-offs of the Chosen Dimensions

Number of Dimensions	0	1	2	3	4	5	6	7	8	9	10	**Total**
Percentage of Poor (%)	3.1	10.0	14.5	17.3	17.6	16.3	11.7	6.4	2.4	0.5	0.1	100

As we can see from Table 10, only 3.1 percent of all rural population is not deprived in any dimension. If identification of the poor is based on the union approach, then 96.9 percent of all rural people live in poverty. Recall that a household is identified as poor by the union approach if it is deprived in at least one dimension, whereas a household is identified as poor according to the inter-section approach if it is deprived in every dimension. Nearly 32 percent of the rural population are deprived in either two or three dimensions. Roughly a third of the rural population is deprived in either four or five dimensions. Also observe that any poverty index based on the intersection approach would judge India as almost poverty free

(0.1%). The BPL process neither follows the union approach nor the intersection approach, but an intermediate approach in a peculiar way. We have already presented the methodology earlier.

Under-coverage Rate and Over-coverage Rate

In comparing our measure with the pseudo-BPL measure, it is useful to identify persons who are classified as poor according to one measure, and non-poor by the other. These can be called over-coverage and under-coverage.

Table 11: Definition of Over-coverage and Under-coverage

		Poor by AF Method		
		Yes	**No**	**Total**
Poor by pseudo-BPL method	**Yes**	p_{yy}	p_{Byn}	p_{By}
	No	p_{Myn}	p_{nn}	p_{Bn}
	Total	p_{My}	p_{Mn}	1

Let us denote the total household population by N. Let the number of poor based on the pseudo-BPL approach be denoted by N_{By} and the number of non-poor be denoted by $N_{Bn} = N - N_{By}$. We define $p_{By} = N_{By}/N$ and $p_{Bn} = N_{Bn}/N$, where p_{By} and p_{Bn} are the proportion of poor and non-poor identified by the pseudo-BPL method. Let the proportion of poor and non-poor identified by the AF method be denoted by p_{My} and p_{Mn}, respectively.

These concepts are summarized in Table 11. The rows denote the proportion of households that are identified as poor versus those identified as non-poor by the pseudo-BPL method. The columns, on the other, denote the proportion of households that are poor versus those are not poor according to the AF method. The following four variables denote the interaction between these two distinct methodologies.

p_{yy} : The proportion of households that are identified as poor by both methodologies

p_{nn} : The proportion of households that are identified as non-poor by both methodologies

p_{Byn} : The proportion of households that are identified as poor by the pseudo-BPL method but are classified as non-poor in terms of the AF method

p_{Myn} : The proportion of households that are identified as non-poor by the pseudo-BPL method but are classified as poor in terms of the AF method

The *under-coverage rate* is defined to be the ratio of the percentage of the sample population that is identified as non-poor by the pseudo-BPL method but are actually classified as poor by the AF method to the percentage of the population that are classified as poor by the AF method. Similarly, the *over-coverage rate* is defined to be the ratio of the percentage of the sample population that are identified as non-poor by the AF method, but are classified as poor by the pseudo-BPL method to the percentage of the population that are identified as poor according to the pseudo-BPL method. Thus, from Table 11, the under-coverage rate is p_{Myn}/p_{My}; whereas, the over-coverage rate is p_{Byn}/p_{By}. Intuitively, if a hundred individuals are identified as poor by the AF method and five of them are misidentified as non-poor by the pseudo-BPL method, then the under-coverage rate is five percent. Similarly, if a hundred individuals are identified as poor by the pseudo-BPL method and ten of them are actually non-poor according to the AF methodology, then the over-coverage rate is ten percent.

Coverage Rates for the Alternative Methodology

In this section, we compare the coverage rates for both methodologies, and find that the AF methodology identifies the poor differently from the BPL methodology. To illustrate the differences in coverage rates, we generate a pseudo-BPL score.[63] The highest possible score for any household is 38. A household is classified as poor based on these ten dimensions if it fails to make a certain score, say z, out of 38 such that $0 \leq z \leq 38$. In Table 12, we summarize the pseudo-BPL poverty rates for various poverty cut-off scores. The first row of Table 12 reports various poverty cut-offs (z). If a household fails to meet a score that is greater than the cut-off, the household is classified as poor (analogous to what is done in the BPL 2002 process). In the second row, we report the poverty rates based on the corresponding poverty cut-off reported in the first row.

Table 12: BPL Poverty Rates Calculated from the NFHS-3 Dataset

Poverty Line (z)	14	15	16	17	18	19	20	21	22	23	24
Pseudo-BPL Pov. Rate (%)	16.8	21.6	26.9	32.9	39.1	44.9	51.1	56.8	62.2	67.2	72.1

For clarity, simplicity, and to match our analysis with the pseudo-BPL identification method, we primarily restrict the analysis to the multidimensional headcount ratio. According to the AF identification methodology, a household is identified as poor if the household is deprived in a certain number or weighted sum of dimensions only. For the purposes of comparison with the existing BPL measure, we further match the BPL assumption of weighting the dimensions equally. Hence, if the second cut-off (k) is, say, four out of ten dimensions, then a household is identified as poor if the household is deprived in at least four dimensions.

[63] To see the score structure of these ten dimensions, please refer to Appendix I.

We present the multidimensional headcount ratio (MD Headcount) in Table 13.

Table 13: India: Multidimensional Poverty Measures

Poverty Cut-Off (k)	**MD Headcount (H)**	**Matched Pseudo-BPL Poverty Rate**	**Under Coverage Rate**	**Over Coverage Rate**
3	0.724	0.721 (z = 24)	5.70%	5.30%
4	0.551	0.568 (z = 21)	7.70%	10.40%
5	0.375	0.391 (z = 18)	12.40%	16.10%
6	0.212	0.216 (z = 15)	20.60%	22.10%
7	0.094	0.092 (z = 12)	33.00%	31.10%

In the first column of the table, we report the second cut-off (k), which establishes the minimum number of dimensions a household must be deprived in order to be considered as poor. In the second column, we report the fraction of people that are deprived in at least that many dimensions. For example, 55 percent of the sample population are poor in at least four out of ten dimensions. If the poverty cut-off is five out of ten dimensions, then 37.5 percent of the sample population are poor.

The next obvious question is how the AF identification method compares to the pseudo-BPL method. In the third column of Table 13, we report the pseudo-BPL poverty rates that match as closely as possible to the corresponding multidimensional poverty rates. For example, the multidimensional poverty rate for $k = 3$ (0.72) is close to the pseudo-BPL poverty rate corresponding to z = 24 (0.72) from Table 12. In the fourth and the fifth columns of Table 13, we report the under-coverage rate and the over-coverage rate for the multidimensional headcount method. This is analogous to what we defined in the last subsection and in Table 11.

The findings are striking. The k cutoff that comes closest to approximating the actual

2002 BPL headcount ratio is $k = 5$. At this headcount, over 12 percent of the poor do not receive BPL cards, and 16 percent of those with BPL cards are not poor. However now we focus on the poorest households among the BPL population – those deprived in 7 or more dimensions ($k = 7$). Here we find that 33 percent of the extreme poor do not receive BPL cards. Whereas we might have expected the persons that were borderline on either measures to be mis-identified, in fact we find that mis-identification increases with the depth of poverty, which is a disturbing feature. More generally, in the fourth column, we report the percentage of the population residing in households that are classified as non-poor by the pseudo-BPL method among the total population residing in households that are identified as poor by the AF method. Similarly, in the fifth column, we report the percentage of the population residing in households that are classified as poor by the pseudo-BPL method but are identified as non-poor by the AF method. The under-coverage rate and the over-coverage rate for M_0 increases because as the cut-off k becomes more stringent, the non-deprived dimensions partially compensate for the deprived dimensions.[64] Even in an environment with no data corruption, the BPL 2002 method would not allocate BPL cards to some of the extreme poor and instead would distribute them among the non-poor.

We can conclude from the analysis in this section that the AF approach is more powerful than the BPL 2002 approach in terms of the identification of poor households. Note that the BPL method has also been criticized due to the data content. It has been argued earlier that the poor households cannot be identified properly even if the methodology is implemented without any corruption. In the next section, we propose choosing the dimensions based on the capability approach. We also propose the adjusted headcount ratio as a measurement of

[64] Note that the under-coverage rate and the over-coverage rate would have been identical if we were able to choose the pseudo-BPL poverty rate as exactly identical to the multidimensional headcount ratio.

overall poverty instead of the multidimensional headcount ratio.

Towards an Improved Measure: Reflecting Multiple Deprivations

In the last section, we matched the dimensions and weights used in the BPL census to identify the poor households. However, as we observed, the BPL census data content and weighting is subject to serious and reasonable criticism. In August 2008, the Deputy Chairperson of the Planning Commission of India stated that an index of deprivation might be constructed to better represent the many faces of poverty (Chauhan, 2008). The dimensions might include education, health, infrastructure, a clean environment, and benefits for women and children – thus some dimensions not used in the 2002 BPL method. Moreover, poverty should be measured by the deprivation of capabilities (Reddy, 2008). Therefore, in this last section, we explore an illustrative improved multidimensional poverty measure that uses existing data, but still might better reflect multiple deprivations across India. Naturally, the choice of dimensions, poverty cut-offs, and weights for such an improved measure are value judgements, and should be influenced by the public debate, as well as by the needs of policy and public sector institutions. If such a set of dimensions were widely agreed on, then it might be a reasonable expectation that accurate and robust measures of all relevant dimensions would be implemented in national survey processes such as the BPL, NSS, and/or NFHS. The process of public discussion and debate, and the enriched data set, would contribute to a measure of poverty that reflects people's multiple deprivations. Using existing data and illustrative dimensions, this final section demonstrates the characteristics of such a measure if it employs the adjusted headcount methodology (M_0) proposed by Alkire and Foster.

Dimensions, Indicators, and Cut-offs

First, we present the tentative dimensions, indicators, and cut-offs that will be used in the following analysis. We use NFHS-3 data to select the indicators for nine dimensions, drawing on the article mentioned above but selecting these indicators merely as an illustrative example. We choose nine dimensions that are based on eleven indicators. We presume that infrastructural facilities should be an important dimension while measuring deprivation and the the dimension consists of two crucial indicators, housing and access to electricity, with equal importance. Similarly, sanitation and access to drinking water together create another important dimension for the same purpose. Other dimensions are measured using only a single indicator. The set of dimensions and the respective indicators are summarized in Table 14 and their detailed descriptions can be found in Appendix J.

