UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2017/206

Robustness tests for multidimensional poverty comparisons

Sam Jones*

December 2017

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: This paper provides practical tests for the robustness of multidimensional comparisons of well-being. Focussing on counting-type multidimensional poverty measures, I draw on the properties of positive Boolean threshold functions to prove that the space of feasible poverty definitions is finite and can be partitioned into at most $(D^2+D)/2$ parts, where D is the number of dimensions of well-being spanned by the measure. This provides the foundation for two complementary tests: (i) a bounding approach, which weights each dimension equally; and (ii) stochastic search, where coverage of the space of poverty definitions is assessed via Good-Turing estimates of missing mass. The two methods are applied to a measure encompassing nine dimensions of well-being in Mozambique, revealing persistent regional asymmetries.

Keywords: multidimensional poverty, pointwise dominance, robustness, Mozambique
JEL classification: I32, O55, C15

Acknowledgements: Thanks to Paul Anand, James Foster, Simon Quinn, Vincenzo Salvucci, Suman Seth, Finn Tarp, Erik Thorbecke, and Gaston Yalonetzky, as well as seminar participants at Queen Elizabeth House, University of Oxford and UNU-WIDER, Helsinki for comments and encouragement on earlier versions of this paper. All errors and omissions are my own.

* University of Copenhagen, Denmark; email sam.jones@econ.ku.dk .

This study has been prepared within the project on 'Inclusive growth in Mozambique—scaling-up research and capacity' implemented in collaboration between UNU-WIDER, University of Copenhagen, University Eduardo Mondlane, and the Mozambican Ministry of Economics and Finance. The project is financed through specific programme contributions by the governments of Denmark, Finland, Norway, and Switzerland.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-432-2 https://doi.org/10.35188/UNU-WIDER/2017/432-2

Typescript prepared by Ans Vehmaanperä.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Denmark, Finland, Sweden, and the United Kingdom.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Interest in multidimensional poverty has moved toward the policy mainstream.[1] Since 2010, the UN's Human Development Report has included a multidimensional poverty index based on survey data for more than 100 countries and covering ten dimensions of well-being. In 2011, Colombia officially adopted a multidimensional poverty index to guide and track progress in its national poverty reduction plans (Angulo et al., 2016). Despite the growing use of these metrics, an unresolved challenge is how to adequately account for potential differences in the relative importance of the various dimensions of well-being that make up a multidimensional poverty measure. Colombia's index captures social and health-related aspects of poverty across five broad areas, using indicators spanning 15 different dimensions.[2] While the five areas receive an equal weight, the individual dimensions are treated differently – e.g., the percentage in long-term unemployment attracts a weight of 10%, while access to a clean public water system attracts a weight of 4%. The concern is that comparisons of progress based on one definition of multidimensional poverty (weighting structure) may not hold up if a different definition (set of weights) is used (Ravallion, 2012).

Various studies have sought to assess the robustness of multidimensional welfare comparisons, focussing on the stability of differences in well-being between groups or over time to changes in the measure of welfare employed. Following progress in the unidimensional literature on consumption and income poverty, a prevailing approach employs metrics of distributional dominance (Duclos et al., 2006). The idea here is to compare the properties of the multivariate empirical distribution of the underlying data on well-being (attainments). In so doing, dominance results obtained from such approaches should apply to a wide range of plausible social welfare functions and, therefore, will not depend on any subjective choices used to construct specific outcome metrics.

By dint of its generality, evaluation of robustness via multivariate stochastic dominance is attractive. Nonetheless, applying measures of distributional dominance to complex joint distributions – which may include variables of a binomial, ordinal and continuous nature – remains work in progress. As Arndt et al. (2012a) note, existing approaches tend to place restrictions on the nature of the underlying social welfare functions, particularly as regards the permitted sign of cross-derivatives between dimensions. Moreover, despite important advances in extending multivariate measures of stochastic dominance to ordinal variables (Yalonetzky, 2013), these approaches do not transfer easily to the popular counting-type measures of multidimensional

[1] Seminal contributions to the literature on constructing multidimensional well-being measures include Maasoumi (1986); Atkinson (2003); Bourguignon and Chakravarty (2003); Alkire and Foster (2011).

[2] For further details see: http://ophi.org.uk/policy/national-policy/colombia-mpi/.

well-being (see also Atkinson, 2003). Indeed, Yalonetzky (2011) concludes that: " traditional dominance conditions ... are not appropriate for poverty counting measures ... except when extreme poverty identification approaches are considered." (p. 18).

A reliance on distributional dominance to assess robustness can be questioned on other grounds. First, when applied to cases that combine data from a large number of dimensions, these techniques are extremely demanding in terms of data and computational time (Crawford, 2005). As a result, existing applications of these methods typically consider only a few dimensions (Maasoumi and Millimet, 2005; McCaig and Yatchew, 2007), which contrasts to the much larger number of dimensions typically included in popular implementations of multidimensional poverty indexes (c.f., Colombia). Second, dominance tests typically generate a single overall statistic or binary conclusion of whether dominance holds. As such, they may not be very informative for policy-makers where the extent of dominance and sources of distributional overlap are often of interest (Bennett and Mitra, 2013). Third, and as noted by Hansen et al. (1978), first order stochastic dominance does not necessarily imply pointwise dominance. Even if the (multivariate) cumulative distribution function of deprivations for group j stochastically first order dominates that for group j', this does not mean that multidimensional poverty is higher in group j for *every* feasible poverty definition. However, from the perspective of investigating whether poverty comparisons are stable (in direction), assessment of pointwise dominance is important and meaningful (see also Foster et al., 2013).[3]

An alternative route to assessing robustness is via sensitivity or uncertainty analysis.[4] These methods do not investigate the distribution of the underlying data directly; rather, they consider the sensitivity of the output from a chosen measurement approach to alternative assumptions. The aim is to characterize the empirical probability density function of the (univariate) outcome distribution and, thereby, ascertain whether inferences drawn from a specific set of modelling assumptions are fragile. This type of analysis holds advantages in complex situations (as here), where closed form analytical tests regarding a model's properties are elusive. It also has the practical benefit of being intuitive and transparent. By providing an estimate of the empirical distribution of outputs (point-by-point), pointwise dominance can be evaluated directly. Furthermore, one can explore specific sources of variation or interesting features of the same outcome distribution.

[3] For simplicity, throughout this paper I focus on comparisons of poverty – i.e., whether multidimensional poverty in group j is greater than in group j'. Where this obtains, it is equivalent to saying that well-being in group j' is higher than in group j.

[4] For a general discussion of sensitivity analysis see Saltelli et al. (2009). The distinction between uncertainty and sensitivity analysis is not strict. Helton et al. (2006) (*inter alia*), however, suggest that uncertainty analysis seeks to quantify the degree of uncertainty in outputs due to uncertainty regarding the appropriate values or form of inputs; sensitivity analysis, on the other hand, seeks to determine the sources of variation in outputs due to differences (uncertainty) in inputs.

A main disadvantage and source of criticism of sensitivity analysis is that it is often perfunctory. That is, a general critique is that these methods fail to consider a sufficient combination of different inputs; or, at least, it is unclear when the outcome distribution has been characterized comprehensively. For instance, many practical sensitivity analyses rely on varying inputs one at a time. However, as Saltelli and Annoni (2010) show, this is only justifiable if the model is linear in these inputs. Taking the case of a counting-type multidimensional poverty measures (discussed below), these concerns are pertinent. Alkire et al. (2010) explicitly ask whether the multidimensional poverty measure adopted in the UN's 2010 Human Development Report is robust to different weights. However, to answer this, they only evaluate the stability of country rankings under three alternative weighting structures. In similar exercises, both Lasso de la Vega (2010) and Bennett and Mitra (2013) limit their analysis of robustness to verifying the impacts of variation in the poverty cut off, which is just one (variable) input into the definition of such poverty measures.[5]

Cognizant of the limitations of existing methods, the aim of this study is to set out a rigorous approach to evaluating the robustness of multidimensional poverty comparisons, focussing on counting-type measures where assessment via stochastic dominance has achieved limited progress. In doing so, I contribute theoretically and practically. As a theoretical contribution, I note the equivalence between the popular multidimensional headcount poverty measures developed in Alkire and Foster (2011) (hereafter, AF) and positive Boolean threshold functions. This is important because examination of the properties of these functions provides viable directions for assessing robustness. I show that the space of unique poverty definitions is countably finite and can be partitioned according to the maximum and minimum lengths of its prime implicants. The latter means that robustness can be analysed by comparing bounds on multidimensional headcounts for each partition; and the former means that a stochastic search of the space of unique poverty definitions can be complemented by measures of the degree of coverage of the search space, for which the Good-Turing estimator of missing mass is appropriate.

As a practical contribution, I show how these two approaches can be implemented. I do so via the example of the evolution of multidimensional poverty in Mozambique, calculated from a series of household surveys spanning around 20 years. The results provide strong evidence that multidimensional poverty has declined over time at the national level, for any possible poverty definition. That is, headcount poverty in 1996/97 strongly pointwise dominates poverty rates today. Even so, I show that headcount poverty rates in the Northern and Central regions in 2014/15 overlap with those in the South in 1996/97 – i.e., no dominance relation holds. This confirms the presence of very persistent regional asymmetries in well-being.

[5] A similar critique applies to the robustness analyses set out in Batana (2013) and Alkire and Santos (2014).

By way of structure, Section 2 sets out a general definition of robustness in the context of poverty comparisons. Section 3 adopts a Boolean perspective. It clarifies the equivalence between the AF headcount and positive Boolean threshold functions; and derives three propositions about these functions. These propositions directly inform the two robustness procedures – bounding and stochastic search – which are described in Section 4. Section 5 moves on to their application. Here, I begin by describing the data and context. Next, I set out the practical steps required to implement the methods. I then present the main results of the robustness analysis, followed by brief comments on potential analytical extensions. As a complement to the application, Appendix A describes a set of user-friendly functions, written in R, that implement these procedures and which are also supplied with this paper. Finally, Section 6 concludes.

2 Defining robust comparisons

To begin it is helpful to specify the precise meaning of the so-called robust comparisons problem. To do so, the following general definitions are used:

Definition 2.1. A multidimensional poverty *definition* is a unit-wise mapping $f : (\theta_t, \vec{z}_{ij}) \rightarrow y_{ij} \in [0, 1]$, where:

- $\vec{z}_{ij}$ is a $D \times 1$ vector of raw observations for unit i belonging to group j and represents the indicators of deprivation to be combined/aggregated across each unit. Throughout, this raw data is taken as given;
- Unless otherwise stated, $Z_j = (\vec{z}_{1j}, \vec{z}_{2j}, ..., \vec{z}_{Ij})$ is taken as a binary matrix; in which case each $\vec{z}_{ij}$ is a draw from the set of deprivation domains $\mathcal{D} = \{z_1, z_2, ..., z_D\}$ and where $z_d \in \{0, 1\}$;
- θ_t is a draw from the set $\mathcal{S}$ of feasible input parameters, which is assumed to be compact and takes an associated probability distribution Ω. Constraints on the parameters, such as adding-up requirements, are implicit in the definition of $\mathcal{S}$; and
- y_i is assumed to follow a stable distribution, implying f is well-behaved.

Definition 2.2. A multidimensional poverty *measure* is a positive monotone aggregation of the vector of unit-wise outcomes to yield a group-wise summary statistic $g : y_{ij} \rightarrow \bar{y}_j \in [0, 1]$. Typically, g is the (weighted) mean operator.

