Econometric Reviews

ISSN: 0747-4938 (Print) 1532-4168 (Online) Journal homepage: https://www.tandfonline.com/loi/lecr20

Multidimensional Poverty: Measurement, Estimation, and Inference

Christopher J. Bennett & Shabana Mitra

To cite this article: Christopher J. Bennett & Shabana Mitra (2013) Multidimensional Poverty: Measurement, Estimation, and Inference, Econometric Reviews, 32:1, 57-83, DOI: 10.1080/07474938.2012.690331

To link to this article: https://doi.org/10.1080/07474938.2012.690331

Published online: 01 Aug 2012.

Econometric Reviews, 32(1):57–83, 2013

ISSN: 0747-4938 print/1532-4168 online
DOI: 10.1080/07474938.2012.690331

MULTIDIMENSIONAL POVERTY: MEASUREMENT, ESTIMATION, AND INFERENCE

Christopher J. Bennett[1] **and Shabana Mitra**[2]

[1]*Department of Economics, Vanderbilt University, Nashville, Tennessee, USA*
[2]*Poverty Reduction and Equity Unit, The World Bank, Washington, District of Columbia, USA*

□ *Multidimensional poverty measures give rise to a host of statistical hypotheses that are of interest to applied economists and policy-makers alike. In the specific context of the generalized Alkire–Foster (Alkire and Foster, 2008) class of measures, we show that many of these hypotheses can be treated in a unified manner and also tested simultaneously using a minimum p-value approach. When applied to study the relative state of poverty among Hindus and Muslims in India, these tests reveal novel insights into the plight of the poor which are not otherwise captured by traditional univariate approaches.*

Keywords FGT measures; Multidimensional poverty measurement; Multiple testing.

JEL Classification C1; C12; I3; I32; O1.

1. INTRODUCTION

Multidimensional poverty measures give rise to a rich set of testable hypotheses. In this article, we formulate a variety of these hypotheses—in the specific context of the Alkire and Foster (AF) (2008) measure—that are likely to be of particular interest to applied economists and policy-makers alike. More importantly, we introduce a unified framework for developing statistical tests of these and other related hypotheses.

Governments of several nations, including those of India and Mexico, as well as numerous nongovernmental agencies are in the process of adopting multidimensional measures of poverty to complement their traditional income (or consumption) analysis. The adoption of

The views expressed here are those of the author and should not be attributed to the World Bank or any affiliated organization.

Address correspondence to Christopher J. Bennett, Department of Economics, Vanderbilt University, 415 Calhoun Hall, Nashville, TN 37240, USA; E-mail: chris.bennett@vanderbilt.edu

a multidimensional approach is largely in response to arguments that income alone does not completely identify the poor, and that there are other dimensions which are relevant to the wellbeing of individuals. The goal of a multidimensional approach to poverty analysis, therefore, is to move beyond the traditional univariate approach to incorporate additional relevant indicators of wellbeing.

Following Sen (1976), poverty measurement has been viewed as a two step procedure involving both an identification and an aggregation step. Identification grapples with the question: *Who is poor*? This involves the notion of poverty lines, whereby the individuals below a poverty line are identified as poor. In the multidimensional case, however, two cutoffs must be considered for identification. First, for each dimension, a dimension-specific poverty line identifies the individuals deprived in that particular dimension. The second cutoff determines the number of dimensions, k, in which one must be deprived before they are considered (multidimensionally) poor. The measures of Bourguingnon and Chakravarty (2005) and Tsui (2002), for example, adopt a *union* approach to identification whereby any individual who is deprived in at least one dimension is considered poor. In other words, their second cutoff is simply one dimension of deprivation. In practice, however, the union approach often identifies substantially high proportions of various populations as poor. In some instances, the union approach has been found to identify more than 90% of a population as poor (Mitra, 2011).

Alkire and Foster (2008) have recently proposed a new class of multidimensional poverty measures based on the FGT (Foster et al., 1984) class of unidimensional poverty meaures. The AF measure is remarkably simple, both conceptually and computationally. In the identification stage, the AF measure involves selecting the second cutoff k to be any value between one (the union approach) and the maximum number of dimensions d (the intersection approach). The aggregation stage is then based on the FGT framework and thus retains many of the desirable properties of the FGT class of measures, including decomposability of the overall poverty measures among subgroups of the population. This property is essential, for example, when one wishes to compare poverty across subregions or ethnic groups.