In the last columns of Table 14, we report unidimensional headcount ratios. It is evident that most of the rural Indians (77%) in the sample are deprived in sanitation. This is, as might be expected, slightly higher than the national average, which, according to the HDR 2007, was 67 percent. On the contrary, most of the villagers (84%) have access to safe drinking water.

Weighting

We use equal weights, again for illustrative purposes. Note that two of our dimensions have two indicators. Therefore, all of our following nine dimensions receive equal weights of 11/9: living standard, sanitation/water, fuel, asset, education, livelihood, electricity, child status

[65] See Basu and Foster (1998).

[66] Among the 68.9% of households having at least one child in the age group of 5-14, 24.6% of households are deprived in terms of child status.

Table 14: Dimensions, Indicators, and the Headcount Ratios

	Dimensions	**Indicators**	**Headcount (NFHS)**
1.	Living Standard	Housing type	0.18
		Access to electricity	0.44
2.	Health	The minimum BMI of one woman in the household	0.44
3.	Water & Sanitation	Access to improved sanitation	0.77
		Access to improved drinking water source	0.16
4.	Air Quality	Sources of fuel for cooking	0.31
5.	Assets	Asset holding	0.31
6.	Education	Maximum year of education completed by any member[65]	0.26
7.	Livelihood	Occupation of the respondent and her partner	0.29
8.	Child Status	Child labor and/or child school attendance	0.20[66]
9.	Empowerment	Empowerment of women in the household	0.59

and empowerment. The reason behind such a choice is that the total weight sums up to eleven, which is the total number of indicators. We provide a weight of 11/18 to each of the following four indicators: housing, electricity, sanitation, and water.

Table 15: Multidimensional Poverty Measures

Poverty Cut-Off (k)	**Headcount Ratio (HCR)**	M_0	$A = M_0/HCR$
3	0.676	0.308	0.456
4	0.463	0.244	0.527
5	0.275	0.166	0.603
6	0.200	0.128	0.642

Results

In Table 15, we present the number of poor in multiple dimensions, the cut-off based headcount ratios, the adjusted headcount ratios, and average deprivation among the poor using the nested weight. The union approach would identify 92.4 percent of the rural population as poor. On the other hand, the intersection approach leads to an almost poverty free India. If the poverty cut-off is four out of eleven dimensions, 46 percent of the rural population belongs to poor households, which is the multidimensional headcount ratio for this particular cut-off. The main criticisms of the multidimensional headcount ratio are that it does not take into account the breadth of multidimensional poverty, it does not satisfy dimensional monotonicity, and it is not decomposable. Therefore, we propose the adjusted headcount ratio (M_0) as a measure of poverty instead of the multidimensional head count. For the theoretical properties of M_0, see the earlier methodological section.

We use the cut-off of four out of eleven subsequently because the multidimensional headcount ratio of 46 percent is somewhat close to the headcount ratio of 42 percent estimated by the World Bank for a poverty line of \$1.25 per day (Chen and Ravallion, 2005). The third column of Table 15 reports the adjusted headcount poverty rates for different cut-offs. If the poverty cut-off is four out of ten dimensions, then M_0 is 0.244. Recall that $M_0 = HCR \cdot A$. For the poverty cut-off of four out of ten dimensions, HCR is equal to 0.463 and A is equal to $0.244/0.463 = 0.527$. A can be interpreted as the poor being deprived in 52.7 percent of all dimensions on average. If the union approach is employed, then the poor are deprived in 37.9 percent of all dimensions on average. Thus, the fourth column reports the average depth of poverty among the population from the poor households.

Until now, our discussion was at the country level. We now move to state level analysis. In

Table 16: State-wise Decomposition of Poverty for 4/11 Cut-off

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
States	**Populn. Share of States**	**Headcount Ratio (H)**	**HC Rank**	**M_0 Poverty Ratio**	**M_0 Rank**	**NSS Income Poverty**[67]	**NSS Rank**
Kerala	2.41%	0.056	1	0.026	1	0.132	6
Sikkim	0.06%	0.073	2	0.033	2	0.223	14.5
Mizoram	0.05%	0.088	3	0.04	3	0.223	14.5
Himachal Pradesh	0.73%	0.1	5	0.046	4	0.107	4
Manipur	0.18%	0.1	6	0.046	5	0.223	14.5
Goa	0.07%	0.098	4	0.049	6	0.054	2
Punjab	2.25%	0.149	7	0.071	7	0.091	3
Nagaland	0.13%	0.161	8	0.079	8	0.223	14.5
Tripura	0.41%	0.227	9	0.114	9	0.223	14.5
Jammu & Kashmir	0.88%	0.242	10	0.116	10	0.046	1
Uttaranchal	0.82%	0.244	11	0.118	11	0.408	25
Meghalaya	0.25%	0.258	12	0.129	12	0.223	14.5
Tamil Nadu	3.72%	0.293	13	0.142	13	0.228	19
Haryana	2.10%	0.306	14	0.152	14	0.136	7
Gujarat	4.14%	0.325	15	0.159	15	0.191	9
Karnataka	4.80%	0.345	17	0.172	16	0.208	10
Maharashtra	6.82%	0.342	16	0.173	17	0.296	21
Andhra Pradesh	6.79%	0.382	18	0.192	18	0.112	5
Arunachal Pradesh	0.11%	0.388	19	0.203	19	0.223	14.5
Assam	2.94%	0.395	20	0.205	20	0.223	14.5
West Bengal	8.54%	0.466	21	0.246	21	0.286	20
Bihar	10.62%	0.503	22	0.254	22	0.421	26
Chhattisgarh	2.62%	0.541	25	0.281	23	0.408	24
Rajasthan	6.51%	0.535	23	0.286	24	0.187	8
Orissa	4.23%	0.537	24	0.288	25	0.468	28
Uttar Pradesh	17.86%	0.612	26	0.332	26	0.334	22
Madhya Pradesh	6.97%	0.629	27	0.344	27	0.369	23
Jharkhand	2.97%	0.823	28	0.489	28	0.463	27
India	**-**	**0.463**	**-**	**0.244**	**-**	**0.283**[68]	**-**

our NFHS sub-sample, India has twenty eight states. Table 16 ranks states according to their adjusted headcount poverty ranks, where a household is identified as poor if it is deprived in four out of eleven dimensions. Kerala has the least poverty and Sikkim, a state in the eastern part of India, registers the second lowest poverty rate according to the M_0 measure. Jharkhand, where more than eighty percent of population are identified as members of poor households, ranks last. The overall M_0 ranks for states do not vary significantly from the headcount ranks. Spearman's rank correlation coefficient between these two rankings is 0.99. Conversely, the M_0 rank and the NSS income poverty rank among states varies significantly. Spearman's rank correlation coefficient between these two rankings is merely 0.58. Andhra Pradesh, which ranks fifth in terms of the NSS income poverty line, ranks eighteenth in terms of the adjusted headcount ratio. Similarly, Rajasthan ranks eighth in terms of the NSS income poverty but twenty-fourth in terms of M_0.

Table 17: Poverty Decomposition by Dimensions

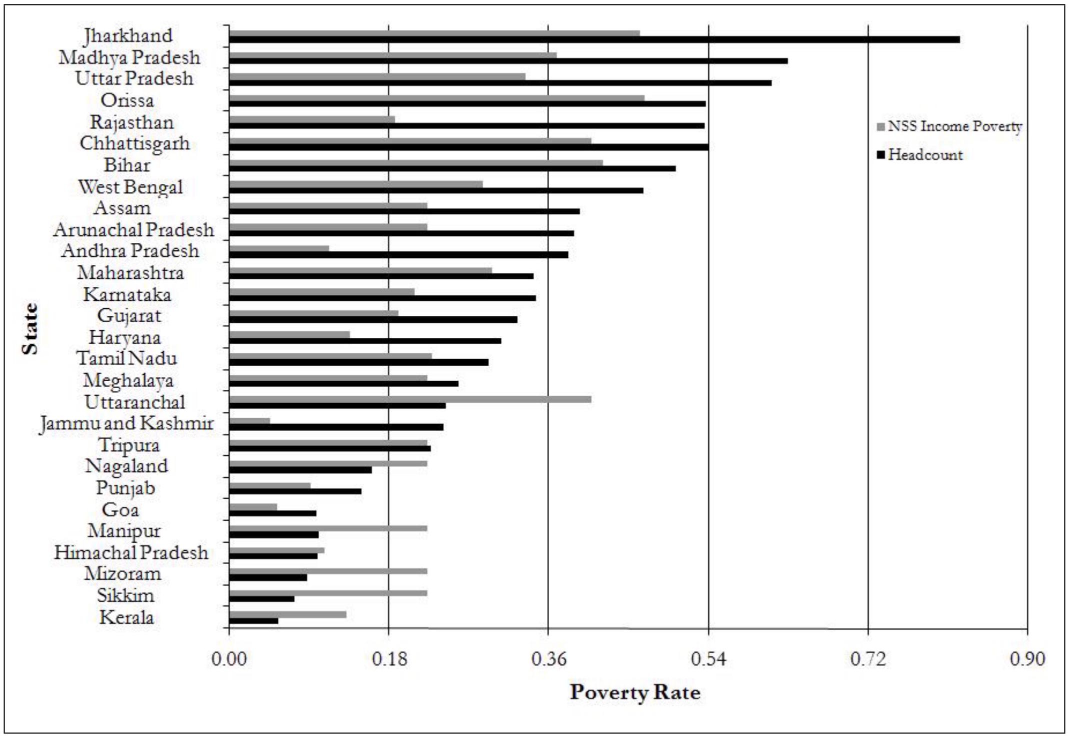
M_0 Rank	State	House	Elect.	Health	Sanit.	Water	Fuel	Asset	Educ.	Liveli.	Child Sta.	Emp.	M_0
1	**Kerala**	0.012	0.034	0.036	0.021	0.031	0.022	0.043	0.004	0.028	0.012	0.038	**0.026**
	Break Down	*2.7%*	*7.4%*	*15.4%*	*4.5%*	*6.7%*	*9.4%*	*18.4%*	*1.9%*	*12.1%*	*5.2%*	*16.3%*	***100%***
2	**Sikkim**	0.032	0.038	0.021	0.057	0.036	0.044	0.023	0.059	0.018	0.025	0.026	**0.033**
	Break Down	*5.4%*	*6.4%*	*7.1%*	*9.6%*	*6.1%*	*15.0%*	*7.7%*	*19.7%*	*6.0%*	*8.5%*	*8.6%*	***100%***
21	**West Bengal**	0.161	0.4	0.326	0.374	0.056	0.11	0.363	0.285	0.208	0.12	0.305	**0.246**
	Break Down	*3.6%*	*9.1%*	*14.7%*	*8.5%*	*1.3%*	*5.0%*	*16.4%*	*12.9%*	*9.4%*	*5.4%*	*13.8%*	***100%***
22	**Bihar**	0.273	0.465	0.358	0.488	0.032	0.249	0.038	0.297	0.173	0.174	0.371	**0.254**
	Break Down	*6.0%*	*10.2%*	*15.6%*	*10.7%*	*0.7%*	*10.9%*	*1.7%*	*13.0%*	*7.5%*	*7.6%*	*16.2%*	***100%***
25	**Orissa**	0.27	0.418	0.344	0.528	0.147	0.286	0.197	0.282	0.213	0.14	0.446	**0.288**
	Break Down	*5.2%*	*8.1%*	*13.3%*	*10.2%*	*2.8%*	*11.0%*	*7.6%*	*10.9%*	*8.2%*	*5.4%*	*17.2%*	***100%***
28	**Jharkhand**	0.064	0.697	0.488	0.813	0.472	0.544	0.696	0.375	0.492	0.259	0.52	**0.489**
	Break Down	*0.7%*	*7.9%*	*11.1%*	*9.2%*	*5.4%*	*12.4%*	*15.8%*	*8.5%*	*11.2%*	*5.9%*	*11.8%*	***100%***
	India	**0.145**	**0.311**	**0.289**	**0.439**	**0.096**	**0.24**	**0.263**	**0.22**	**0.218**	**0.152**	**0.319**	**0.244**
	Break Down	*3.3%*	*7.1%*	*13.2%*	*10.0%*	*2.2%*	*11.0%*	*12.0%*	*10.0%*	*9.9%*	*6.9%*	*14.5%*	***100%***