Definition 2.3. Random variable $\bar{y}_j$ is said to *weakly pointwise dominate* random variable $\bar{y}_{j'}$ if, for a given choice of the multidimensional mapping function f and aggregation function g, it holds that $\forall\ \theta_t \in \mathcal{S} : \bar{y}_j \geq \bar{y}_{j'}$ and there is at least one draw such that $\bar{y}_j > \bar{y}_{j'}$.

Definition 2.4. Random variable $\bar{y}_j$ is said to *strongly pointwise dominate* random variable $\bar{y}_{j'}$ if, for a given choice of the multidimensional mapping function f and aggregation function g, $\forall\, \theta_t \in \mathcal{S} : \bar{y}_j > \bar{y}_{j'}$.

Definition 2.5. A comparison of multidimensional poverty between two groups j and j' is said to be *strongly robust* if, $(\bar{y}_j \succ_{sp} \bar{y}_{j'}) \vee (\bar{y}_{j'} \succ_{sp} \bar{y}_j)$, where $\succ_{sp}$ denotes strong pointwise dominance. Weak robustness is defined analogously.

The first two definitions state that a generic measure of multidimensional poverty for a given group involves two operations on the raw data matrix Z_j, composed of D marginal dimensions (columns or deprivation domains) and I units (rows, typically households or individuals). The first is the row-wise operation that combines the outcomes in each deprivation domain into a single composite outcome; the second is the simple aggregation over all members of the group. Definition (2.3) refers to comparisons of poverty measures between two groups and provides an intuitive definition of dominance. It states that multidimensional poverty in group j dominates (is larger than) that of j' if their difference is never less than zero and is positive for at least one vector of inputs θ_t.

The key aspect of this definition of a robust comparison is that it is defined on the parameter space $\mathcal{S}$. That is, pointwise dominance is defined here for a given form of the mapping function f and aggregation function g, essentially meaning we are interested in the difference between $\bar{y}_j$ and $\bar{y}_{j'}$ taken over all feasible values of θ in the parameter space. In turn, Definition (2.5) simply states an equivalence between strong pointwise dominance and strong robustness. As elaborated by Hansen et al. (1978), pointwise dominance (PD) is a strict form of dominance that immediately implies first order stochastic dominance (but not *vice versa*). This notion of robustness also is consistent with existing uses in the literature (Foster et al., 2013; Alkire et al., 2015).

3 A Boolean perspective

The previous section set out a general definition of what constitutes a robust poverty comparison. Evidently, since robustness holds for *given* choices of f and g, as well as raw data matrices $(Z_j, Z_{j'})$, the nature of these functional forms is fundamental. In fact, it is precisely the form of f that tends to differentiate between alternative multidimensional poverty measures employed in previous empirical work (Atkinson, 2003). One of the most established of these is the

Alkire-Foster (AF) class of metrics. The simple AF headcount measure is defined as follows:

$$z_{id} = \begin{cases} 0 \text{ if } x_{id} \geq \bar{x}_d \\ (\bar{x}_d - x_{id})/\bar{x}_d \text{ otherwise} \end{cases} \tag{1a}$$

$$y_i = \mathcal{I}\left(\sum_{d\in\mathcal{D}} \omega_d [z_{id}]^0 > k\right) \tag{1b}$$

$$\sum_{d\in\mathcal{D}} \omega_d = 1, \ \forall d \ : \omega_d > 0, \ 0 < k \leq 1$$

$$H(\vec{\omega}, k) = N^{-1} \sum_{i=1}^{N} y_i \tag{1c}$$

As shown, the headcount is calculated in three steps. The first applies poverty thresholds, denoted $\bar{x}_d$ to each column of the raw data, yielding a matrix of poverty gaps, with elements z_{id}. Second, these are combined row-wise to classify units as either poor ($y_i = 1$) or not ($y_i = 0$). Third, the prevalence of poverty, denoted H, is just the average of vector y_i, which may be calculated using sample weights (not shown). So, in relation to the definitions of Section 2, equation (1b) corresponds to the row-wise mapping function f; and equation (1c) corresponds to the group-wise aggregation g.

The principal insights of this study stem from recognizing that the mapping operation f applied by the AF headcount measure (H) is a positive Boolean threshold function. Formally, these are defined as follows:

Definition 3.1. A positive Boolean threshold function f is a mapping $\{0,1\}^D = \mathcal{B}^D \mapsto \{0,1\}$ with the property that it is positive in each of its inputs and there exist D positive weights, $(w_1, w_2, \ldots, w_D) \in \mathbb{R}^D_+$, and a threshold $0 < k \leq \sum_{d=1}^{D} \omega_d$, such that for all inputs, $(z_{i1}, z_{i2}, \ldots, z_{iD}) \in \mathcal{B}^D$, it holds:

$$f(z_{i1}, z_{i2}, \ldots, z_{iD}) = \begin{cases} 1 \text{ if } \sum_{d=1}^{D} w_d z_{id} \geq k \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

The equivalence between Definition (3.1) and equation (1b) is obvious. What is important is that positive Boolean threshold functions have a number of interesting properties, which in turn provide viable pathways for evaluation of robustness.

3.1 Finiteness

The first property of positive Boolean threshold functions is that they are finite in number. Formally:

Proposition 1. *When f (Definition 2.1) corresponds to a positive Boolean threshold function, then for any given number of deprivation domains, the number of unique poverty definitions is countably finite.*

Proof. This derives trivially from the properties of positive Boolean threshold functions. A standard result in Boolean algebra (see Theorems 1.04, 1.13 and 1.23 in Crama and Hammer, 2011) is that each positive Boolean function can be represented by its unique Blake canonical form (complete disjunctive normal form, DNF), which is given by the disjunction of all its prime implicants. That is, it can be represented by a unique sequence of union (OR) and intersection (AND) operations on the individual inputs (literals), none of which enter in complemented form. In the present context, the literals are individual elements taken from the set of deprivation domains $\mathcal{D}$.

Consider the power set, $\mathcal{P}(\mathcal{D})$, the elements of which represent all unique conjunctions of literals in $\mathcal{D}$. It follows that a poverty definition is a member of the hyper-power set populated by all feasible complete DNFs formed from unique disjunctions of members of $\mathcal{P}(\mathcal{D})$. This constitutes the poverty space, which is a (finite) subset of all Boolean functions in D variables. □

A main implication of the above is that any combination of weights ($\vec{\omega}$) and thresholds (k) that enter as inputs into the poverty definition (c.f., equation 1b), can be restated as a Boolean expression that combines the literals via a sequence of AND or OR operations. So, although there may be an infinite number of choices for the weights and cut-off, the set of all possible Boolean functions to which they correspond is finite. From this, it directly follows that the number of unique points in the parameter space (see Definition 2.1, where $\theta = \{\vec{\omega}, k\}$) is finite. So, since g is a fixed operator, the set of outcome measures also is countably finite.

3.2 Classifiability

What I term the classifiability of positive Boolean threshold functions refers to the fact that one can classify any given function according to restrictions it places on the combination of literals required to map to a positive outcome (a logical result of one). This is evident in the simplest case as follows:

Proposition 2. *When f (Definition 2.1) corresponds to a positive Boolean threshold function containing all literals of interest, a multidimensional poverty measure (Definition 2.2) has greatest lower bound when θ corresponds to the conjunction (intersection) of all literals in $\mathcal{D}$, and has least upper bound when θ corresponds to the disjunction (union) of all literals in $\mathcal{D}$.*

Proof. Consider a draw from the set of deprivation domains $\mathcal{D}$, denoted $\vec{z}_i \in \{0,1\}^D$. The positive Boolean threshold function corresponding to the union of all elements of $\vec{z}_i$ takes a value of one if $\exists : z_i = 1$; in contrast, the Boolean threshold function corresponding to the intersection of all elements of $\vec{z}_i$ takes a value of one if and only if all elements equal one. Without loss of generality or need for proof, the intersection of two sets is a subset of each. So, the cardinality of the intersection of D subsets (here, positive elements) must be less than or equal to the cardinality of their union. Since this proves that, for any draw $\vec{z}_i$, the union of all literals is the least restrictive poverty definition and the intersection of all literals is the most restrictive, the same property carries over to the aggregate measure taken over any vector of outcomes $(y_1, y_2, ..., y_N)$ derived from the application of the *same* threshold function to N draws from $\mathcal{D}$. □

The above is just a formal statement of the established idea that the AF headcount measure is able to combine both intersection and union type definitions of poverty (Alkire and Foster, 2011). More specifically, for any given matrix of binary deprivation measures (Z_i), the lower and upper bounds on the headcount measure can be derived in a straight-forward manner. That is, for all possible choices of weights and cut-offs, the upper bound on the headcount obtains when each unit is classified as poor if it is poor in any one domain; and the lower bound obtains when each unit is classified as poor only if it is poor in all domains simultaneously.

The same logic can be extended to more complex positive Boolean threshold functions, namely those containing a mix of union and intersection operations on the literals. To investigate these, the length of (number of literals contained in) the prime implicants in the DNF of the expression are of interest. Drawing on the weighted voting game literature (Taylor and Zwicker, 1992; Elkind et al., 2009; Zuckerman et al., 2012), it is helpful to focus on the smallest and largest winning coalitions (SWC, LWC) required to switch the Boolean function from zero to one.[6] In the proposition described above, the union of all literals constitutes a case where both the smallest and largest winning coalitions contain just one element; and the intersection of literals is the case where the SWC and LWC both contain D elements. Any positive Boolean threshold function

[6] An SWC of size s implies that a minimum of s votes are required for a positive outcome; a LWC of size l states that at most l votes are required (i.e., it is the largest coalition required in which there are zero redundant players). Applied to the present context, the SWC gives the smallest number of dimensions on which a person must be deprived to be classified as poor.

can be classified according to the size/length of the SWC (denoted s) and LWC (denoted l) it contains. In turn, this means that any poverty definition given by the input vector θ_t (Definition 2.1) can be classified in the same way.

To denote the partition to which a given poverty definition belongs I use the notation $\theta_{s,l}$, which states that θ belongs to class SWC $= s$ and LMC $= l$.[7] Alongside previous notation, the results of Proposition (2) thus can be re-stated as follows:

$$\forall(s,l) : \bar{y}\big(Z, \tilde{\theta}_1\big) \geq \bar{y}\big(Z, \theta_{s,l}\big) \geq \bar{y}\big(Z, \tilde{\theta}_D\big) \tag{3}$$

and where $\tilde{\theta}_d$ is the special case in which $f(\theta, \cdot)$ corresponds to all feasible conjunctions of prime implicants of length d. This corresponds to the unique Boolean threshold function containing a vector of equal weights ($\forall d : \omega_d = 1/D$) and cut-off $k = d/D$ in equation (1b). This leads to the final proposition, which further tightens these bounds:

Proposition 3. *When f (Definition 2.1) corresponds to a positive Boolean threshold function containing all literals of interest, then a multidimensional poverty measure (Definition 2.2) denoted $\bar{y}\big(Z, \theta_{s,l}\big)$, with shortest prime implicant $1 \leq s \leq l$, longest prime implicant $s \leq l \leq D$ and no redundant literals, is bounded as follows:*

$$\forall(s,l) : \bar{y}\big(Z, \tilde{\theta}_s\big) \geq \bar{y}\big(Z, \theta_{s,l}\big) \geq \bar{y}\big(Z, \tilde{\theta}_{l^*}\big), \tag{4}$$

$$\textit{where } l^* = \begin{cases} l \textit{ if } s < l \wedge 1 < l < D \\ \min(l+1, D) \textit{ otherwise} \end{cases}$$

Proof. Again, this is a straightforward application of set algebra. By definition, each prime implicant of $\tilde{\theta}_{l^*}$ is a subset of at least one prime implicant of $\theta_{s,l}$. Equally, every prime implicant of $\theta_{s,l}$ is a subset of one of the prime implicants of $\tilde{\theta}_s$. So, by the rules of the cardinality of sets, the expression holds. □

The above propositions can be easily understood via a simple example. Consider the set of three deprivation domains, $\mathcal{D} = \{z_1, z_2, z_3\}$ and a chosen poverty definition represented in complete DNF as $f(\theta_{1,2}) := \{z_1\} \cup \{z_2 \cap z_3\}$. The smallest prime implicant of $f(\theta_{1,2})$ has one element and the longest prime implicant two elements. It follows that the cardinality of a more restrictive definition where all prime implicants are of size 2, namely $f(\theta_{2,2}) := \{z_1 \cap z_2\} \cup \{z_1 \cap z_3\} \cup \{z_2 \cap z_3\}$, must be less than or equal to that of $\theta_{1,2}$ for any draw of $\mathcal{D}$. Similarly, the cardinality of the least restrictive definition $f(\theta_{1,1}) := \{z_1\} \cup \{z_2\} \cup \{z_3\}$ must be greater than or equal to that of $f(\theta_{1,2})$ for any draw of $\mathcal{D}$.