After we compare the ranks of states under different methodologies, it is interesting to analyze the source and contribution of different dimensions to overall poverty. In Table 17,

[67] We report the poverty rates based on Uniform Recall Period (URP) rather than the Mized Recall Period (MRP) since the URP method is the same as the traditional method used in 1993-94 and different from the method pursued in 1999-00. The MRP based method yielded an extremely low level of rural poverty (22%). See Government of India (2007)

[68] Out of the 68.9 percent of households containing at least one child within the age group of 5-14 years, 9.04 percent contain at least one child that does not attend school.

Figure 6: Adjusted Headcount Poverty Ranking Vs. NSS Income Poverty Ranking



we present the decomposition of poverty across different dimensions. It is evident that Sikkim and Kerala have almost the same M_0 poverty rates, but the source differs radically. For example, the contribution of the education dimension towards the overall poverty in Kerala is merely 1.9 percent. On the contrary, the contribution of education to Sikkim poverty is nearly 20 percent. Kerala also performs better in terms of sanitation and fuel, but performs much worse in nutrition, assets, and livelihood compared to Sikkim. West Bengal and Bihar comparisons are more similar, although stark differences appear with respect to assets, where West Bengal is much worse, and with respect to clean air, where Bihar performs poorly. Comparing Orissa and Jharkhand, we find that women's disempowerment is starkly more prominent in Orissa, where poverty is also more strongly driven by poor housing and nutrition. Jharkhand has far higher contributions to poverty from poor asset holdings and livelihoods. This type of decomposition enables policy makers to make proper policy recommendations.

To have a graphical visualization of the difference in ranking between the Alkire and Foster methodology and the NSS income poverty ranking, see Figure 6.[69]

Table 18: Spearman's Rank Correlation Matrix for Different M_0 Rankings

Cut-off (k)	**3**	**4**	**5**	**6**	**7**
4	1	-	-	-	-
5	0.99	1	-	-	-
6	0.99	1	**1**	-	-
7	0.97	0.97	**0.98**	0.98	-
8	0.96	0.96	**0.96**	0.97	0.98

The final concern is about how robust the poverty rankings are for varying cut-offs. One might argue that the choice of cut-off is arbitrary and might wonder if the M_0 rankings change drastically due to a change in the cut-off. To address this legitimate query, we calculate the M_0 measures for all states for different cut-offs and then we calculate the Spearman's rank correlation coefficients between each pair of rankings for k = 3, . . .,8. From Table 18, it can be seen that the minimum correlation is 0.98 between $k = 3$ and $k = 8$. Therefore, we can conclude that the rankings for varying poverty cut-offs are highly robust. We did not calculate the rankings beyond $k = 8$ because the value of M_0 is very low and with so few observations the rankings could be biased.

[69] It can be seen from Tables 9 and 15 that the multidimensional headcount ratio for $k = 5$ (28%) is very close to the NSS 2004-05 poverty rate (28%). Therefore, a comparison of the rankings for $k = 5$ would have made more sense. However, a subsequent analysis of rank correlation between the rankings generated by various k values (Table 18) ensures that a choice of different k would not alter our analysis.

Conclusion

This chapter first identified the various criticisms that have been leveled against the 2002 Below the Poverty Line (BPL) measure in rural India. The criticisms fall into three kinds: (i) problems of data quality, data coverage, and corruption, (ii) problems with the aggregation method, and (iii) issues of data content and periodicity. This chapter endeavours to isolate the criticisms about the identification methodology, because errors of 12-33 percent would remain due to these problems even if the data were strong.

To address the problems of *identification* and *aggregation*, using the same NFHS matching dimensions, we applied dimension-specific cut-offs, and computed a multidimensional headcount and adjusted headcount measure (M_0), using the methodology proposed by Alkire and Foster (2008). The resulting measure – which matched the BPL dimensions but with better data and a more defensible aggregation technique – was then compared with the poverty status identified by a pseudo-BPL approach at the national level. Significant differences appeared, with under-coverage and over-coverage rates of up to 33 percent, which, despite the differences in dimensions, bears consideration. We also illustrated the policy value of having an aggregation method that generates decomposable multidimensional poverty measures because they can immediately reveal to any policy maker the poverty priorities in her or his area. If census data were available, such a measure could be calculated at the local level or for different population groups, so as to identify local priorities for public investment and hence to inform multisectoral planning.

Finally, this chapter addressed the issue of *data content*, and also sought to affirm the possibility of a multidimensional index that transparently represents the multiple deprivations people suffer. Naturally, the final selection of dimensions, weights, and cutoffs for a national poverty measure requires significant public discussion as well as the generation of new data to match the dimensions of interest. However, for illustrative purposes, we tentatively selected nine dimensions, and eleven indicators, that may improve upon the BPL dimensions. We included empowerment because of its intrinsic importance, although data for this dimension remains weak.

The nine dimensions were – living standard (housing, electricity), health, water and sanitation, air quality the household members breathe, assets, education, livelihood, child status, and empowerment. We compute our measure using these dimensions, compare it with the 2004/5 NSS levels, and decompose it by state. The results are striking and informative. For example, multidimensional poverty in Jharkhand is driven by asset deprivation, low air quality, and poor quality of work, with nutritional deficits and disempowerment also contributing significantly. In Gujarat, nutrition ranks as the leading contributor to poverty, followed by deprivations in women's empowerment and air quality.

While clearly further analysis is required, the multidimensional poverty methodology implemented in this chapter can be used not only to identify the poor (as the NSS or BPL do), but also to see easily what dimensions are most important for multidimensional poverty among different groups of people.

Appendix

G. Below Poverty Line Survey Questions (2002)

Sl. No	Characteristics/ Questions	Scores: 0	1	2	3	4
1	**Size group of operational holding of land**	Nil	Less than 1 ha of un-irrigated land (or less than 0.5 ha of irrigated land)	1-2 ha of un-irrigated land (or 0.5-1 ha of irrigated land)	2 -5 ha of un-irrigated land (or 1.0 -2.5 ha of irrigated land)	More than 5 ha of un-irrigated land (or 2.5 ha of irrigated land)
2	**Type of house**	Houseless	Kutcha	Semi-pucca	Pucca	Urban type
3	**Average availability of normal wear clothing (per household in pieces)**	Less than 2	2 or more, but less than 4	4 or more, but less than 6	6 or more, but less than 10	10 or more
4	**Food Security**	Less than one square meal per day for major part of the year	Normally, one square meal per day, but less than one square meal occasionally	one square meal per day throughout the year	two square meals per day with occasional shortage	Enough food throughout the year
5	**Sanitation**	Open defection	Group latrine with irregular water supply	Group latrine with regular water supply	Clean group latrine with regular water supply and regular sweeper	Private latrine
6	**Ownership of Consumer durables: Do you own (tick) – TV, electric fan, radio, pressure cooker**	Nil	Any one	Two items only	Any three or all items	All items and/or any one of the following items - computer, telephone, refrigerator, colour TV, electric kitchen appliances, expensive furniture, LMV@/ LCV@, tractor, mechanised two-wheeler/ three-wheeler, power tiller, combined thresher/ hawestor [@ 4-wheeled mechanised vehicle]
7	**Literacy status of the highest literate adult**	Illiterate	Upto Primary (Class V)	Completed Secondary (Passed Class X)	Graduate/ Professional Diploma	Post Graduate/ Professional Graduate
8	**Status of the Household Labour Force**	Bonded labor	Female and children labor	Only adult females and no child labor	Adult males only	Others
9	**Means of livelihood**	Casual Labor	Subsistence cultivation	Artisan	Salary	Others
10	**Status of children (5-14 years) [any child]**	Not going to school and working	Going to School and working			Going to school and not working
11	**Type of indebtedness**	For daily consumption purposes from informal sources	For production purpose from informal sources	For other purpose from informal sources	Borrowing only from Institutional agencies	No indebtedness and possess assets
12	**Reason for migration from household**	Casual work	Seasonal employment	Other forms of livelihood	Non-migrant	Other purposes
13	**Preference of Assistance**	Wage Employment/ TPDS (Targeted Public Distribution System)	Self Employment	Training and Skill Upgradation	Housing	Loan/Subsidy more than Rs. One lakh or No assistance needed

*Source: Government of India, Ministry of Rural Development (2002) and Sundaram (2003)

H. Dimensions, Indicators, and Poverty Cut-Offs Analogous to Year 2002 BPL Questions

1. Land: Acres of irrigated and un-irrigated agricultural land holdings

   This dimension corresponds to Question 1 in the BPL questionnaire and is asked directly in the NFHS-3 survey.

   Question HV244: If owns land usable for agriculture

   Question SH60H: Hectares of agricultural land holding

   Question SH61H: Hectares of land irrigated

   Poverty Cut-off - Less than one hectare of un-irrigated land and 0.5 hectare of irrigated land

2. Housing: Type of House

   This dimension corresponds to Question 2 in the BPL questionnaire and is asked directly in the NFHS-3 survey.

   Question SHNFHS2: House type (Kachha, Semi-pucca, Pucca)

   Poverty Cut-off - Live in a Kachha House

3. Land: Acres of irrigated and un-irrigated agricultural land holdings

   This dimension corresponds to Question 2 in the BPL questionnaire that asks how many times the household eats during a day. The NFHS-3 does not contain this question, but it does collect information on nutritional intake and the body mass index (BMI) of the respondents in the household. We prefer BMI to the nutritional intake of the respondents not merely for convenience, but also for the following reasons. First, it is difficult to match the BPL question with NFHS questions regarding specific food types consumed. Second, the body mass index directly represents the nutritional state of a household – which is arguably the desired outcome for which the BPL meal resources are a proxy. Note that BMI data are present for the female only, which is not optimal, but may be acceptable because of the importance of women's health in general. Also, malnutrition among women has not improved over the past decade despite a high rate of growth and reduction in income poverty. See Jose and Navaneetham (2008).

Question V445: Body mass index for the female respondent

Poverty Cut-off - The minimum BMI of the women in the household is less than 18.5 Kg/m^2

4. Sanitation: Type of toilet facility

This dimension corresponds to Question 5 in the BPL questionnaire and is asked directly in the NFHS-3 survey.

Question HV205: Type of toilet facility (1. Flush - to piped sewer system, 2. Flush - to septic tank, 3. Flush - to pit latrine, 4. Flush - to somewhere else, 5. Flush - don't know where, 6. Pit latrine – ventilated, 7. Pit latrine - with slab, 8. Pit latrine - without slab, 9. No facility/uses bush/field, 10. Composting toilet, 11. Dry toilet, 96. Other)

Poverty Cut-off - Uses Pit latrine – w/o slab, No facility/uses bush/field, Composting toilet, Dry toilet, OTHER

5. Asset: Access to different assets

This dimension corresponds to Question 6 in the BPL questionnaire and the NFHS-3 collects information on the ownership of most of these items.

Question SH47B: Has mattress	Question SH47V: Has thresher
Question SH47C: Has pressure cooker	Question SH47W: Has tractor
Question SH47F: Has table	Question HV207: Has radio
Question SH47G: Has electric fan	Question HV209: Has refrigerator
Question SH47I: Has black & white TV	Question HV211: Has motor cycle
Question SH47J: Has colour	Question HV212: Has car
Question SH47N: Has computer	Question HV221: Has phone

Poverty Cut-off - Owns any one of the following assets: a b/w television, an electric fan, a pressure cooker, or a radio. At the same time, does not own any of the following assets: a refrigerator, a motor cycle, a car, a phone, a mattress, a table, a colour TV, a computer, a thresher, or a tractor.

6. Education: Highest education level attained by the family members

   This dimension corresponds to Question 7 in the BPL questionnaire and the NFHS-3 survey contains enough information to replicate this dimension.

   Question HV108: Education completed in single years

   Poverty Cut-off - Maximum year of education completed by any member is less than 5 years

7. Labor: Number of hours the children worked for household and non-household members [age: 5-14]

   This dimension corresponds to Question 8 in the BPL questionnaire that asks about bonded labour and the labour status of women and children in the household, implying that a household is most deprived if any worker is bonded, or if women and the children work. The NFHS-3 does not have data on bonded labour. Further, many would dispute the view that women's work-force participation should be treated as a deprivation. However, there is widespread agreement in treating a child's labour force participation as a deficiency for the household. Therefore, we substitute the eighth BPL question by the dimension named 'existence of child labour in the household within the age group of 5-14'.

Question SH24: In past week, number of hours worked for non-HH member [age 5-14]

Question SH27: In past week, number of hours helped with HH chores [age 5-14]

Question SH29: In past week, number of hours did other family work [age 5-14]

Question HV105 Age of household members

Poverty Cut-off - There is at least one incidence of child labour within the age group of 5-14.

8. Occupation: Occupation of the respondent and her partner

This dimension corresponds to Question 9 in the BPL questionnaire that asks respondents to categorize the means of livelihood for the family. The NFHS survey contains enough information to identify a household by the major occupation of its members.

Question V716: Respondent's occupation

Question V704: Partner's occupation

Poverty Cut-off - The respondent and her partner both fall into the following occupation categories: unemployed, agricultural labourer, plantation labourers, simply labourers, and new workers seeking jobs

9. Child Status: The reason why the children do not go to school (5-14)

This dimension corresponds to Question 10 in the BPL questionnaire that asks about the status of children in the household – whether they are in school and whether they are working. We have already created a dimension on child labour. Therefore, we replicate the tenth question by creating a dimension based only on whether the children in the age group of 5-14 go to school.

Question SH22: Main reason not attending school [age 5-18] (1. School too far away, 2. Transport not available, 3. Further education not considered, 4. Required for household work, 5. Required for work on farm, 6. Required for outside work, 7. Costs too much, 8. No proper school facilities, 9. Not safe to send girls, 10. No female teacher, 11. Required for care of sibling, 12. Not interested in studies, 13. Repeated failures, 14. Got married, 15. Did not get admission, 96. Other)

Question HV105: Age of household members

Poverty Cut-off - A household is classified as deprived in the child-status dimension, if any of the children in the age group of 5-14 does not go to school for any reason.

10. Loan: Any one in the household has a Bank or Post Office account

This dimension corresponds to Question 11 in the BPL questionnaire that asks for what purposes the household has become indebted and whether the loan is from an informal sector or from institutional agencies. The NFHS does not contain analogous questions but it has information on whether any member of the household has a bank or a postal account. A household that has access to such account is more likely to obtain an institutional loan, but a household without it is more inclined to obtain loan from an informal sector, if at all.

Question HV247: Owns a bank account or post office account

Poverty Cut-off - None of the household members holds a bank or post office account

I. Score Structure of the Ten Matched NFHS-3 Questions

Sl No	Characteristic/ Questions	Scores 0	1	2	3	4
1	**Size group of operational holding of land**	Nil	Less than 1 ha of un-irrigated land (or less than 0.5 ha of irrigated land)	1-2 ha of un-irrigated land (or 0.5-1 ha of irrigated land)	2 -5 ha of un-irrigated land (or 1.0 -2.5 ha of irrigated land)	More than 5 ha of un-irrigated land (or 2.5 ha of irrigated land)
2	**Type of house**		Kutcha	Semi-pucca	Pucca	
3	**Minimum BMI of the respondent in the household**	Less than 16 Kg/m2	Higher than 16 Kg/m2 but less than 18.5 Kg/m2		Higher than 18.5 Kg/m2	
4	**Sanitation**	No facility/uses bush/field or others	Composting toilet or Dry toilet or share the following type of facilities with others: Pit latrine – ventilated, Pit latrine - with slab, Pit latrine - without slab	Pit latrine - without slab or share the following facilities with others: Flush - to piped sewer system, Flush - to septic tank, Flush - to pit latrine, Flush - to somewhere else, Flush - don't know where	Pit latrine – ventilated, Pit latrine - with slab	Flush - to piped sewer system, Flush - to septic tank, Flush - to pit latrine, Flush - to somewhere else, Flush - don't know where
5	**Ownership of Consumer durables: Do you own (tick) – B/W TV, electric fan, radio, pressure cooker**	Nil	Any one	Two items only	Any three or all items	All items and/or any one of the following items - refrigerator, motor cycle, car, phone, mattress, table, colour TV, computer, thresher, and tractor
6	**Literacy status of the highest literate adult**	Illiterate	Upto Primary (Class V)	Completed Secondary (Passed Class X)	Graduate/ Professional Diploma	Post Graduate/ Professional Graduate
7	**Status of the Household Labour Force**	Only children work and no adult work or no one works	Female and child labor	Only adult females and no child or adult male works	Adult males only	Both adult male and adult female work but no child works
8	**Means of livelihood**	Laborer, others, and no occupation	Agricultural laborer and Plantation laborers	Other unskilled and manual except laborer	Clerical and Salary	Professional, Technical, Management, Sales, other agricultural employee
9	**Status of children (5-14 years) [any child]**	Not going to school irrespective of working	Going to School and working			Going to school and not working
10	**Bank Account**	No one in the household has bank account				Has bank account

J. Dimensions, Indicators, and Cut-Offs for the Deprivation Measure[70]

1. Living Standard

   The first dimension represents the living standard of the households. The indicators used to measure this dimension are the type of house and access to electricity.[71]

   Question SHNFHS2: House type (Kachha, Semi-pucca, Pucca)

   Question HV206: Has electricity

   Poverty Cut-off – (i) A household is deprived in terms of housing if the household lives in a kaccha house. (ii) A household is deprived of electricity if it does not have access to electricity.

2. Health

   This dimension is same as the food security dimension (3) in Appendix H..

   Question V445: Body mass index for the female respondent.

   Poverty Cut-off – The minimum BMI of the women in the household is less than 18.5 Kg/m^2.

3. Water and Sanitation

   This dimension measures the quality of a household's access to water and sanitation.

   Question HV201: Source of drinking water

   Question HV205: Type of toilet facility

   Poverty Cut-off – (i) A household is classified as deprived in terms of access to safe drinking water supply if the sources of water are a unprotected well and spring, river,

[70] The following questions or indicators were gathered from the NFHS-3 questionnaire. Poverty cut-off denotes the situation under which a household is deprived in that dimension.

[71] The NFHS-3 dataset does not allow us to incorporate the size of the house, which might be an important factor. We do not rely on land holding since the quality of land differs from place to place and not all households own land.

dam, lake, ponds, stream, tanker truck, cart with small tank, bottled water, other. (ii) A household is classified as deprived in the sanitation dimension if the household uses one of the following: pit latrine without slab, no facility/uses bush/field, composting toilet, dry toilet, other.

4. Air Quality

   More than 90 percent of the rural households use solid waste matter as their source of fuel while cooking. But the use of solid waste matter is harmful for the environment and indeed harmful for household members if they breathe it regularly.[72] Some rural households cook outside or in a separate building; others cook inside, but some, unfortunately, do not have a separate room for cooking. The households that cook inside their living room using solid waste matters face clear respiratory hazards.

   Question HV242: Household has separate room used as kitchen

   Question HV226: Type of cooking fuel (1. Electricity, 2. LPG/Natural gas, 4. Biogas, 5. Kerosene, 6. Coal, lignite, 7. Charcoal, 8. Wood, 9. Straw/shrubs/grass, 10. Agricultural crop , 11. Animal dung, 96. Other)

   Poverty Cut-off – The household does not have a separate room used as kitchen and the sources of fuel are coal, lignite, charcoal, wood, Straw/shrubs/grass, agricultural crop, animal dung, and other.