[7] This is for notational purposes only and is not to say that all positive Boolean threshold functions can be uniquely indexed by s and l.

4 Evaluating robustness

I now reflect on how these insights can be used to evaluate robustness in practice. Specifically, I set out a bounding procedure that identifies the partitions of the poverty space in which robustness obtains. I then set out a more general approach, based on a stochastic search of the poverty space complemented by a measure of search coverage.

4.1 Bounding approach

The proposed bounding approach to assessing robustness draws directly on Propositions (2) and (3). As noted, any given multidimensional poverty measure of the AF headcount type can be classified by the size/length of the SWC and LWC implied by the choice of weights and cut-off. Note that while Proposition (3) requires that the same poverty measure (when expressed in complete disjunctive normal form) contains all literals, this simply amounts to a requirement that no single deprivation domain is redundant – i.e., each literal has non-zero power such that, at least in combination with some other literals, its presence can switch the Boolean function from zero to one. In effect, this requirement restricts us to poverty definitions that are informative and meaningful across all domains, which is pertinent in most practical applications.

The important point concerning the classifiability of positive threshold Boolean functions is that the poverty space can be partitioned, whereby each partition is characterized by a unique combination or class of the SWC and LWC sizes underlying different poverty definitions. Furthermore, for each partition, equation (4) can be used to establish the upper and lower bounds on the headcount corresponding to all poverty definitions within the same class. Returning to the previous example with just three domains, all (meaningful) poverty definitions can be classified into one of four feasible zones: $\theta_{1,1}, \theta_{1,2}, \theta_{2,2}, \theta_{3,3}$.[8] In fact, in this simple case we can enumerate all positive Boolean threshold functions containing no redundant literals. These are set out in Table 1, where the first column gives the relevant partition class (SWC=s, LWC=l) and the second column is the canonical sum of products form for the logical combination of literals.[9] The final two columns indicate the upper and lower bounds on the given headcount, as per equation (4).

Two main implications fall out of this framework. First, and most obviously, the D Boolean

[8] If the LWC is equal to the number of domains, no other prime implicants with fewer than D literals can be present unless one of the literals is redundant.

[9] As before, $z_i \in \{0, 1\}$ and that the outcome of the expression takes a value of one if and only if the sum of products is non-zero.

Table 1: Positive Boolean threshold functions in exactly three variables

Class	Boolean expression	Bounds	
(s, l)	Sum of products	Upper	Lower
1, 1	$z_1 + z_2 + z_3$	$\tilde{\theta}_1$	$\tilde{\theta}_1$
1, 2	$z_1 + z_2.z_3$	$\tilde{\theta}_1$	$\tilde{\theta}_2$
1, 2	$z_2 + z_1.z_3$	$\tilde{\theta}_1$	$\tilde{\theta}_2$
1, 2	$z_3 + z_1.z_2$	$\tilde{\theta}_1$	$\tilde{\theta}_2$
2, 2	$z_1.z_2 + z_1.z_3 + z_2.z_3$	$\tilde{\theta}_2$	$\tilde{\theta}_2$
2, 2	$z_1.z_2 + z_1.z_3$	$\tilde{\theta}_2$	$\tilde{\theta}_3$
2, 2	$z_1.z_2 + z_2.z_3$	$\tilde{\theta}_2$	$\tilde{\theta}_3$
2, 2	$z_1.z_3 + z_2.z_3$	$\tilde{\theta}_2$	$\tilde{\theta}_3$
3, 3	$z_1.z_2.z_3$	$\tilde{\theta}_3$	$\tilde{\theta}_3$

Notes: 'Class' indicates the partition, defined by the smallest and longest prime implicants (SWC = s; LWC = l), to which the Boolean expression belongs, where the latter is given in the sum of products form; 'Bounds' denote the corresponding theoretical upper and lower bounds on the headcounts (see equation 4).
Source: own elaboration.

threshold functions formed from combining a fixed vector of equal weights with the series of D natural cut-offs: $k \in \{1/D, 2/D, ..., 1\}$ provide a sufficient basis to calculate bounds for any class $\theta_{s,l}$. Second, the same bounds can be directly applied to undertake poverty comparisons between groups. To see this, note that for any two headcount distributions with lower and upper bounds given by $H_j \in [x_j, y_j]$ and $H_{j'} \in [x_{j'}, y_{j'}]$, their difference is contained in: $(H_j - H_{j'}) \in [x_j - y_{j'}, y_j - x_{j'}]$. So, if the lower bound of the first distribution is greater than the upper bound of the second distribution, the pairwise difference is always in the positive domain. It follows that a maximal lower bound on the difference between AF headcount measures calculated over two groups, for any choice of weights and cut-offs, is given by: $\bar{y}(Z_j, \tilde{\theta}_D) - \bar{y}(Z_{j'}, \tilde{\theta}_1)$.

Admittedly, this maximal bound is likely to be wide. However, in accordance with equation (4), we can make further progress. Specifically, the difference in AF headcounts also will always be positive (i.e., $H_j > H'_j$) if the same bounding condition holds for all unique classes (partitions) of the poverty space. Formally, for all feasible classes of $\theta_{s,l}$ such that $1 \leq s \leq l \leq D$, the poverty comparison is weakly robust if:

$$\Delta_{s,l^*} \equiv \bar{y}(Z_j, \tilde{\theta}_{l^*}) - \bar{y}(Z_{j'}, \tilde{\theta}_s) \geq 0 \quad (5)$$

A useful corollary is that even if this condition is not fulfilled for all feasible combinations of s and

l, we can identify the subset of classes, and corresponding constraints on θ, where the condition does (not) hold. Similarly, if there is extreme overlap in the distribution of headcount rates for the two groups, due to equality or non-dominance, this should be apparent from switches in the sign of equation (5) according to different choices of s and l.

4.2 Stochastic approach

The advantage of assessing robustness via a bounding approach is that poverty headcounts only need be estimated using a vector of equal weights and corresponding cut-offs. This entails $(D^2 + D)/2$ comparisons. The main drawback is that these bounds may be rather uninformative (see Section 5). Another approach is to search the poverty space directly. Since by Property (1) the search space is finite, one strategy is to enumerate all positive Boolean threshold functions in D variables and evaluate robustness point-by-point (as shown in Table 1 for 3 variables). The immediate complication of brute force enumeration is that the number of unique points to visit explodes as the number of deprivation dimensions increases. Counting the number of positive (monotone) Boolean functions is known as Dedekind's problem, for which no concise closed-form general expression exists. An active research agenda is just to count and/or list all unique Boolean threshold functions (e.g., Korshunov and Shmulevich, 2002; de Keijzer et al., 2012; Kurz, 2012). A trivial upper bound is 2^{2^D}, which is the number of all possible Boolean functions in D variables. This hints that, in all but relatively simple situations (e.g., fewer than around five dimensions), full enumeration is not feasible.

A practical alternative to brute force enumeration is stochastic search of a Monte Carlo variety. A well known result from the optimization literature is that pure random search will converge to a neighbourhood ϵ of a global optimum y^* (target) with probability one (Solis and Wets, 1981). So, when designed carefully – i.e., in light of the structure of the problem – these techniques can be extremely powerful. The disadvantage is that the expected number of iterations required to achieve a given level of convergence is inversely proportional to the likelihood of visiting a point in the target region, which typically is exponential in the number of dimensions of the problem.[10] This implies that achieving a desired degree of coverage of the space of definitions (in probability) may be extremely computer-intensive.

Assuming that a stochastic search employs (quasi-)random draws, then a representative sample of the search space should be achieved. This means that observed sample estimates, such as the mean, are expected to be well-behaved in the sense of displaying strong consistency and

[10]As Zabinsky (2011) shows, the probability of failing to sample a point in the target region after g trials is equal to $(1 - \Pr(y^* + \epsilon))^g$.

minimum variance. Even so, strictly speaking, robustness refers to all points in the search space, so even one reversal would be sufficient to induce a conclusion of non-robustness. Consequently, to make rigorous claims about robustness based on stochastic search, some measure of the degree of uncertainty associated with the given set of comparisons is required. A natural metric of uncertainty is the proportion of the search space that remains unseen – i.e., as this declines, the likelihood of a reversal becomes smaller. Although the dimensions of the search space are unlikely to be known in advance, estimation of the unseen part of a finite space has been extensively studied. It is a variant of the balls-in-boxes problem, in which a sample of balls is thrown independently (with no misses) at a fixed number of boxes such that, on each throw, the n^{th} box will be hit with probability p_n. Based on the observed number of boxes occupied after N throws, the problem is to estimate the remaining number of unseen and unoccupied boxes.

An early answer to the unseen mass problem was provided by Alan Turing and Irving John Good during the Second World War, as part of their work to crack ciphers for the Enigma machine during World War II (see Good, 1953; Gale and Sampson, 1995). As shown by Jones (2016), the Good-Turing estimator is simply given by: $\hat{u}_{\text{GT}} = |N_1|/N$, where $|N_1|$ is the number of unique objects that have been seen *only* once; and N is the total number of objects that have been seen, typically the total number of iterations (also see Church and Gale, 1991). The intuition is that as the coverage of the search space increases, the rate at which previously unseen objects are discovered declines and $\hat{u}_{\text{GT}}$ approaches zero asymptotically. While various extensions to the Good-Turing estimator have been proposed, such as to allow the probability measure associated with the search space to be non-uniform (Gandolfi and Sastri, 2004; Jones, 2016), these are not germane in the present application. *A priori*, the search space is probabilistically flat and there is no reason to 'prefer' certain poverty definitions over others to determine robustness.