5. Assets

   This dimension is same as the Asset dimension (5) in Appendix H.

[72] See Duflo et al. (2008).

Question SH47B: Has mattress	Question SH47V: Has thresher
Question SH47C: Has pressure cooker	Question SH47W: Has tractor
Question SH47F: Has table	Question HV207: Has radio
Question SH47G: Has electric fan	Question HV209: Has refrigerator
Question SH47I: Has black & white TV	Question HV211: Has motor cycle
Question SH47J: Has colour	Question HV212: Has car
Question SH47N: Has computer	Question HV221: Has phone

Poverty Cut-off - Owns any one of the following assets: a b/w television, an electric fan, a pressure cooker, and a radio. At the same time, does not own any of the following assets: a refrigerator, a motor cycle, a car, a phone, a mattress, a table, a colour TV, a computer, a thresher, and a tractor

6. Education

This dimension is same as the Asset dimension (6) in Appendix H.

Question HV108: Education completed in single years

Poverty Cut-off - Maximum year of education completed by any member is less than 5 years

7. Livelihood

This dimension is same as the Occupation dimension (8) in Appendix H.

Question V716: Respondent's occupation

Question V704: Partner's occupation

Poverty Cut-off – The respondent and her partner both fall into the following occupation categories: unemployed, agricultural labourer, plantation labourers, simply labourers, and new workers seeking jobs.

8. Child Status

For any country, one of the biggest assets is the children. Therefore, we incorporate a dimension regarding the status of the child. This dimension consists of the labour status and school attendance status of the children.

Question SH24: In past week, number of hours worked for non-HH member [age 5-14]

Question SH27: In past week, number of hours helped with HH chores [age 5-14]

Question SH29: In past week, number of hours did other family work [age 5-14]

Question HV105: Age of household members

Question SH22 Main reason not attending school [age 5-18] (1. School too far away, 2. Transport not available, 3. Further education not considered, 4. Required for household work, 5. Required for work on farm, 6. Required for outside work, 7. Costs too much, 8. No proper school facilities, 9. Not safe to send girls, 10. No female teacher, 11. Required for care of sibling, 12. Not interested in studies, 13. Repeated failures, 14. Got married, 15. Did not get admission, 96. Other)

Poverty Cut-off – There is at least one incidence of child labour and/or at least one child aged 5-14 does not attend school.[73]

9. Women's Empowerment

The final dimension is the empowerment of women. It has been very difficult to find a variable that adequately represents the empowerment of women. In the NFHS-3 sample survey, respondents were asked several questions related to empowerment and violence, such as: 1) if the woman faces severe, less severe, emotional, or sexual violence; 2) if the

[73] The NFHS-3 does not allow us to incorporate the labor status of the children in the age group of 15-18. Also, the households that do not have any child are assumed not to be deprived in this dimension.

woman has the final say in household decision making; 3) when the women respondent justifies beating; and 4) if the women is allowed to freely go to certain places. The first two sets of question reduce the number of observations drastically. Given that on some occasions other households were present in the household during interview, the fourth question seems to be a better proxy for woman empowerment than the third as it is more objective. The fourth question asks if they are freely allowed to go to certain places, like, a market, a health facility, and out of the village. We use this dimension but acknowledge that stronger data are necessary to reflect the degree and kinds of empowerment among all household members.

Question S824A: Allowed to go to: market (1. Alone, 2. With someone else only, 3. Not at all)

Question S824B: Allowed to go to: health facility (1. Alone, 2. With someone else only, 3. Not at all)

Question S824C: Allowed to go to: places outside this village/community (1. Alone, 2. With someone else only, 3. Not at all)

Poverty Cut-off – If any woman in the household does not have the right to go alone to the market, a health facility, and somewhere outside of the village.

K. Weighted and Unweighted Population

State	Dataset Comparing BPL and M_0		Dataset for the Deprivation Measure	
	Number of Observations	Weighted by Population	Number of Observations	Weighted by Population
Andhra Pradesh	8,415	16,235	8,455	16,357
Arunachal Pradesh	3,972	250	4,149	262
Assam	8,648	7,009	8,725	7,091
Bihar	9,449	25,449	9,470	25,577
Chhattisgarh	9,310	6,231	9,392	6,304
Goa	4,623	162	4,837	170
Gujarat	7,656	9,694	7,849	9,966
Haryana	8,214	4,997	8,272	5,046
Himachal Pradesh	7,388	1,732	7,476	1,757
Jammu and Kashmir	7,066	2,062	7,267	2,126
Jharkhand	7,404	7,126	7,409	7,151
Karnataka	12,830	11,381	12,990	11,555
Kerala	7,317	5,709	7,405	5,794
Madhya Pradesh	12,352	16,662	12,399	16,772
Maharashtra	9,443	16,218	9,537	16,425
Manipur	7,681	432	7,855	443
Meghalaya	4,123	594	4,190	605
Mizoram	2,857	130	2,877	131
Nagaland	6,584	309	6,607	311
Orissa	10,171	10,005	10,326	10,186
Punjab	8,401	5,326	8,520	5,416
Rajasthan	10,652	15,574	10,694	15,679
Sikkim	3,959	141	3,978	142
Tamil Nadu	8,292	8,907	8,324	8,967
Tripura	4,555	973	4,597	985
Uttar Pradesh	27,862	42,550	28,088	43,014
Uttaranchal	7,546	1,967	7,600	1,986
West Bengal	11,408	20,355	11,493	20,564
India	**238,178**	**238,178**	**240,781**	**240,781**

L. Contribution of Each Dimension by State: M_0 for Multiple Deprivations, $k = 4$, Nested Weights

M0 Rank	States	Popul.	House	Electri.	Health	Sanit.	Water	Fuel	Asset	Educat.	Liveli.	Child Status	Empow.	M_0
1	**Kerala**	5,794	0.012	0.034	0.036	0.021	0.031	0.022	0.043	0.004	0.028	0.012	0.038	0.026
-	*Break Down*	-	*2.65%*	*7.38%*	*15.41%*	*4.45%*	*6.74%*	*9.42%*	*18.43%*	*1.92%*	*12.09%*	*5.17%*	*16.34%*	*100%*
2	**Sikkim**	142	0.032	0.038	0.021	0.057	0.036	0.044	0.023	0.059	0.018	0.025	0.026	0.033
-	*Break Down*	-	*5.42%*	*6.35%*	*7.11%*	*9.56%*	*6.05%*	*14.98%*	*7.70%*	*19.72%*	*6.01%*	*8.46%*	*8.63%*	*100%*
3	**Mizoram**	131	0.073	0.069	0.035	0.041	0.061	0.083	0.012	0.059	0.02	0.019	0.01	0.04
-	*Break Down*	-	*10.18%*	*9.55%*	*9.65%*	*5.74%*	*8.44%*	*22.96%*	*3.28%*	*16.50%*	*5.50%*	*5.40%*	*2.80%*	*100%*
4	**Himachal Pradesh**	1,757	0.015	0.016	0.059	0.098	0.033	0.035	0.062	0.026	0.058	0.018	0.075	0.046
-	*Break Down*	-	*1.81%*	*1.91%*	*14.19%*	*11.84%*	*3.97%*	*8.49%*	*15.07%*	*6.35%*	*13.93%*	*4.25%*	*18.18%*	*100%*
5	**Manipur**	443	0.058	0.045	0.04	0.075	0.081	0.058	0.027	0.047	0.012	0.039	0.061	0.046
-	*Break Down*	-	*6.97%*	*5.44%*	*9.55%*	*9.00%*	*9.78%*	*14.07%*	*6.51%*	*11.43%*	*3.01%*	*9.43%*	*14.81%*	*100%*
6	**Goa**	170	0.013	0.034	0.069	0.086	0.057	0.049	0.05	0.042	0.067	0.019	0.054	0.049
-	*Break Down*	-	*1.51%*	*3.79%*	*15.52%*	*9.69%*	*6.46%*	*11.11%*	*11.25%*	*9.34%*	*15.06%*	*4.23%*	*12.04%*	*100%*
7	**Punjab**	5,416	0.026	0.025	0.074	0.121	0.004	0.07	0.047	0.071	0.113	0.058	0.114	0.071
-	*Break Down*	-	*2.05%*	*1.95%*	*11.72%*	*9.49%*	*0.35%*	*10.98%*	*7.47%*	*11.24%*	*17.77%*	*9.10%*	*17.88%*	*100%*
8	**Nagaland**	311	0.129	0.119	0.066	0.125	0.08	0.059	0.063	0.114	0.02	0.055	0.112	0.079
-	*Break Down*	-	*9.02%*	*8.29%*	*9.27%*	*8.77%*	*5.57%*	*8.23%*	*8.74%*	*15.96%*	*2.75%*	*7.72%*	*15.68%*	*100%*
9	**Tripura**	985	0.09	0.171	0.149	0.087	0.13	0.074	0.122	0.132	0.101	0.071	0.139	0.114
-	*Break Down*	-	*4.37%*	*8.30%*	*14.53%*	*4.25%*	*6.34%*	*7.21%*	*11.82%*	*12.84%*	*9.86%*	*6.92%*	*13.56%*	*100%*
10	**Jammu & Kashmir**	2,126	0.089	0.056	0.128	0.238	0.123	0.085	0.062	0.089	0.171	0.097	0.162	0.116
-	*Break Down*	-	*4.23%*	*2.67%*	*12.19%*	*11.38%*	*5.90%*	*8.13%*	*5.96%*	*8.52%*	*16.29%*	*9.31%*	*15.44%*	*100%*
11	**Uttaranchal**	1,986	0.079	0.135	0.147	0.231	0.071	0.108	0.092	0.065	0.101	0.103	0.187	0.118
-	*Break Down*	-	*3.70%*	*6.38%*	*13.86%*	*10.88%*	*3.33%*	*10.19%*	*8.70%*	*6.15%*	*9.56%*	*9.65%*	*17.62%*	*100%*
12	**Meghalaya**	605	0.14	0.21	0.07	0.188	0.184	0.055	0.118	0.202	0.15	0.085	0.124	0.129
-	*Break Down*	-	*6.00%*	*8.99%*	*5.98%*	*8.05%*	*7.92%*	*4.73%*	*10.16%*	*17.31%*	*12.88%*	*7.33%*	*10.65%*	*100%*
13	**Tamil Nadu**	8,967	0.111	0.096	0.173	0.287	0.024	0.16	0.242	0.116	0.197	0.055	0.075	0.142
-	*Break Down*	-	*4.34%*	*3.78%*	*13.51%*	*11.22%*	*0.92%*	*12.53%*	*18.99%*	*9.07%*	*15.45%*	*4.31%*	*5.87%*	*100%*
14	**Haryana**	5,046	0.045	0.085	0.184	0.279	0.019	0.154	0.166	0.116	0.185	0.121	0.229	0.152
-	*Break Down*	-	*1.66%*	*3.12%*	*13.44%*	*10.21%*	*0.70%*	*11.23%*	*12.10%*	*8.47%*	*13.52%*	*8.81%*	*16.74%*	*100%*