5 Application

This section summarises the results of an empirical application of the bounding and stochastic search procedures to assess the evolution of multidimensional poverty in Mozambique. This country is not merely chosen for illustration. Mozambique has achieved one of the world's most rapid and sustained rates of per capita economic growth since the end of conflict in 1992. However, recent consumption poverty estimates, based on household survey data from both 2008/09 and 2014/15, have raised concerns as to how well aggregate growth is being translated into broad-based welfare gains (DNEAP, 2010; Arndt et al., 2012b; DEEF, 2016). The latest survey indicates that headcount consumption poverty affected 46% of the population, versus

Table 2: Dashboard of deprivations experienced by households, national means

	1996/97	2002/03	2008/09	2014/15
Member has primary education	64.0	53.3	40.4	32.4
Head can read	47.8	45.6	44.7	41.6
Access to electricity	93.9	91.1	84.8	72.9
Access to clean water	73.0	58.6	57.5	47.7
Quality sanitation	95.6	86.0	82.0	71.6
Quality roofing	78.2	70.8	67.3	58.0
Access to means of transport	82.4	65.6	54.7	55.6
Access to communication tech.	63.1	42.8	37.4	24.6
Owns consumer durables	87.3	79.5	68.7	49.6
Column average	76.1	65.9	59.7	50.5

Notes: Cells indicate the share of households deprived on a given dimension.
Source: own estimates.

53% in 2002.[11] So there has been relatively slow progress in aggregate poverty reduction recently, despite high levels of real GDP growth over the period. Nonetheless, the same surveys indicate more consistent gains in non-consumption dimensions, including access to public goods such as education services. Consequently, it is appropriate to investigate what has happened to multidimensional poverty in Mozambique.

In order to apply the methods developed in Section 4, it is necessary to select the deprivation dimensions to be included. This decision itself can be controversial; however, the feasible choices are often limited by the availability of consistent data over time, as well as exclusion of highly correlated dimensions. The domains selected for the present exercise are summarised in Table 2, which shows the share of households deprived in each of nine individual domains across the four available survey waves. The domains cover: human capital (education); access to electricity, water and sanitation; housing quality; and ownership of productive and durable assets. For each variable, households which are classified as deprived receive a score of one and zero otherwise. Inclusion of dimensions such as these are largely conventional in the field of multidimensional poverty (Alkire et al., 2015).[12]

The table indicates changes in well-being have been heterogeneous. While we see progress in all dimensions over the full 18 year period, the pace of change is inconsistent. This implies that when constructing a multidimensional indicator, the relative importance attributed to different dimensions is likely to matter. A closer look at the data further reveals acute regional differentials. The Southern region, where the capital city is located, and also the region most proximate to

[11] Survey weights are applied in all estimates in this section.
[12] Further details on the construction of the underlying deprivation domains is available on request from the author. See also DEEF (2016).

South Africa, not only shows the lowest deprivation rates on virtually all indicators, but also more consistent improvements over time. In contrast, both the Central and Northern regions display much higher levels of deprivation and relatively slower progress on some indicators, such as access to electricity and housing quality. These regional differences, shown in Appendix Table B1, raise fundamental questions regarding the depth and persistence of regional inequalities. Concretely, despite a generally positive trend across the indicators, the levels of deprivation in the Southern region almost 20 years ago were lower than they are today in the Central and Northern regions across a number of indicators. This motivates a formal analysis of the robustness of multidimensional poverty comparisons, focussing in particular on the regional-temporal comparison of the South in 1996/97 and 2014/15 (denoted S96, S14) versus outcomes in the Centre and North in the same years (denoted C96, C14; and N96, N14).

5.1 Robustness via bounding

Applying the procedure described in Section 4.1, the starting point is to compare upper and lower bounds on the multidimensional poverty headcount for any two groups or time periods. This is implemented as per equation (1c), using a vector of equal weights and thresholds (k) corresponding to the sequential sum of weights. For reference, the raw output from this procedure is reported in Appendix Table B2, showing both national and regional headcount estimates. All headcounts are multiplied by 100 to facilitate presentation.

Focussing on the national results, the 2014/15 headcount rates always appear lower than those of 1996/97; but there is a fairly sharp decline in poverty as the cut-off increases (i.e., as we move from a pure union definition of poverty to a pure intersection definition). Table 3 employs the same national estimates to perform a bounding analysis. To interpret the results, it helps to recall how the bounds are constructed. Recall from Proposition (3) that for a given group j, a poverty definition with SWC of size 2 and LWC of size 4 is bounded by: $\bar{y}(Z_j, \tilde{\theta}_2) \geq \bar{y}(Z_j, \theta_{2,4}) \geq \bar{y}(Z_j, \tilde{\theta}_4)$. To compare poverty between two groups, such as that poverty in j is always greater than or equal to that of j', one takes the difference between the lower bound of j and the upper bound of j'. So, for the same poverty definition (SWC = 2; LWC = 4), the relevant gap is: $\Delta_{2,4} = \bar{y}(Z_j, \tilde{\theta}_4) - \bar{y}(Z'_j, \tilde{\theta}_2)$. Each cell of the table reports comparisons of this form, defining j as 1996/97 and j' as 2014/15 (nationwide). The rows indicate the magnitude of the LWC entering the expression, corresponding to the lower bound headcount estimate for 1996/97, taken from the vector of equal weights with LWC (=SWC) of the indicated size. The columns indicate the magnitude of the SWC entering the expression, referring to the upper bound headcount estimate for 2014/15, taken from the vector of equal weights with SWC (=LWC) of the indicated size. The diagonal reports the difference in headcounts for the equal

Table 3: Dominance comparisons for national headcounts 1996/97 vs. 2014/15

	SWC (s)								
LWC	1	2	3	4	5	6	7	8	9
1	5.1								
2	3.3	15.1							
3	0.6	12.4	21.6						
4	−2.6*	9.1	18.3	27.8					
5	−7.8*	3.9	13.2	22.7	34.1				
6	−15.8*	−4.0*	5.2	14.7	26.1	38.3			
7	−28.5*	−16.8*	−7.6*	1.9	13.3	25.6	38.7		
8	−46.8*	−35.1*	−25.9*	−16.4*	−5.0*	7.2	20.4	31.9	
9	−69.1*	−57.4*	−48.1*	−38.6*	−27.2*	−15.0*	−1.9*	9.6	18.5

* significant at 5% level, after Bonferonni correction

Notes: For each partition defined by the smallest prime implicant s and longest prime implicant l, cells indicate the difference between the lower bound estimate for the headcount poverty in 1996/97 and the upper bound in 2014/15; null hypothesis is that the difference is in the positive domain.

Source: own estimates.

weight vectors of the same size, which corresponds to a unique poverty definition.[13]

While the sign (and magnitude) of the gaps reported in each cell are of primary interest, formal tests against a null hypothesis are helpful. Given the conservative nature of the bounding procedure, I use a one-sided null hypothesis running in the direction of the plausible prior that multidimensional poverty in 1996/97 weakly dominates that in 2014/15.[14] This can be tested using the conventional statistic:

$$T_{s,l^*} = \frac{\Delta_{s,l^*}}{\sqrt{\sigma^2_{\bar{y}_j} + \sigma^2_{\bar{y}_{j'}}}} \sim \mathcal{N}(0,1) \tag{6}$$

in which $\sigma^2_{\bar{y}_j}$ refers to the (estimated) standard error of the headcount estimate for group j; and $\mathcal{N}(0,1)$ is the standard normal distribution with cumulative density Φ. Following Kakwani (1993), the standard errors are approximated as:

$$\sigma_{\bar{y}_j} = \sqrt{\frac{\bar{y}_j(1-\bar{y}_j)}{N_j}} \tag{7}$$

where N_j is the number of observations in group j. Under the one-sided null hypothesis, we assume $\forall(s,l^*) : \Delta_{s,l^*} \geq 0$ and focus on the rejection region associated with poverty definitions

[13] This differs to the bounds on the poverty definition with SWC and LWC of equal size. For instance, the bounds for SWC and LWC of size 3 are: $\bar{y}(Z_j, \tilde{\theta}_3) \geq \bar{y}(Z_j, \theta_{3,3}) \geq \bar{y}(Z_j, \tilde{\theta}_4)$.

[14] Null hypotheses in dominance tests can be constructed in a variety of ways, including sequential tests, as set out in Section 4.2.

where T_{s,l^*} is below a critical value, thereby providing evidence against the null. Applying this approach, Table 3 identifies whether the one-way null hypothesis can be rejected using a probability threshold of $\alpha = 0.001$.[15]

Overall, the table reveals that the null hypothesis cannot be rejected for a large proportion of the partitions of the poverty space. All direct comparisons on the diagonal indicate that poverty in 1996/97 was higher than in 2014/15. Only for those partitions with the property that $l^* - s \geq 3$ is the null hypothesis rejected. Effectively, these cases correspond to poverty definitions with comparatively extreme differences in how different dimensions are weighted – e.g., the partition $s = 1, l = 4$ includes definitions that contain a prime implicant with just one element and another prime implicant with four elements, which implies a factor 4 difference in weights. Thus, if we are willing to concern ourselves with less extreme intersection/union combinations, then the procedure directly indicates the null hypothesis cannot be rejected, meaning headcount poverty in 1996/97 weakly dominates that of 2014/15.

Appendix Tables B3 and B4 replicate the same procedure at the regional level, comparing multidimensional headcount poverty rates in the South in 1996/97 against those in the North and Centre in 2014/15. These findings are less clear-cut. For most partitions (intervals) tested, the gap is in the negative domain, suggesting we cannot be certain that a weak dominance condition holds. By way of contrast, and as shown in Appendix Table B5, a comparison of multidimensional poverty in the North in 1996/97 against that in the South in 2014/15 provides the clearest result. Unsurprisingly, for all but the most extreme partitions (e.g., $s = 2, l^* = 8$), the null hypothesis of weak dominance cannot be rejected. And for all poverty definitions in which the SWC is larger than two, the null hypothesis also cannot be rejected.[16] This confirms the bounding procedure is informative, but principally where headcount estimates between two groups diverge considerably. Nonetheless, the results also suggest the procedure can be helpful to focus stochastic search toward specific partitions of the search space where there is some uncertainty. I now turn to how such search can be conducted.

[15] This choice of threshold is based on a Bonferonni correction, taking into account the fact that $(D^2+D)/2$ = 45 separate hypotheses are being tested; thus: $\alpha = 0.05/45$. As Guo (2009) notes, the Bonferonni correction is highly conservative and is robust to arbitrary dependence across the individual tests.

[16] Note that for all partitions with LWC = 9 the assumption is that the largest winning coalition requires unanimity. Thus, if there are to be no redundant dimensions, it must be that SWC=9 as well. Consequently, all cells in the bottom row of the table *excluding* the equal weight special case can be ignored. (They are only shown for completeness).

5.2 Robustness via stochastic search

To implement the stochastic search, I proceed in four steps: (i) I create a 10,000 $\times$ 9 matrix of weights, in which each row $\vec{\omega}_t$ is a pseudo-random draw on the closed standard 8-simplex from the Dirichlet distribution;[17] (ii) I shrink the weights on each row toward the vector of equal weights by the square of a row-specific random factor distributed uniformly on the unit interval;[18] (iii) associated with the matrix of weights, I choose a vector of cut-offs with elements k_t, each of which is drawn from the uniform distribution and constrained according to pre-specified limitations on the permitted sizes of the SWCs and LWCs (see Section 5.3);[19] and (iv) I apply the combination of weights and cut-offs, which constitute individual poverty definitions $\theta_t = \{\vec{\omega}_t, k_t\}$, to equation (1b); and, for each definition, I collapse over relevant population sub-groups as per equation (1c). For each sub-group (e.g., nationwide, regions, survey rounds etc.), this yields a final vector of 10,000 headcount estimates distributed relatively evenly over the various partitions of interest (see Appendix Table B6).[20] Note that the raw data on deprivations is always held fixed and, for each draw θ_t, the same weights and cut-offs are applied to all sub-groups, implying one can then inspect the variation in headcount estimates between groups for each definition (i.e., to make pointwise comparisons).

Figure 1 plots the distributions of pointwise differences in poverty headcounts derived from this procedure, focussing on the four sub-group comparisons discussed in the previous subsection. The differences are expressed on the natural log scale (approximately percentages), which is helpful as the magnitude of poverty headcount estimates vary significantly with the chosen cut-offs. Visually, some clear patterns emerge. In plots (a) and (b), the distribution of headcounts is almost entirely in the positive domain. In fact, further inspection (see below) reveals that in both cases there is no definition in which poverty in group j' is higher than that in j. Thus appears consistent with the condition of weak pointwise dominance, at least for all 10,000 randomly selected points. In contrast, plots (c) and (d) reveal mass in *both* the positive and negative domains, implying that poverty comparisons do not take a consistent sign and dominance is unlikely to hold.

[17] This is required so that the weights sum to one. Drawing weights for each dimension independently from a uniform distribution will not yield even coverage of the weight space. Note that the first 9 weights and cut-offs are fixed to give the special cases of equal weights and series of natural cut-offs.

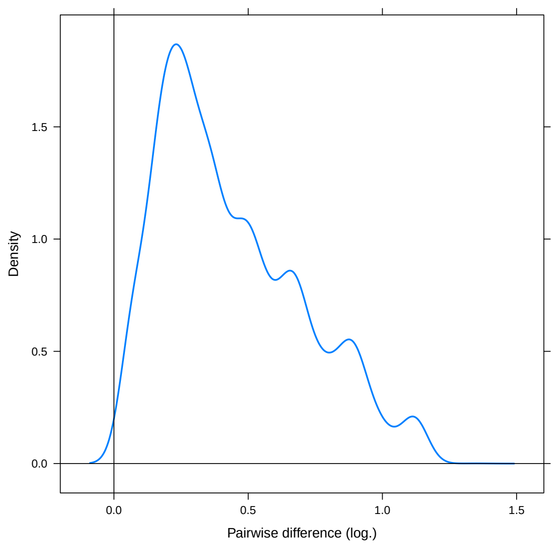
[18] This has the result of permitting different deprivation dimensions to take more similar weights. In the absence of shrinkage, differences in weights between dimensions are more extreme.

[19] In this application the minimal limitations are: $k_t \in [w_{i=1,t}, \sum_{i=1}^{8} w_{i,t}]$, where i ranks weights from smallest to largest. This states that the lower bound on the cut-off is just the smallest weight; the upper bound is the sum of the smallest $D - 1 = 8$ weights, which simply avoids re-sampling the unique definition in which an individual must be deprived in all dimensions to be classified as poor.

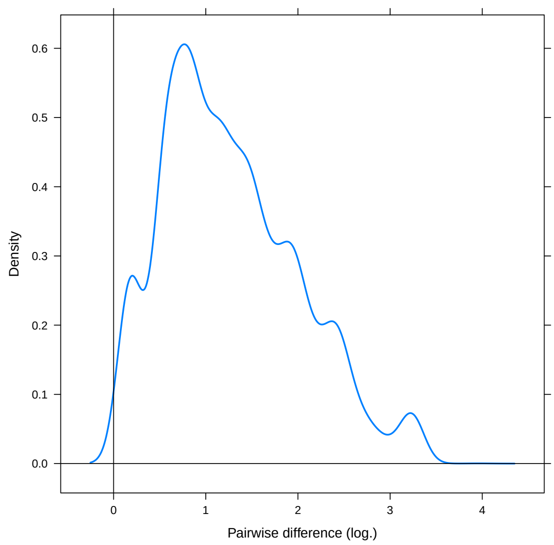
[20] Some sub-groups contain the same members; but, the same poverty definitions are used throughout.

Figure 1: Kernel density plots of pointwise differences in headcounts from stochastic search

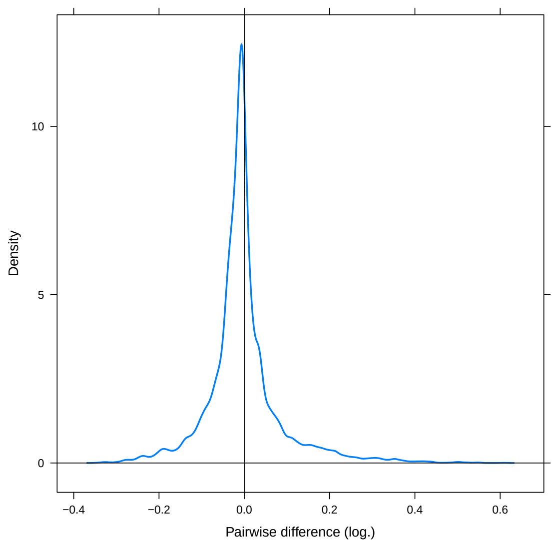
(a) National: 1996/97 - 2014/15



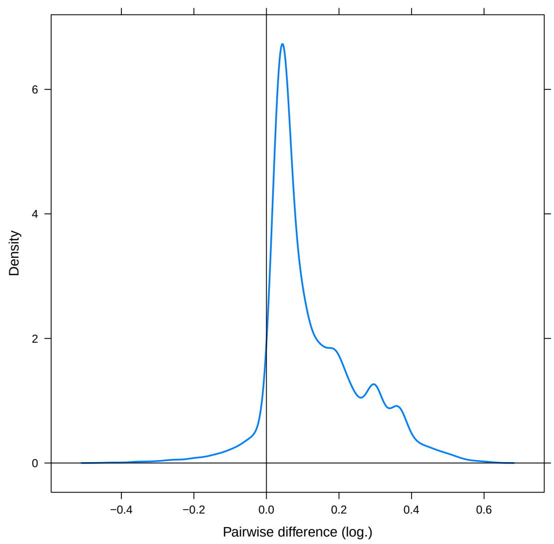
(b) Regional: North 1996/97 - South 2014/15



(c) Regional: South 1996/97 - North 2014/15



(d) Regional: South 1996/97 - Centre 2014/15



Notes: differences are calculated pairwise, based on 10,000 random search points (the same in each plot); in plots (a) and (b) the density plot enters the negative domain only due to smoothing; x-axis is the difference in the natural logarithm of the estimated headcounts.
Source: own estimates.

Table 4: Sequential hypothesis testing procedure

	Hypothesis	Rejection rule
H0	$\forall s : \bar{y}_{s,j} = \bar{y}_{s,j'}$	$\exists s : p_t < \alpha \vee p_t > 1 - \alpha$
H1.1	$\forall s : \bar{y}_{s,j} \geq \bar{y}_{s,j'}$	$\exists s : p_t < \alpha$
H1.1.1	$\forall s : \bar{y}_{s,j} > \bar{y}_{s,j'}$	$\exists s : p_t \leq 1 - \alpha$
H1.2	$\forall s : \bar{y}_{s,j} \leq \bar{y}_{s,j'}$	$\exists s : p_t > 1 - \alpha$
H1.2.1	$\forall s : \bar{y}_{s,j} < \bar{y}_{s,j'}$	$\exists s : p_t \geq \alpha$
H1.3	$\bar{y}_{s,j} \neq \bar{y}_{s,j'}$	-

Notes: H0 is the null hypothesis of equality; H1.1 and H1.2 are level one alternative hypothesis; H1.1.1 and H1.2.1 are level two alternative hypothesis; and H1.3 is the residual alternative hypothesis, corresponding to non-dominance.
Source: own elaboration.

To formally test these visual results, three additional steps are required. First, a relevant hypothesis testing framework must be adopted; second, one or more global test statistics must be chosen; and, third, adjustments must be made to account for unseen elements of the search space (i.e., the missing mass). With respect to the hypothesis testing framework, I adopt a sequential intersection-union testing framework (Bishop et al., 1994; Tse and Zhang, 2004), in which the global null hypothesis is defined as the intersection of local (pointwise) null hypotheses; and the overall alternative hypothesis is the logical union of their complement. In keeping with the previous subsection, local tests are performed using equation (6), providing a vector of associated estimated (lower-tail) probabilities, denoted $p_t = \Pr(Z \leq T_t)$, where Z follows the standard normal distribution; and t denotes the draw from the space of poverty definitions.

To formalize the hypothesis testing procedure, Table 4 sets out the set of relevant null and alternative hypotheses, as well as their corresponding rejection rules, evaluated via the vector of estimated probabilities. The overall null hypothesis (H0) is that the two vectors of headcount estimates being compared are equivalent, such that the distribution of test statistics taken from the vector of pointwise differences will be indistinguishable from a standard normal distribution. This hypothesis can be rejected if either the upper or lower tails of the pointwise difference distribution contain mass above critical thresholds, defined as $\mid T_t \mid \geq \Phi^{-1}(\alpha)$. As before, the critical significance level, α, can be selected to account for the presence of multiple comparisons. The conservative Bonferonni correction is applied again here, yielding the critical threshold $\alpha = .05/10^3$.

If the global null is rejected, the next step is to investigate whether a situation of dominance obtains. This is addressed via the alternative hypotheses H1.1 and H1.2. H1.1 tests for weak pointwise dominance in the direction $j \geq j'$, and is rejected if there is mass in the standardized pointwise difference distribution below $\Phi^{-1}(\alpha)$. If this cannot be rejected we go on to evaluate strict pointwise dominance, which obtains if all estimated probabilities fall above the $1-\alpha$ threshold, meaning that all standardized differences T_t are found in the extreme upper tail of a standard normal distribution. If H1.1 is rejected, we go on to evaluate H1.2, which concerns dominance in the opposite direction. In the case that both H1.1 and H1.2 are rejected, then the distribution of standardized pointwise differences contains mass both above $\Phi^{-1}(1-\alpha)$ and below $\Phi^{-1}(\alpha)$, leading to a conclusion of non-dominance.

Despite the various cases described by Table 4, in practice one only needs to evaluate the share of standardized pointwise differences falling in both of the defined critical regions. Returning to the four comparisons of interest (see Figure 1), Table 5 summarises the main results. The first three rows respectively report basic summary statistics pertaining to the distribution of pointwise differences, namely: their mean (log scale), the proportion in the positive domain, and the mean of the corresponding vector of probabilities. The fourth and fifth rows report the mass observed in the critical regions ($p_t \leq \alpha$, $p_t \geq 1-\alpha$). These provide clear guidance. In the first two columns, all 10,000 pointwise differences are located in the extreme upper tail (critical region) of the standard normal distribution. Correspondingly, we cannot reject hypothesis H1.1.1 and conclude that the multidimensional poverty headcount in 1996/97 strongly dominates that in 2014/15 for all chosen definitions. The same goes for the multidimensional poverty headcount in the North in 1996/97 versus the South in 2014/15. In contrast, the final two columns indicate non-zero mass in *both* the upper and lower critical regions. For instance, in the case of the South 1996/97 versus the North in 2014/15, more than one fifth of all outcomes are found in each critical region. In the case of the South 1996/97 versus the Centre in 2014/15, around 4% of outcomes are found in the lower critical region and 90% are found in the upper critical region. While this indicates that, on average, poverty was higher in the South (1996/97) than in the Centre today (2014/15), we cannot rule out that there are some poverty definitions where the reverse holds.