Appendix L. Continued

M0 Rank	States	Popul.	House	Electri.	Health	Sanit.	Water	Fuel	Asset	Educat.	Liveli.	Child Status	Empow.	M_0
15	**Gujarat**	9,966	0.019	0.102	0.236	0.32	0.094	0.16	0.072	0.159	0.138	0.191	0.206	0.159
-	*Break Down*	-	*0.65%*	*3.56%*	*16.54%*	*11.19%*	*3.28%*	*11.23%*	*5.04%*	*11.11%*	*9.63%*	*13.37%*	*14.39%*	*100%*
16	**Karnataka**	11,555	0.047	0.09	0.227	0.338	0.076	0.184	0.197	0.177	0.136	0.112	0.24	0.172
-	*Break Down*	-	*1.52%*	*2.92%*	*14.69%*	*10.92%*	*2.45%*	*11.87%*	*12.71%*	*11.42%*	*8.77%*	*7.22%*	*15.52%*	*100%*
17	**Maharashtra**	16,425	0.029	0.175	0.245	0.332	0.062	0.173	0.224	0.138	0.166	0.105	0.203	0.173
-	*Break Down*	-	*0.92%*	*5.63%*	*15.80%*	*10.70%*	*1.99%*	*11.16%*	*14.41%*	*8.90%*	*10.65%*	*6.79%*	*13.06%*	*100%*
18	**Andhra Pradesh**	16,357	0.107	0.089	0.251	0.37	0.038	0.167	0.238	0.232	0.2	0.101	0.24	0.192
-	*Break Down*	-	*3.09%*	*2.57%*	*14.50%*	*10.69%*	*1.11%*	*9.64%*	*13.78%*	*13.42%*	*11.54%*	*5.82%*	*13.85%*	*100%*
19	**Arunachal Pradesh**	262	0.321	0.206	0.114	0.301	0.126	0.321	0.227	0.21	0.059	0.206	0.212	0.203
-	*Break Down*	-	*8.81%*	*5.64%*	*6.24%*	*8.24%*	*3.44%*	*17.60%*	*12.45%*	*11.51%*	*3.21%*	*11.27%*	*11.60%*	*100%*
20	**Assam**	7,091	0.186	0.376	0.242	0.327	0.168	0.089	0.183	0.229	0.245	0.137	0.191	0.205
-	*Break Down*	-	*5.05%*	*10.20%*	*13.14%*	*8.86%*	*4.56%*	*4.83%*	*9.94%*	*12.40%*	*13.28%*	*7.43%*	*10.33%*	*100%*
21	**West Bengal**	20,564	0.161	0.4	0.326	0.374	0.056	0.11	0.363	0.285	0.208	0.12	0.305	0.246
-	*Break Down*	-	*3.63%*	*9.05%*	*14.72%*	*8.46%*	*1.27%*	*4.97%*	*16.40%*	*12.89%*	*9.42%*	*5.41%*	*13.80%*	*100%*
22	**Bihar**	25,577	0.273	0.465	0.358	0.488	0.032	0.249	0.038	0.297	0.173	0.174	0.371	0.254
-	*Break Down*	-	*5.97%*	*10.15%*	*15.63%*	*10.66%*	*0.69%*	*10.88%*	*1.65%*	*13.00%*	*7.54%*	*7.62%*	*16.21%*	*100%*
23	**Chhattisgarh**	6,304	0.031	0.244	0.355	0.535	0.17	0.288	0.453	0.252	0.194	0.12	0.374	0.281
-	*Break Down*	-	*0.61%*	*4.82%*	*14.06%*	*10.59%*	*3.37%*	*11.39%*	*17.92%*	*9.99%*	*7.67%*	*4.77%*	*14.81%*	*100%*
24	**Rajasthan**	15,679	0.26	0.359	0.303	0.528	0.173	0.253	0.34	0.27	0.134	0.237	0.38	0.286
-	*Break Down*	-	*5.05%*	*6.96%*	*11.75%*	*10.25%*	*3.36%*	*9.83%*	*13.18%*	*10.47%*	*5.19%*	*9.21%*	*14.75%*	*100%*
25	**Orissa**	10,186	0.27	0.418	0.344	0.528	0.147	0.286	0.197	0.282	0.213	0.14	0.446	0.288
-	*Break Down*	-	*5.21%*	*8.06%*	*13.30%*	*10.20%*	*2.83%*	*11.04%*	*7.61%*	*10.90%*	*8.23%*	*5.40%*	*17.21%*	*100%*
26	**Uttar Pradesh**	43,014	0.16	0.506	0.338	0.581	0.063	0.378	0.376	0.213	0.365	0.208	0.454	0.332
-	*Break Down*	-	*2.67%*	*8.47%*	*11.31%*	*9.71%*	*1.05%*	*12.67%*	*12.60%*	*7.14%*	*12.23%*	*6.97%*	*15.19%*	*100%*
27	**Madhya Pradesh**	16,772	0.235	0.299	0.375	0.627	0.233	0.419	0.439	0.284	0.239	0.189	0.454	0.344
-	*Break Down*	-	*3.80%*	*4.83%*	*12.12%*	*10.13%*	*3.76%*	*13.53%*	*14.18%*	*9.17%*	*7.72%*	*6.11%*	*14.65%*	*100%*
28	**Jharkhand**	7,151	0.064	0.697	0.488	0.813	0.472	0.544	0.696	0.375	0.492	0.259	0.52	0.489
-	*Break Down*	-	*0.72%*	*7.92%*	*11.10%*	*9.24%*	*5.36%*	*12.36%*	*15.83%*	*8.53%*	*11.19%*	*5.90%*	*11.84%*	*100%*
-	**India**	**240,780**	**0.145**	**0.311**	**0.289**	**0.439**	**0.096**	**0.24**	**0.263**	**0.22**	**0.218**	**0.152**	**0.319**	**0.244**
-	***Break Down***	**-**	***3.31%***	***7.08%***	***13.15%***	***10.00%***	***2.20%***	***10.95%***	***11.96%***	***10.01%***	***9.93%***	***6.90%***	***14.52%***	***100%***

Chapter VII

Conclusion

This dissertation consists of three projects focusing on different aspects of aggregation while constructing indices based on multiple attributes of well-being. In particular, this dissertation is concerned with three different types of indices: multidimensional welfare indices, composite indices, and multidimensional poverty indices. An index that summarizes the state of a society by aggregating achievements of individuals or households based on multiple attributes is called a multidimensional index. Whereas an index that summarizes the state of a society by aggregating indicators of multiple attributes is called a composite index.

The first project develops a class of multidimensional social welfare indices that is sensitive to two different forms of inequality across the population. One form of inequality is concerned with the dispersion of the distribution of attributes and the other is concerned with the correlation or association across attributes. The sensitivity to the first form of inequality requires that the welfare index decreases if the dispersion of the attributes across the population increases, while the dimensional averages remain unchanged. The sensitivity to the second form of inequality requires that if the distribution of each attribute remains unaltered but there is an increase in the association across attributes, then the social welfare index decreases (resp. increases) if the attributes are substitutes (resp. complements). In summary, the level of welfare decreases as a consequence of increase in any form of multidimensional inequality. In Chapter II of the dissertation, the class of multidimensional social welfare indices is characterized with the help of certain reasonable axioms. An application of this class to the Indian context is illustrated in Chapter V.

The second project develops a tool that is useful in verifying the robustness of rankings generated by composite indices. These rankings are highly contingent upon the choice of initial weights. A choice other than the one initially selected often alters the rankings and creates ambiguous comparisons, raising an obvious concern as to how robust each comparison is to this choice. To find an answer to this question, a natural measure of robustness is developed in Chapter III, which can gauge the level of robustness on a 0-100 percent scale. A comparison of hundred percent robustness implies that the comparison is never reversed no matter what alternative weights are chosen and thus the comparison is completely robust. However, requiring complete robustness is too stringent and, so, we introduce a concept of partial robustness, where the required level of robustness may strictly lie between zero and hundred. This idea is closely related to the model of Knightian uncertainty (Bewley, 2002) and the multiple prior model of Gilboa and Schmeidler (1989). It is discussed how the required level of robustness depends to the confidence one places on the initial weights. In the same chapter, we also introduce the concept of prevalence of robustness, and in chapter IV, we show how the association across the indicators of the attributes is important in explaining the relationship between the prevalence and the required level of robustness.

Finally, the third project of this dissertation is concerned with measuring poverty in the Indian context, when there are multiple attributes of well-being and the attributes are either categorical or dichotomous. The Indian government has realized the need for a multidimensional approach in addition to the existing income based approach for measuring poverty. In 2002, the government has proposed identifying the poor using a questionnaire containing thirteen non-income related questions. However, this approach has encountered several criticisms in terms of its methodology and the contents of the questionnaire. The third project proposes a new approach developed by Alkire and Foster (2008) for both identifying the poor and measuring poverty, which may amend some of the methodological shortcomings. Moreover, an illustrative index is proposed using some alternative attributes, each of which represent policy goals in the eleventh five-year national plan (Government of India, 2008).

This dissertation contributes to the literature of welfare and poverty measurement both theoretically and empirically. The first project not only proposes a class of welfare indices, but also discusses its theoretical properties. Moreover, this project shows how various classes of indices proposed at different points of time and for different purposes are closely related to each other. Bourguignon (1999), Foster et al. (2005), and Decancq and Ooghe (2009) all have proposed different classes of welfare indices based on generalized means. Bourguignon (1999) proposes his class while constructing a multidimensional inequality index; Foster et al. (2005) propose their class while proposing an inequality-adjusted human development index; and Decancq and Ooghe (2009) propose their class while explaining how correlation between attributes may affect the level of social welfare. All three classes are closely related to the class introduced in the second chapter of this dissertation. In fact, the classes proposed by Foster et al. (2005) and Decancq and Ooghe (2009) are subclasses of our class, and each index in the Bourguignon (1999) class is a monotonic transformation of an index in our class. Chapter II not only establishes clear links across these indices, but also critically evaluates their theoretical properties.

Besides discussing the theoretical features of this class of indices, Chapter V analyzes the empirical applicability of this class by applying it to the Indian context and developing the confidence intervals for testing the statistical significance of the indices. It is shown how the consideration of inequality across the individuals may yield completely different state rankings. Recently, the United Nations Development Programme has taken the initiative of proposing an improved human development index that is sensitive to the inter-personal inequality. The statistical tools developed in Chapter V may be useful for this purpose, because they are also applicable to the indices in the Foster et al. (2005) class.

The second project also contributes to the literature both theoretically and empirically. Chapter III characterizes a measure of robustness linking it to several theoretical concepts in economics, such as, partial ordering, Knightian uncertainty, epsilon contamination, and multiple priors. The measure is also useful for policy analysis. It has been discussed how popular and important these composite indices based rankings are. This natural measure adds another dimension to the cross-country comparisons.

The focus of the third project is however primarily empirical. It applies a theoretically improved measure to the Indian context to analyze the current state of its multidimensional poverty. It appears from the above discussion that the first two projects of this dissertation heavily applies the concept of association to the measurement issues.

Now, let us discuss the possible avenues for future studies and some extensions of the current research. First, using the class of welfare indices in the first project, it is depicted in Chapter V how the state rankings are altered while considering the two forms of inter-personal inequality. However, it is not clear from the analysis if one form of inequality is more prevalent than the other form. The present study does not allow us to isolate the separate effects of these two distinct forms of inequality. The question, differently phrased, is – can we state which form of inequality has greater impact on the level of welfare for a particular society? A technique may be developed that would enable us to answer this question. If we think of an achievement matrix as a joint statistical distribution, then the impact of any one form of inequality should be inferred by eliminating the other form of inequality from that distribution altogether. One possible way to do this is to eliminate the distribution sensitive inequality from the joint distribution, which can be obtained by smoothing all marginal distributions so that every person has the same set of achievements. The problem with this approach is that there would not be any association sensitive inequality either because the attributes would be perfectly positively correlated to each other. The other possibility is to keep the distribution sensitive inequality intact and eliminate the association sensitive inequality. This is equivalent to purging the correlation or association from the joint distribution, while keeping the marginal distributions unaltered. This approach requires constructing a new joint distribution with the same marginal distributions that are independent of each other. The only way to construct such an independent joint distribution from the existing one is using the copula. The copula is a statistical technique for generating numerous joint distributions from a fixed set of marginal distributions.

Second, the Chapter VI is concerned with the measurement of poverty in the multidimensional framework because the measurement of poverty in a single dimension is increasingly recognized as being inadequate. Similarly, the measurement of poverty over a single period of time also seems to be inadequate as it provides only a narrow portrayal of what poverty truly is. Over the past few decades, the measurement of poverty has evolved in two directions. The first is by incorporating more than one dimension of well-being, which is known as the multidimensional poverty measurement — what we discuss in this dissertation. A second direction in which poverty measurement has evolved is by considering a single dimension of well-being over more than one period of time, which is called chronic poverty measurement. However, the progress in these two areas of measurement has been made in separate and disconnected ways, without properly acknowledging the common challenges faced by both approaches. Further research is required in order to establish the connection between these two branches of poverty measurement, addressing the commonalities and differences.

Third, the debate over robustness versus redundancy has not been completely resolved in this dissertation. It is understood that if the association across dimensions is higher, then the rankings are more robust. However, so is the extent of redundancy of composite indices. Thus, further research is required in this area to analyze this link more thoroughly and to investigate if there are other important factors in addition to multidimensional association.

Finally, the empirical application in Chapter V is developed only for an illustrative purpose and further research is required to be conducted. First, the National Family Health Survey data set does not contain any information on income. As a result, we can not directly compare our multidimensional results directly with the results using an income based approach. Secondly, we use a proxy indicator of health risks to measure the quality of health of each household. Further research is required to construct an indicator that can measure household health directly.

Bibliography

Alesina, A. and R. Perotti (1996). Income distribution, political instability, and investment. *European Economic Review 40*, 1203–1228.

Alkire, S. and J. E. Foster (2008). Counting and multidimensional poverty measurement. Working Paper 7, Oxford Poverty & Human Development Initiative, University of Oxford.

Alkire, S. and M. B. Sarwar (2009). Multidimensional measures of poverty & well-being. Report Working Paper, Oxford Poverty & Human Development Initiative, University of Oxford.

Anand, S. and M. Ravallion (1993). Human development in poor countries: On the role of private incomes and public services. *Journal of Economic Perspectives 7*, 133–150.

Asano, T. (2008). Portfolio inertia and epsilon-contaminations. *Theory and Decision 68*, 341–365.

Atkinson, A., B. Cantillon, E. Marlier, and B. Nolan (2002). *Social Indicators: The EU and Social Inclusion*. Oxford: Oxford University Press.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory 2*, 244–263.

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

Atkinson, A. B. and J. Micklewright (1983). On the reliability of income data in the family expenditure survey. *Journal of the Royal Statistical Society 146*, 33–61.

Bartram, J., K. Lewis, R. Lenton, and A. Wright (2005). Focusing on improved water and sanitation for health. *Lancet 365*, 810–812.

Basu, K. and J. E. Foster (1998). On measuring literacy. *Economic Journal 108*, 1733–1749.

Beggs, P. J. and F. Siciliano (2001). Spatial relationship between dwelling crowding and selected causes of morbidity in Sydney, Australia, 1994-1997. *Australian Geographer 32*, 377–401.

Ben Porath, E., I. Gilboa, and D. Schmeidler (1997). On the measurement of inequality under uncertainty. *Journal of Economic Theory 75*, 194–204.

Bewley, T. F. (2002). Knightian decision theory. Part I. *Decisions in Economics and Finance 25*, 79–110.

Bickel, P. J. and D. A. Freedman (1984). Asymptotic normality and the bootstrap in stratified sampling. *The Annals of Statistics 12*, 470–482.

Bierens, H. J. (2004). *Introduction to the Mathematical and Statistical Foundations of Econometrics*. Cambridge University Press.

Blackorby, C., D. Donaldson, and J. A. Weymark (1984). Social choice with interpersonal utility comparisons: A diagrammatic introduction. *International Economic Review 25*, 327–356.

Blackorby, C., D. Primont, and R. R. Russell (1978). *Duality, Separability, and Functional Structure: Theory and Economic Applications*. New York: North-Holland.

Boland, P. J. and F. Proschan (1988). Multivariate arrangement increasing functions with applications in probability and statistics. *Journal of Multivariate Analysis 25*, 286–298.

Bourguignon, F. (1999). Comment on 'multidimensioned approaches to welfare analysis' by E. Maasoumi. In J. Silber (Ed.), *Handbook of Income Inequality Measurement*, pp. 477–484. Boston: Kluwer Academic.

Bourguignon, F., A. Bénassy-Quéré, S. Dercon, A. Estache, J. W. Gunning, R. Kanbur, S. Klasen, S. Maxwell, J.-P. Platteau, and A. Spadaro (2008). Millennium development goals at midpoint: Where do we stand and where do we need to go? European Report on Development.

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

Cahill, M. B. (2005). Is the human development index redundant? *Eastern Economic Journal 31*, 1–5.

Carlier, G. and R. A. Dana (2008). Two-persons efficient risk-sharing and equilibria for concave law-invariant utilities. *Economic Theory 36*, 189–223.

Carlier, G., R. A. Dana, and N. Shahidi (2003). Efficient insurance contracts under epsilon-contaminated utilities. *The GENEVA Papers on Risk and Insurance - Theory 28*, 59–71.

Chakravarty, S. R. (2003). A generalized human development index. *Review of Development Economics 7*, 99–114.

Chakravarty, S. R. (2009). *Inequality, Polarization and Poverty*. New York: Springer.

Chateauneuf, A., J. Eichberger, and S. Grant (2006). Choice under uncertainty with the best and worst in mind. *Journal of Economic Theory 137*, 538–567.

Chauhan, C. (2008). Poverty redefined. Hindustan Times, New Delhi. August 19.

Cherchye, L., W. Moesen, N. Rogge, T. Van Puyenbroeck, M. Saisana, A. Saltelli, R. Liska, and S. Tarantola (2008). Creating composite indicators with DEA and robustness analysis: the case of the technology achievement index. *Journal of the Operational Research Society 59*, 239–251.

Cherchye, L., E. Ooghe, and T. V. Puyenbroeck (2008). Robust human development rankings. *Journal of Economic Inequality 6*, 287–321.

Datt, G. and M. Ravallion (2002). Is India's economic growth leaving the poor behind? *Journal of Economic Perspectives 16*, 89–108.

de Barros, R. P., F. H. Ferreira, J. R. M. Vega, and J. S. Chanduvi (2009). *Measuring Inequality of Opportunities in Latin America and the Caribbean* (Conference ed.). Washington D.C.: The World Bank.

Decancq, K. (2009). Copula-based measurement of dependence between dimensions of well-being. Center for Economic Studies, Katholieke Universiteit Leuven.

Decancq, K. and M. A. Lugo (2008). Setting weights in multidimensional indices of well-being. Working Paper Number 18, Oxford Poverty & Human Development Initiative, University of Oxford.

Decancq, K. and M. A. Lugo (2009). Measuring inequality of well-being with a correlation-sensitive multidimensional Gini index. ECINEQ Working Paper 2009-124.

Decancq, K. and E. Ooghe (2009). Has the world moved forward? a robust multidimensional evaluation. Center for Economic Studies, Katholieke Universiteit Leuven.

Drèze, J. and A. Sen (2002). *India, Development and Participation.* New Delhi and New York: Oxford University Press.

Duflo, E., M. Greenstone, and R. Hanna (2008). Cooking stoves, indoor air pollution and respiratory health in rural orissa. *Economic and Political Weekly 43*, 71–78.

Dutta, I., P. K. Pattanaik, and Y. Xu (2003). On measuring deprivation and the standard of living in a multidimensional framework on the basis of aggregate data. *Economica 70*, 197–221.

Eichhorn, W. (1978). *Functional Equations in Economics.* Reading, MA: Addison-Wesley.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics 75*, 643–669.

Eremin, E. V. (1999). A central limit theorem for linear estimators of measurement results. *Measurement Techniques 42*, 1009–1015.

Esty, D., C. Kim, T. Srebotnjak, M. Levy, A. de Sherbinin, and V. Mara (2008). 2008 environmental performance index. Yale Center for Environmental Law and Policy, Yale University.