The final step is to address the uncertainty associated with the stochastic search of the poverty space. Although the number of distinct poverty definitions is finite, they are expected to be very large in number and not amenable to enumeration. Thus, as an approximation to the proportion of the poverty space that has not been seen, I calculate the Good-Turing estimate of missing mass for each vector of pointwise differences. In doing so, it is practical to apply a simple rounding constraint. Indeed, in most empirical analyses, continuous outcomes can be treated as discrete without loss of meaningful information, reflecting the point that outcomes are hardly ever meaningful to an infinite degree of accuracy (for discussion see Bedeian et al., 2009). The rounding rule I use is chosen via Sheppard's approximate correction for calculating distributional moments, based on a discrete approximation to a continuous random variable

Table 5: Results from stochastic search of multidimensional poverty headcount definitions

	M96-M14	N96-S14	S96-N14	S96-C14
Mean gap (log.)	0.437	1.276	0.002	0.125
Proportion gaps > 0	1.000	1.000	0.631	0.943
Mean p_s	1.000	1.000	0.618	0.943
Mean $[p_s \leq \alpha]$	0.000	0.000	0.214	0.039
Mean $[p_s \geq 1-\alpha]$	1.000	1.000	0.366	0.900
Missing mass ($\hat{u}_{\text{GT}}$)	0.001	0.001	0.008	0.009
Adj. mean p_s	0.999	0.999	0.613	0.934
Likelihood $\bar{y}_j > \bar{y}_{j'}$	0.998	0.997	0.226	0.868

Notes: columns refer to group comparisons, where M is Mozambique, N is the Northern region, C is the Central region, S is the Southern region, and numbers refer to the (first) year of the survey; rows report summary measures from the distribution of pointwise differences.
Source: own estimates.

(Wilrich, 2005; Jones, 2016). This suggests headcount poverty estimates (ranging from zero to 100) can be safely rounded to the nearest decimal place. Based on this, the estimates for $\hat{u}_{\text{GT}}$ are all below 1%, suggesting that after 10,000 draws at least 99% of all quantitatively distinct poverty headcount definitions have been seen.

Estimates of unseen mass taken from the GT estimator can be applied directly in conjunction with other estimates. For instance, Manski-type bounds on the mean probability ($\bar{p}$) are given by:

$$\bar{p} \in \left[\hat{\bar{p}}(1-\hat{u}_{\text{GT}}), \hat{\bar{p}} + \hat{u}_{\text{GT}}(1-\hat{\bar{p}})\right] \tag{8}$$

The lower bound from this expression is shown in the penultimate row of Table 5. And the final row transforms this adjusted mean to provide a simple metric of likelihood that poverty in group j exceeds that of j' for any poverty definition. This is calculated as $2p_t - 1$, reflecting the point that when the pointwise gap is zero, the probability associated with the test statistic is 0.5. Both metrics confirm the general conclusion that for the first two columns, there is very strong evidence to suggest the poverty comparisons are strongly robust; for the third column there is no evidence that poverty in the South in 1996/97 was higher than in the North in 2014/15; and in the final column, there is some but not overwhelming evidence to support the conjecture that poverty in the South in 1996/97 was higher than in the Centre in 2014/15. More generally, these range of metrics demonstrate the value of focussing analysis on pointwise differences and how these can be adjusted to address uncertainty due to unseen points.

5.3 Extensions

A number of analytical extensions to the basic applications demonstrated above can be contemplated. First, and in contrast to the bounding approach, the stochastic search procedure is not restricted to choices for f (Definition 2.1) that are only positive Boolean threshold functions (i.e., the AF headcount measure).[21] In fact, *any* functional form can be evaluated in this way as long as the search space is finite and/or the space can be suitably discretized. Consequently, the stochastic search procedure combined with estimates of missing mass provides a quite general approach to evaluating robustness. A natural extension is to the various adjusted AF multidimensional poverty measures. These modify the headcount above to account for the intensity of poverty. Specifically, the row-wise mapping function is adjusted via an intermediate step:

$$\tilde{y}_i = y_i \sum_{d \in D} \omega_d \left(z_{id}\right)^{\alpha}, \ \alpha \in \mathbb{Z}_{\geq 0} \tag{9}$$

and the final outcome is defined in similar fashion to before, as:

$$M_{\alpha}(\alpha, \vec{\omega}, k) = N^{-1} \sum_{i=1}^{I} \tilde{y}_i \tag{10}$$

Table 6 applies the M_0 (adjusted AF headcount measure) to investigate the same comparisons as reviewed previously. Replicating the structure previous table, the same principal findings hold – i.e., there is clear evidence that the comparisons indicated in the first two columns display strong robustness; while the final two columns again indicate non-dominance. The close similarity between the results for the headcount and its adjusted counterpart remains consistent with Theorem 2 in Alkire and Foster (2011), which proves that a dominance relation found under measure H implies a dominance relation under M_0, but only for any choice of the cut-off.

A second extension is to constrain the search space according to specific limits on the poverty definition, such as the minimum intersection of dimensions required for a unit to be classified as poor. In the previous application no limits were applied, meaning that the cut-off was only constrained to be at least as large as the smallest element of the weight vector. Other constraints or rules for the choice of k are outlined in Appendix Table B7, in which the subscripts on the weights indicate their rank order, from smallest to largest – i.e., $w_1 = \min(\vec{w}); w_D = \max(\vec{w})$. Thus, row 2 states that as long as $k \leq 1 - w_d$, then a majority/consensus is not required to classify any unit as poor.[22]

[21] The bounding approach relies on the specific properties of Boolean threshold functions and the cardinality of sets. These do not immediately extend to other choices for f.

[22] In the R code developed to implement the analysis in the present study, functionality is included to select poverty definitions with smallest prime implicant s and longest prime implicant l.

Table 6: Results from stochastic search of M_0 definitions

	M96-M14	N96-S14	S96-N14	S96-C14
Mean gap (log.)	0.592	1.628	-0.001	0.171
Proportion gaps > 0	1.000	1.000	0.595	0.932
Mean p_s	1.000	1.000	0.564	0.930
Mean $[p_s \leq \alpha]$	0.000	0.000	0.154	0.027
Mean $[p_s \geq 1 - \alpha]$	1.000	1.000	0.175	0.822
Missing mass ($\hat{u}_{\text{GT}}$)	0.002	0.005	0.004	0.005
Adj. mean p_s	0.998	0.995	0.562	0.926
Likelihood $\bar{y}_j > \bar{y}_{j'}$	0.995	0.990	0.124	0.851

Notes: columns refer to group comparisons, where M is Mozambique, N is the Northern region, C is the Central region, S is the Southern region, and numbers refer to the (first) year of the survey; rows report summary measures from the distribution of pointwise differences.
Source: own estimates.

Finally, in cases where non-dominance is found, it may be of interest to identify whether there are particular domains of the poverty space in which reversals in the direction of the headcount difference take place. This can be explored in a number of ways, such as calculation of the Euclidean distance between a baseline poverty definition (weight vector and cut-off) and that associated with a reversal. Another simple alternative is to classify the weight vectors according to which deprivation dimension is ranked highest or lowest. This exercise is undertaken in Table 7, which reports the mean log. headcount gap when each dimension (denoted one to nine; in the order shown in Table 2) is ranked lowest or highest in the vector of weights. Focussing on the four group comparisons of earlier interest, the table immediately reveals the particularly crucial role played by the sixth dimension (quality of roofing). When this specific dimension receives the highest weight, the headcount comparisons S96-N14 and S96-C14 fall in the negative domain or are close to zero. Also, when literacy of the household head receives a large weight, the comparisons S96-N14 is significantly negative on average. This highlights the particularly high levels of deprivation associated with access to roofing in regions outside of the South, as well as the gap between the South and other regions on this indicator (see Table B1).

6 Conclusion

The objective of this study was to provide a rigorous framework for evaluating the robustness of multidimensional poverty comparisons. The motivation was that existing approaches have not made sufficient progress. Multivariate stochastic dominance has been found to be generally unwieldy and poorly suited to counting-type multidimensional measures. Alternative approaches

Table 7: Mean headcount gaps (log.) by groups of weight vectors

	When dimension takes largest weight				When dimension takes smallest weight			
Dim.	M96-M14	N96-S14	S96-N14	S96-C14	M96-M14	N96-S14	S96-N14	S96-C14
1	0.587	1.590	-0.006	0.160	0.592	1.749	0.014	0.203
2	0.602	1.663	0.024	0.190	0.585	1.541	-0.062	0.122
3	0.588	1.686	-0.007	0.185	0.604	1.512	0.018	0.151
4	0.589	1.588	0.001	0.178	0.585	1.673	-0.000	0.157
5	0.586	1.694	-0.021	0.171	0.587	1.507	0.028	0.162
6	0.591	1.563	0.067	0.234	0.599	1.740	-0.144	0.027
7	0.596	1.715	-0.041	0.112	0.590	1.429	0.075	0.287
8	0.588	1.518	-0.003	0.154	0.596	1.871	0.021	0.227
9	0.596	1.618	-0.020	0.157	0.585	1.632	0.047	0.207
Mean	0.592	1.626	-0.001	0.171	0.591	1.628	-0.000	0.171

Notes: cells give the mean headcount difference for groups of poverty definitions (taken from the stochastic search procedure), where groups are created according to which dimension has the largest weight (first four columns) or smallest weight (final four columns); rows indicate the relevant dimensions, given as per the order in Table 2; sub-columns refer to group comparisons, where M is Mozambique, N is the Northern region, C is the Central region, S is the Southern region, and numbers refer to the (first) year of the survey.
Source: own estimates.

based on sensitivity-type analysis have become popular, but suffer from being rather limited in scope. The contribution provided herein stemmed from situating the analysis of robustness within the domain of Boolean threshold functions, which were shown to be equivalent to the Alkire-Foster headcount metric. Two particular properties of these functions stand out – they are countably finite in number and can be classified according to the smallest and largest number of prime implicants they contain. It was argued that these properties naturally suggest two complementary approaches to evaluating robustness, defined in terms of pointwise dominance. These were: (i) a bounding approach, in which upper and lower bounds on headcount estimates can be evaluated using a vector of equal weights across at most $(D^2 + D)/2$ partitions; and (ii) stochastic search, supported by the Good-Turing measure of missing mass. The latter is crucial – it estimates the proportion of the space of poverty definitions that remains unvisited, from which Manski-type bounds on poverty headcounts (or test statistics) can be constructed. Functions to apply these approaches are provided and set out in Appendix A.

The analytical value of these two approaches was examined in an application to poverty in Mozambique. Based on nine dimensions of deprivation, the methods confirm that, on aggregate, multidimensional poverty in 1996/97 was strictly higher than it is today, for any possible counting definition. However, the same methods reveal persistent regional asymmetries. Specifically, we cannot rule out that poverty in 2014/15 in the Central and Northern regions of the country is at least as high as was observed in the Southern region almost 20 years ago. Supplementary analysis showed that switches in the direction of these poverty comparisons is driven by weights applied to two particular dimensions – housing quality and literacy of the head of household. In sum, the methods set out in this paper constitute a practical and quite general framework for evaluating the robustness of multidimensional poverty comparisons.

References

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

Alkire, S., Roche, J.M., Ballon, P., Foster, J., Santos, M.E. and Seth, S. (2015). *Multidimensional poverty measurement and analysis*. Oxford University Press, USA.

Alkire, S. and Santos, M.E. (2014). Measuring acute poverty in the developing world: Robustness and scope of the multidimensional poverty index. *World Development*, 59:251–274.

Alkire, S., Santos, M., Seth, S. and Yalonetzky, G. (2010). Is the multidimensional poverty index robust to different weights? OPIH Research in Progress, Oxford Poverty & Human Development Initiative (OPHI), University of Oxford. URL www.ophi.org.uk/publications/ophi-research-in-progress.

Angulo, R., Díaz, Y. and Pardo, R. (2016). The Colombian multidimensional poverty index: Measuring poverty in a public policy context. *Social Indicators Research*, 127(1):1–38.

Arndt, C., Distante, R., Hussain, M.A., Østerdal, L.P., Huong, P.L. and Ibraimo, M. (2012a). Ordinal welfare comparisons with multiple discrete indicators: A first order dominance approach and application to child poverty. *World Development*, 40(11):2290–2301.

Arndt, C., Hussain, M.A., Jones, E.S., Nhate, V., Tarp, F. and Thurlow, J. (2012b). Explaining the evolution of poverty: the case of Mozambique. *American Journal of Agricultural Economics*. doi:10.1093/ajae/aas022.

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

Batana, Y.M. (2013). Multidimensional measurement of poverty among women in Sub-Saharan Africa. *Social Indicators Research*, 112(2):337–362.

Bedeian, A.G., Sturman, M.C. and Streiner, D.L. (2009). Decimal dust, significant digits, and the search for stars. *Organizational Research Methods*, 12(4):687–694.

Bennett, C. and Mitra, S. (2013). Multidimensional poverty: measurement, estimation and inference. *Econometric Reviews*, 32(1):57–83.

Bishop, J., Chakraborti, S. and Thistle, P. (1994). Relative inequality, absolute inequality, and welfare: Large sample tests for partial orders. *Bulletin of Economic Research*, 46(1):41–59.

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

Church, K. and Gale, W. (1991). A comparison of the enhanced Good-Turing and deleted estimation methods for estimating probabilities of English bigrams. *Computer Speech & Language*, 5(1):19–54.

Crama, Y. and Hammer, P.L. (2011). *Boolean Functions: Theory, Algorithms, and Applications*. Cambridge: Cambridge University Press.

Crawford, I. (2005). A nonparametric test of stochastic dominance in multivariate distributions. School of Economics Discussion Papers 1205, School of Economics, University of Surrey.

DEEF (2016). Pobreza e bem-estar em Moçambique: Quarta avaliação nacional (IOF 2104/15). Technical report, Direcção de Estudos Económicos e Financeiros, Ministério de Economia e Finanças, República de Moçambique. URL https://www.wider.unu.edu/sites/default/files/Final_QUARTA%20AVALIA%C3%87AO%20NACIONAL%20DA%20POBREZA_2016-10-26_2.pdf.

DNEAP (2010). Pobreza e bem-estar em Moçambique: Terceira avaliação nacional. Technical report, Ministry of Planning and Development, Government of Mozambique. URL www.dneapmpd.gov.mz/index.php?option=com_docman&task=doc_download&gid=133&Itemid=54.

Duclos, J., Sahn, D. and Younger, S. (2006). Robust multidimensional poverty comparisons. *The Economic Journal*, 116(514):943–968.

Elkind, E., Goldberg, L.A., Goldberg, P.W. and Wooldridge, M. (2009). On the computational complexity of weighted voting games. *Annals of Mathematics and Artificial Intelligence*, 56(2):109–131.

Foster, J.E., McGillivray, M. and Seth, S. (2013). Composite indices: rank robustness, statistical association, and redundancy. *Econometric reviews*, 32(1):35–56.

Gale, W. and Sampson, G. (1995). Good-Turing frequency estimation without tears. *Journal of Quantitative Linguistics*, 2(3):217–237.

Gandolfi, A. and Sastri, C.C. (2004). Nonparametric estimations about species not observed in a random sample. *Milan Journal of Mathematics*, 72(1):81–105.

Good, I. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, 40(3-4):237–264.

Guo, W. (2009). A note on adaptive Bonferroni and Holm procedures under dependence. *Biometrika*, 96(4):1012–1018.

Hansen, L., Holt, C. and Peled, D. (1978). A note on first degree stochastic dominance. *Economics Letters*, 1(4):315–319.

Helton, J., Johnson, J., Sallaberry, C. and Storlie, C. (2006). Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, 91(10):1175–1209.

Jones, S. (2016). Measuring what's missing: practical estimates of coverage for stochastic simulations. *Journal of Statistical Computation and Simulation*, 86(9):1660–1672.

Kakwani, N. (1993). Statistical inference in the measurement of poverty. *Review of economics and statistics*, 75(4):632–639.

de Keijzer, B., Klos, T.B. and Zhang, Y. (2012). Solving weighted voting game design problems optimally: Representations, synthesis, and enumeration. ERIM Report Series Reference No. ERS-2012-006-LIS, Erasmus Research Institute of Management. URL http://arxiv.org/abs/1204.5213.

Korshunov, A. and Shmulevich, I. (2002). On the distribution of the number of monotone Boolean functions relative to the number of lower units. *Discrete mathematics*, 257(2):463–480.

Kurz, S. (2012). On minimum sum representations for weighted voting games. *Annals of Operations Research*, 196:361–369.

Lasso de la Vega, C. (2010). Counting poverty orderings and deprivation curves. In J. Bishop (Ed.), *Research on Economic Inequality*, volume 18, chapter 7, pp. 153–172. Emerald Group Publishing Limited.

Maasoumi, E. (1986). The measurement and decomposition of multi-dimensional inequality. *Econometrica*, 54:991–997.

Maasoumi, E. and Millimet, D.L. (2005). Robust inference concerning recent trends in us environmental quality. *Journal of Applied Econometrics*, 20(1):55–77.

McCaig, B. and Yatchew, A. (2007). International welfare comparisons and nonparametric testing of multivariate stochastic dominance. *Journal of Applied Econometrics*, 22(5):951–969.

Ravallion, M. (2012). Mashup indices of development. *The World Bank Research Observer*, 27(1):1–32.

Saltelli, A. and Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12):1508–1517.

Saltelli, A., Chan, K. and Scott, E.M. (2009). *Sensitivity Analysis*. Wiley.

Solis, F. and Wets, R. (1981). Minimization by random search techniques. *Mathematics of Operations Research*, 6(1):19–30.

Taylor, A. and Zwicker, W. (1992). A characterization of weighted voting. *Proceedings of the American mathematical society*, 115(4):1089–1094.

Tse, Y. and Zhang, X. (2004). A Monte Carlo investigation of some tests for stochastic dominance. *Journal of Statistical Computation and Simulation*, 74(5):361–378.

Wilrich, P.T. (2005). Rounding of measurement values or derived values. *Measurement*, 37(1):21–30.

Yalonetzky, G. (2011). Conditions for the most robust poverty comparisons using the Alkire-Foster family of measures. OPHI Working Paper 44b, Oxford Poverty & Human Development Initiative (OPHI), University of Oxford.

—— (2013). Stochastic dominance with ordinal variables: conditions and a test. *Econometric Reviews*, 32(1):126–163.

Zabinsky, Z.B. (2011). Random search algorithms. In J.J. Cochran, J. Louis Anthony Cox, P. Keskinocak, J.P. Kharoufeh and J.C. Smith (Eds.), *Wiley Encyclopedia of Operations Research and Management Science*. Wiley.

Zuckerman, M., Faliszewski, P., Bachrach, Y. and Elkind, E. (2012). Manipulating the quota in weighted voting games. *Artificial Intelligence*, 180:1–19.

A Appendix: Documentation of functions in R

The R code containing the three primary functions used to implement the stochastic search procedures described in Section 5 are found in the supplementary file <mpdFunctions.R>. Documentation for each function is found below, in standard format. The first two (`genweights`, `unseen`) are auxiliary functions, both of which are called by the main (wrapper) function, `robust(...)`.

Note that, as written, the functions assume binary inputs (deprivations); but this can be easily adapted to suit continuous outcomes and variable poverty thresholds.

The complete code stream associated with the present analysis is also available on request from the author.

`genweights` *Random generator of Boolean threshold functions*

Description

Generates a chosen number of positive monotone Boolean threshold functions (poverty headcount definitions of the Alkire-Foster class), typically for use in multidimensional well-being calculations.

Usage

```
genweights(repsN = 1, k = 2, swc = NA, lwc = NA, shrink = 0, round = 5)
```

Arguments

`repsN`	Number of (unique) threshold functions to generate
`k`	Number of dimensions (literals) spanned by the function
`swc`	Smallest winning coalition permitted (defaults to 1)
`lwc`	Largest winning coalition permitted (defaults to `k`)
`shrink`	Shrinkage factor applied to weights (toward equal weights; defaults to 0)
`round`	Round weights to this number of decimal places (defaults to 5)

Details

This function can be used as an input into stochastic search procedures.

Weights are randomly drawn on the closed standard $k - 1$ simplex from the Dirichlet distribution, meaning they sum to one. By default, the first k rows are vectors of equal weights and natural cut-offs increasing in order $(1/k, \cdots, k/k)$.

The shrinkage factor is a fixed exponent that is applied to a uniformly distributed random variable (which varies by row) and is applied to each row of weights so as to shrink the given weight vector toward the vector of equal weights. Where this factor is zero, no shrinkage is applied.

Value

Outputs a data.frame with `repsN` rows and $k + 4$ columns containing the following:

`weights` A set of k columns, named d.1 to d.k, giving the chosen weights for each dimension

`min` Vector of length `repsN` giving the calculated minimum cut-off associated with the chosen constraint on the smallest winning coalition (SWC) for the given set of weights (in the same row)

`max` Vector of length `repsN` giving the calculated maximum cut-off associated with the chosen constraint on the largest winning coalition (LWC) for the given set of weights (in the same row)

`warning` Vector of length `repsN` indicating TRUE if the calculated LWC for the given row of the weight matrix (i.e., poverty definition) is smaller than the calculated SWC

`cut` Vector of length `repsN` giving the chosen thresholds (cut-offs) associated with the weights (in the same row)

Note

Each row of the output data.frame refers to a distinct Boolean threshold function (poverty definition).

Author(s)

Anonymous for now

Examples

```
# Set seed for replicability
```

```
  set.seed(12345)

# Generates 10 threshold functions spanning 5 dimensions
## NB. first 5 definitions will be equal weight vectors

  genweights(repsN = 10, k = 5, swc = 1, lwc = 5, shrink = 0, round = 5)
```

`unseen` *Estimates of missing or unseen mass*

Description

Provides Good-Turing type estimates of missing mass, which is the proportion of a (finite) vector space that has not yet been observed after N draws

Usage

```
unseen(x, digits = 2)
```

Arguments

`x`	input vector of length N over which function is to be applied
`digits`	round to the nearest number of decimal places (default = 2)

Value

Returns a vector with three elements:

`lb`	The lower bound estimate, which is the Good-Turing estimator
`gs`	The Gandolfi & Sastri (2004) adjusted estimate
`ub`	The upper bound estimate, in which the tuning parameter γ (see Jones, 2016) is set to one

Author(s)

Anonymous for now

References

Gale, W. and Sampson, G. (1995). Good-Turing frequency estimation without tears. Journal of Quantitative Linguistics, 2(3):217–237.

Gandolfi, A. and Sastri, C.C. (2004). Nonparametric estimations about species not observed in a random sample. Milan Journal of Mathematics, 72(1):81–105.

Jones, S. (2016). Measuring what's missing: practical estimates of coverage for stochastic simulations. Journal of Statistical Computation and Simulation, 86(9):1660–1672

Examples

```
# Set seed for replicability

  set.seed(12345)

# draw a vector from the uniform distribution

  random <- runif(1000)

# get the estimates of missing mass, rounding the input vector to two decimal places

  unseen( x = random, digits = 2 )
```

`robust` *Robustness of multidimensional poverty comparisons*

Description

Core function used to analyse the robustness of multidimensional well-being comparisons using the Alkire-Foster class of measures (see Alkire and Foster, 2011).