Esty, D., M. Levy, T. Srebotnjak, and A. de Sherbinin (2005). Environmental sustainability index: Benchmarking national environmental stewardship. Yale Center for Environmental Law and Policy, Yale University.

Fajnzylber, P., D. Lederman, and N. Loayza (2002). Inequality and violent crime. *Journal of Law and Economics 45*, 1–40.

Filmer, D. and K. Scott (2008). Assessing asset indices. Policy Research Working Paper 4605, The World Bank.

Foster, J. E., L. F. López-Calva, and M. Székely (2005). Measuring the distribution of human development: Methodology and an application to Mexico. *Journal of Human Development 6*, 5–29.

Foster, J. E. and A. Sen (1997). *On Economic Inequality.* Oxford: Oxford University Press.

Foster, J. E. and A. F. Shorrocks (1988a). Poverty orderings. *Econometrica 56*, 173–177.

Foster, J. E. and A. F. Shorrocks (1988b). Poverty orderings and welfare dominance. *Social Choice and Welfare 5*, 179–198.

Foster, J. E. and M. Székely (2008). Is economic growth good for the poor? Tracking low incomes using general means. *International Economic Review 49*, 1143–72.

Gajdos, T. and J. A. Weymark (2005). Multidimensional generalized Gini indices. *Economic Theory 26*, 471–496.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*, 141–153.

Gisbert, F. J. G. (2003). Weighted samples, kernel density estimators and convergence. *Empirical Economics 28*, 335–351.

Government of India (1979). Report of the Task Force on Projections of Minimum Needs and Effective Consumption Demand. Perspective Planning Division, Planning Commission, New Delhi, January.

Government of India (2002). Report of the Expert Group on Identification of Households Below Poverty Line (BPL Census 2002). Submitted to the Ministry of Rural Development.

Government of India (2006). Report of the XI plan Working Group on Poverty Elimination Programmes. Planning Commission. New Delhi, December.

Government of India (2007). Poverty estimates for 2004-05. Press Information Bureau. New Delhi, March.

Government of India (2008). *Eleventh Five Year Plan (2007-2012): Inclusive Growth.* New Delhi: Oxford University Press.

Grosh, M. and P. Glewwe (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study.* Washington D.C.: The World Bank.

Hardy, G. H., J. E. Littlewood, and G. Pólya (1934). *Inequalities*. Cambridge University Press.

Heritage Foundation (2008). *Index of Economic Freedom*. Washington, D.C: The Heritage Foundation.

Hicks, D. A. (1997). The inequality adjusted human development index: A constructive proposal. *World Development 25*, 1283–1298.

Hirway, I. (2003). Identification of BPL households for poverty alleviation programmes. *Economic and Political Weekly 38*, 4803–38.

Howe, L. D., J. R. Hargreaves, and S. R. Huttly (2008). Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. *Emerging Themes in Epidemiology 5:3*. DOI:10.1186/1742-7622-5-3.

International Institute for Population Sciences (IIPS) and ORC Macro (2007). *National Family Health Survey-3, 2005-06, India: Volume-I*. IIPS Mumbai.

Jain, S. K. (2004). Identification of the poor: Flaws in government surveys. *Economic and Political Weekly 39*, 4981–84.

Jalan, J. and R. Murgai (2007). An effective "targeting shortcut"? an assessment of the 2002 below-poverty line census method. World Bank, New Delhi.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.

Jose, S. and K. Navaneetham (2008). A factsheet on women's malnutrition in India. *Economic and Political Weekly 43*, 61–67.

Justino, P. (2004). Redistribution, inequality and political conflict. PRUS Working Paper No. 18, University of Sussex.

Kendall, M. and G. J. Dickinson (1990). *Rank Correlation Methods* (Fifth ed.). New York: Oxford University Press.

Khera, R. (2008). Access to the targeted public distribution system in Rajathan. *Economic and Political Weekly 43*, 51–56.

Kockläuner, G. (2006). A multidimensional poverty index. In K.-H. Waldmann and U. M. Stocker (Eds.), *Operations Research Proceedings 2006*, pp. 223–225. Springer.

Kockläuner, G. (2008). On generalized means of censored income gap ratios. Unpublished manuscript, Fachbereiche Wirtschaft, Fachhochschule Kiel.

Kolm, S.-C. (1977). Multidimensional equalitarianisms. *Quarterly Journal of Economics 91*, 1–13.

Kopylov, I. (2009). Choice deferral and ambiguity aversion. *Theoretical Economics 4*, 199–225.

Laderchi, C. R. (2008). Do concepts matter? an empirical investigation of the differences between a capability and a monetary assessment of poverty. In F. Comim, M. Qizilbash, and S. Alkire (Eds.), *The Capability Approach: Concepts, Measures and Applications*, pp. 203–241. Cambridge: Cambridge University Press.

Laderchi, C. R., R. Saith, and F. Stewart (2003). Does it matter that we do not agree on the definition of poverty? a comparison of four approaches. *Oxford Development Studies 31*, 243–274.

Marshall, A. W. and I. Olkin (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

Mayer, S. E. and C. Jencks (1989). Poverty and the distribution of material hardship. *Journal of Human Resources 24*, 88–114.

McGillivray, M. (1991). The Human Development Index: Yet another redundant composite development indicator? *World Development 19*, 1461–1468.

McGillivray, M. (2005). Measuring non-economic well-being achievement. *Review of Income and Wealth 51*, 337–364.

McGillivray, M. and H. White (1993). Measuring development? the UNDP's Human Development Index. *Journal of International Development 5*, 183–192.

McKenzie, D. J. (2005). Measuring inequality with asset indicators. *Journal of Population Economics 18*, 229–260.

Milgrom, P. and J. Roberts (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica 58*, 1255–1277.

Mishra, V. K., R. D. Retherford, and K. R. Smith (1999). Biomass cooking fuels and prevalence of blindness in India. *Journal of Environmental Medicine 1*, 189–199.

Morris, M. D. (1979). *Measuring the Condition of the World's Poor: The Physical Quality of Life Index.* New York: Pergamon Press.

Morris, S. S., C. Carletto, J. Hoddinott, and L. J. M. Christiaensen (2000). Validity of rapid estimates of household wealth and income for health surveys in rural Africa. *Journal of Epidemiology and Community Health 54*, 381–387.

Mukherjee, N. (2005). Political corruption in India's below the poverty line (BPL) exercise: Grassroots' perspectives on BPL in perpetuating poverty and social exclusion. Development Tracks in Research, Training & Consultancy, New Delhi, India.

Nishimura, K. G. and H. Ozaki (2006). An axiomatic approach to ε-contamination. *Economic Theory 27*, 333–340.

Pattanaik, P. K., S. G. Reddy, and Y. Xu (2007). On procedures for measuring deprivation and living standards of societies in a multi-attribute framework. Working Paper 08-02, Andrew Young School of Policy Studies, Georgia State University.

Permanyer, I. (2007). *On the Robust Measurement of Well-Being in a Gender Perspective.* Ph. D. thesis, Universitat Autònoma de Bercelona.

Radhakrishna, R. and S. Ray (Eds.) (2005). *Handbook of Poverty in India, Perspective, Policies, and Programmes.* New Delhi: Oxford University Press.

Reddy, S. G. (2008). Differently distorted: The World Bank's 'updated' poverty estimates. *Economic and Political Weekly 43*, 44–47.

Rehfuess, E., L. Tzala, N. Best, D. Briggs, and M. Joffe (2009). Solid fuel use and cooking practices as a major risk factor for alri mortality among african children. *Journal of Epidemiology and Community Health 63*, 887–892.

Roberts, F. S. (1979). In *Measurement Theory: with Applications to Decision Making, Utility, and The Social Sciences.* Reading, MA: Addison-Wesley.

Saisana, M., A. Saltelli, and S. Tarantola (2005). Uncertainty and sensitivity analysis as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society 168*, 1–17.

Sen, A. (1970). *Collective Choice and Social Welfare.* San Francisco: Holden-Day.

Sen, A. (1976). Poverty: An ordinal approach to measurement. *Econometrica 44*, 219–231.

Sen, A. K. (1985). *Commodities and Capabilities.* Amsterdam: North-Holland.

Sen, A. K. (1996). On the foundations of welfare economics: Utility, capability and practical reason. In F. Farina, F. Hahn, and S. Vannucci (Eds.), *Ethics, Rationality and Economic Behaviour*, pp. 50–65. Oxford: Clarendon Press.

Sen, A. K. (2004). Capabilities, lists and public reason: Continuing the conversation. *Feminist Economics 10*, 77–80.

Seth, S. (2009). Inequality, interactions, and human development. *Journal of Human Development and Capabilities 10*, 375–396.

Stewart, F. (2008). *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies.* London: Palgrave.

Stifel, D. and L. Christiaensen (2007). Tracking poverty over time in the absence of comparable consumption data. *World Bank Economic Review 21*, 317–341.

Stiglitz, J. E., A. Sen, and J.-P. Fitoussi (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. http://www.stiglitz-sen-fitoussi.fr/en/index.htm.

Streeten, P., S. J. Burki, M. U. Haq, N. Hicks, and F. Stewart (1981). *First Things First: Meeting Basic Human Needs in the Developing Countries.* New York: Oxford University Press.

Sundaram, K. (2003). On identification of households below poverty line in BPL census 2002: Some comments on proposed methodology. *Economic and Political Weekly 38*, 4803–08.

Topkis, D. (1998). *Supermodularity and Complementarity.* Princeton University Press, Princeton.

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

Tsui, K.-Y. (1999). Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation. *Social Choice and Welfare 16*, 145–157.

Tsui, K.-Y. (2002). Multidimensional poverty indices. *Social Choice and Welfare 19*, 69–93.

United Nations (2003). *Indicators for Monitoring the Millennium Development Goals: Definitions, Rationale, Concepts and Sources.* New York.

United Nations Development Programme (2000). *Human Development Report.* New York: Oxford University Press.

United Nations Development Programme (2006). *Human Development Report.* New York: Oxford University Press.

Watts, H. (1969). An economic definition of poverty. In D. Moynihan (Ed.), *On Understanding Poverty*, pp. 316–329. New York: Basic Books.

Weymark, J. A. (2006). The normative approach to the measurement of multidimensional inequality. In F. Farina and E. Savaglio (Eds.), *Inequality and Economic Integration*, pp. 303–328. London: Routledge.

World Health Organization (2000). The World Health Report, Health Systems: Improving Performance. Geneva, Switzerland.

World Health Organization and UNICEF (2006). *Core Questions on Drinking-Water and Sanitation for Household Surveys.* Geneva, Switzerland: WHO Press.