Usage

```
robust(data, group, weight = NA, m = 0, reps = 100,
          swc = NA, lwc = NA, shrink = 2, dropWarning = FALSE,
          unseen_type = c("lb","gs","ub"), unseen_digits = 1,
          p_adj = FALSE)
```

Arguments

`data`	A data.frame with N units (rows) containing the raw data on deprivation outcomes, spanning k dimensions, assumed to be binary indicators where 1 indicates the unit is deprived and 0 otherwise
`group`	A vector of length N, encoded as a factor, which classifies units into groups
`weight`	A vector of length N containing sample weights for each unit (ignored if set to NA)
`m`	A binary indicator that takes a value of 0 if the Alkire-Foster headcount measure (H) is to be calcuated and 1 if the Alkire-Foster adjusted headcount (M0) is to be calculated
`reps`	The number of poverty definitions to be evaluated. This parameter is passed to function genweights(...) as `repsN`
`swc`	The smallest winning coalition (SWC) permitted in each poverty definition. This parameter is passed to function genweights(...) as `swc`
`lwc`	The largest winning coalition (LWC) permitted in each poverty definition. This parameter is passed to function genweights(...) as `lwc`
`shrink`	The shrinkage parameter passed to function genweights(...) as `shrink`
`dropWarning`	If set to TRUE, all weight vectors that do not meet the constraints set of SWC and LWC choices will be dropped
`unseen_type`	Choice of which estimator of missing mass to use in subsequent calculations
`unseen_digits`	Rounding rule applied in calculation of estimates of missing mass
`p_adj`	If set to TRUE, the Benjamini and Hochberg (1995) adjustment is applied to the estimated probability values for headcount differences

Value

Outputs a list with the following components:

`weights`	Data.frame containing the weights and cut-offs associated with each poverty definition (in rows), derived from function genweights(...). Additional information includes the length of the SWC and LWC contained in each definition
`outcome`	Data.frame containing the `reps` poverty estimates for each group (rows are poverty definitions, columns are the chosen groups)
`outcome.se`	Data.frame of estimates standard errors that correspond to the data.frame of poverty estimates (above)
`pwise.d`	Data.frame of absolute pairwise differences in poverty measures across groups (rows are poverty definitions, columns are paired groups)

`pwise.dln`	Data.frame of absolute pairwise differences in the natural logarithm of poverty measures across groups (rows are poverty definitions, columns are paired groups)
`pwise.z`	Data.frame of z-scores corresponding to the pairwise differences in poverty measures across groups (rows are poverty definitions, columns are paired groups)
`pwise.p`	Data.frame of probabilities corresponding to the above z-scores (rows are poverty definitions, columns are paired groups), derived from the standard normal distribution function
`pwise.pLB`	Data.frame of lower bound probabilities corresponding to the above z-scores (adjusted for missing mass)
`pwise.pUB`	Data.frame of upper bound probabilities corresponding to the above z-scores (adjusted for missing mass)
`pwise.s`	Data.frame of values that take the value TRUE if the poverty difference is greater than zero and FALSE otherwise
`pwise.sLB`	Data.frame of dummy variables adjusted for missing mass, the column means of which give the lower bound share of poverty definitions for which one group is greater than the other
`pwise.sUB`	Data.frame of dummy variables adjusted for missing mass, the column means of which give the upper bound share of poverty definitions for which one group is greater than the other
`pwise.u`	Data.frame of estimates of missing mass pertaining to each column of pairwise differences in poverty measures. Rows are alternative missing mass estimators, as described under function unseen(...)

Author(s)

Anonymous for now

References

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

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), 57(1): 289–300.

Examples

```
# Set seed for replicability:

  set.seed(set.seed(13254)

# Create matrix of binary deprivations:
## (k=5 dimensions, n=1000 rows, 2 groups -- named A and B)

  k <- 5
  n <- 1000
  correl <- matrix(runif(k^2),nrow=k)
  correl[lower.tri(correl)] <- 0
  means <- runif(5)-.5
  library(MASS)
  data <- mvrnorm(n,means,correl)
  data <- as.data.frame( 1*(data > 0) )

  group <- 1*(runif(n)>.5)
  group[group==1] <- "A"
  group[group==0] <- "B"

# Run function, with 500 poverty definitions:

  output <- robust(data = data, group = group, reps=500)
  ls(output)
  hist(output$pwise.z) # z-scores associated with pairwise difference
```

B Appendix: Additional tables

Table B1: Dashboard of deprivations experienced by households, regional means

	North		Center		South	
	1996/97	2014/15	1996/97	2014/15	1996/97	2014/15
Member has primary edu.	72.0	43.0	70.1	35.0	45.3	11.9
Head can read	53.1	50.9	52.0	41.9	35.0	27.2
Access to electricity	96.6	79.8	97.2	82.9	85.5	43.8
Access to clean water	80.0	57.9	78.3	54.1	56.6	20.4
Quality sanitation	98.5	79.3	98.3	80.1	87.9	44.3
Quality roofing	95.9	77.1	90.4	65.8	38.7	15.2
Access to means of transport	83.9	59.0	81.5	45.2	82.0	70.7
Access to communication tech.	73.8	36.0	69.4	26.2	40.8	5.0
Owns consumer durables	91.2	59.6	93.8	55.2	72.7	24.3
Column average	82.8	60.3	81.2	54.1	60.5	29.2

Notes: Cells indicate the share of households deprived on a given dimension.
Source: own estimates.

Table B2: Estimates of multidimensional headcount poverty, vector of equal weights

	North		Center		South		National	
k	1996/97	2014/15	1996/97	2014/15	1996/97	2014/15	1996/97	2014/15
0.1	100.0	97.4	99.8	95.3	97.4	86.9	99.2	94.1
0.2	99.6	91.7	99.3	87.7	92.0	58.4	97.4	82.3
0.3	98.8	84.4	98.0	80.5	84.9	42.5	94.7	73.1
0.4	97.3	76.3	95.4	71.0	78.4	31.0	91.5	63.6
0.6	94.7	66.2	92.7	57.5	66.6	21.7	86.3	52.2
0.7	89.4	54.2	86.7	43.4	52.3	12.8	78.3	40.0
0.8	77.2	38.2	74.2	28.7	38.7	6.8	65.6	26.9
0.9	56.6	22.8	55.4	16.5	23.8	2.2	47.2	15.3
1.0	31.4	11.3	29.5	5.9	10.5	0.6	25.0	6.5

Notes: Cells indicate the share of households classified as multidimensionally poor for a vector of equal weights and cut-off k, as indicated.
Source: own estimates.

Table B3: Dominance comparisons for national headcounts
South 1996/97 vs. North 2014/15

	SWC								
LWC	1	2	3	4	5	6	7	8	9
1	−0.03								
2	−5.44*	0.3							
3	−12.50*	−6.7*	0.5						
4	−19.01*	−13.3*	−6.0*	2.1					
5	−30.87*	−25.1*	−17.9*	−9.8*	0.3				
6	−45.12*	−39.4*	−32.1*	−24.0*	−13.9*	−1.9*			
7	−58.78*	−53.0*	−45.8*	−37.7*	−27.6*	−15.6*	0.5		
8	−73.68*	−67.9*	−60.7*	−52.6*	−42.5*	−30.5*	−14.4*	0.9	
9	−86.94*	−81.2*	−73.9*	−65.9*	−55.7*	−43.7*	−27.7*	−12.4*	−0.8*

* significant at 5% level, after Bonferonni correction

Notes: For each partition defined by the smallest prime implicant s and longest prime implicant l, cells indicate the difference between the lower bound estimate for the headcount poverty in 1996/97 (Southern region) and the upper bound in 2014/15 (Northern region); null hypothesis is that the difference is in the positive domain.

Source: own estimates.

Table B4: Dominance comparisons for regional headcounts
South 1996/97 vs. Center 2014/15

	SWC								
LWC	1	2	3	4	5	6	7	8	9
1	2.1								
2	−3.3*	4.3							
3	−10.4*	−2.8*	4.4						
4	−16.9*	−9.3*	−2.1*	7.5					
5	−28.7*	−21.2*	−13.9*	−4.4*	9.0				
6	−43.0*	−35.4*	−28.2*	−18.6*	−5.2*	8.9			
7	−56.7*	−49.1*	−41.8*	−32.3*	−18.9*	−4.7*	10.0		
8	−71.6*	−64.0*	−56.7*	−47.2*	−33.8*	−19.6*	−4.9*	7.3	
9	−84.8*	−77.2*	−70.0*	−60.5*	−47.0*	−32.9*	−18.2*	−6.0*	4.6

* significant at 5% level, after Bonferonni correction

Notes: For each partition defined by the smallest prime implicant s and longest prime implicant l, cells indicate the difference between the lower bound estimate for the headcount poverty in 1996/97 (Southern region) and the upper bound in 2014/15 (Central region); null hypothesis is that the difference is in the positive domain.

Source: own estimates.

Table B5: Dominance comparisons for regional headcounts
North 1996/97 vs. South 2014/15

	SWC								
LWC	1	2	3	4	5	6	7	8	9
1	13.1								
2	12.8	41.2							
3	11.9	40.3	56.3						
4	10.5	38.9	54.8	66.3					
5	7.9	36.3	52.2	63.7	73.1				
6	2.5	31.0	46.9	58.4	67.7	76.6			
7	−9.7*	18.8	34.7	46.2	55.5	64.4	70.4		
8	−30.3*	−1.8*	14.1	25.6	34.9	43.8	49.8	54.4	
9	−55.5*	−27.0*	−11.1*	0.4	9.7	18.6	24.6	29.2	30.8

* significant at 5% level, after Bonferonni correction

Notes: For each partition defined by the smallest prime implicant s and longest prime implicant l, cells indicate the difference between the lower bound estimate for the headcount poverty in 1996/97 (Northern region) and the upper bound in 2014/15 (South region); null hypothesis is that the difference is in the positive domain.

Source: own estimates.

Table B6: Coverage of stochastic search procedure

	SWC								
LWC	1	2	3	4	5	6	7	8	9
1	1	0	0	0	0	0	0	0	0
2	571	187	0	0	0	0	0	0	0
3	555	216	179	0	0	0	0	0	0
4	552	263	177	147	0	0	0	0	0
5	508	346	166	137	151	0	0	0	0
6	350	490	299	161	158	163	0	0	0
7	199	485	413	250	167	177	165	0	0
8	76	312	502	461	374	250	203	188	0
9	0	0	0	0	0	0	0	0	1

Notes: For each partition defined by the smallest prime implicant s and longest prime implicant l, cells indicate the number of search points evaluated; partitions with LWC = 9 imply that a winning coalition requires unanimity, so to ensure no redundant dimensions it must be that SWC= 9.

Source: own estimates.

Table B7: Examples of meaningful restrictions on the choice of cut-off (k)

	Restriction	Min ($k \geq$)	Max ($k \leq$)
1.	None	w_1	1
2.	Not unanimous	w_1	$1 - w_d$
3.	No dictator	$w_1 + w_D$	1
4.	SWC = s	$w_1 + \sum_{i=1}^{s-1} w_{D-i-1}$	1
5.	LWC = l	w_1	$\sum_{i=1}^{l} w_i$

Notes: subscripts on weights indicates their rank order, from smallest to largest.
Source: own estimates.