The Alkire–Foster methodology has also recently been applied in several empirical studies; see, e.g., Alkire and Seth (2008), Santos and Ura (2008), and Batana (2008). However, these articles are primarily descriptive in nature due, largely, to a lack of available statistical testing procedures.[1] The present article fills this void not only by formulating a

[1]In contrast, statistical tests relating to the univariate approach to poverty analysis are well established; see, for example, Anderson (1996), Davidson and Duclos (2000), Barrett and Donald (2003), and Linton et al. (2005).

variety of novel and interesting statistical hypotheses in this context, but also by contributing to the literature a general framework for developing statistical tests of these and related hypotheses. A distinguishing feature of our work is our emphasis on multiple testing procedures that enable users to identify from within a collection of hypotheses those which are not supported by the data. It is our contention that multiple testing procedures are of particular relevance in the context of multidimensional poverty analysis. Inferring, for instance, the specific range of poverty lines over which a poverty ordering holds, the subcollection of measures over which a poverty ordering holds, or the specific dimensions (e.g., income, health, education) in which a country or region is underperforming, are of greater policy relevance than whether the ordering fails for some (possibly unidentified) poverty line, measure, or dimension. In contrast, most procedures currently applied in the context of poverty analysis are joint tests which permit us to draw less informative inferences. Batana's (2008), for example, tests whether a poverty ordering based on the headcount ratio is consistent over a collection of poverty lines. Batana (2008) approach, which is based on the empirical likelihood ratio test developed in Davidson and Duclos (2006), allows him to infer only that the hypothesized ordering is violated without necessarily providing any compelling evidence concerning *which* poverty line(s) might suggest a reversal in the hypothesized ordering.

In contrast, we show that the recently introduced multiple testing procedures of Bennett (2010) are particularly well-suited to simultaneous testing of the hypotheses that arise naturally in the context of multidimensional measures.[2] The principle advantage of adopting multiple testing procedures is that, unlike the popular Wald-type tests (e.g., Wolak, 1989; Kodde and Palm, 1986), for example, they offer compelling evidence concerning the source(s) of rejection whenever rejection of the joint intersection hypothesis occurs. The advantage of adopting the MinP procedure of Bennett (2010), in particular, is that this test is shown capable of correctly identifying more false hypotheses (sources of rejection) than related procedures. Specific examples treated in this article include (though are not limited to) simultaneous tests of the poverty orderings for various parameterizations of the AK measure (e.g., robustness to choice to poverty lines and/or k), simultaneous tests of poverty orderings of various populations relative to a benchmark population, and simultaneous tests of dimension-specific (e.g., health, income, education) poverty orderings.

[2]For concreteness, we have chosen to frame our discussion in the context of the AK measure. However, our methodology may also be extended, for example, to test hypotheses that arise from the multidimensional orderings of Maasoumi and Lugo (2008) and Duclos et al. (2006). See also Kakwani and Silber (2008) for an overview of these and other approaches.

To illustrate the methodology developed in this article, we use the National Sample Survey (NSS) 60th round health and morbidity data to study the differences in multidimensional poverty among Hindus and Muslims in urban India. Two separate sets of hypotheses are tested. The first corresponds to a robustness check on the second cutoff (k). We find that for lower values of k, Muslims are poorer than Hindus. This is in accordance with income based poverty comparisons which have generally found Muslims to be more deprived. Interestingly, for higher values of k our results suggest that Hindus are, in fact, poorer than Muslims. In other words, a greater proportion of Hindus suffer from *extreme* poverty. To further our understanding of this reversal, we also investigate which of the dimensions may be responsible. Thus, our second set of tests correspond to a simultaneous test of the component orderings for fixed values of k. Our results here suggest that for higher values of k, the difference in the contribution of income to Hindu and Muslim poverty is small (sometimes even insignificant), and that the reversal in the poverty ordering among the two ethnic groups is driven primarily by dimensions other than income. These results, while interesting in and of themselves, serve to highlight the rich empirical welfare analysis that can be conducted by coupling our statistical methodology with a multidimensional approach to poverty.

The remainder of this article is as follows. In the next section, we formulate a generalized version of the recently proposed Alkire–Foster class of multidimensional poverty measures. We subsequently discuss the formulation of a variety of statistical hypotheses and show that they may be treated in a unified manner. Section 4 develops suitable test statistics and the related asymptotics. Section 5 provides a discussion of the implementation of the minimum p-value methodology, which is followed by our empirical illustration in Section 6.

2. FORMULATION

Let $X = (X_1, \ldots, X_d)$ denote a random draw from a population with joint distribution of achievement F. The components of X may be ordinal or cardinal. Without loss of generality, we assume that the first $d_1 \leq d$ components of the random vector X are ordinal whereas the remaining $d - d_1$ are cardinal. For a fixed k, $1 \leq k \leq d$, a prespecified vector of poverty lines $\ell \in (0, \bar{\ell}]^d$, and a $d \times 1$ vector of "weights" denoted by ω, we formulate the multidimensional headcount ratio and *generalized* AF multidimensional poverty measures as

$$H(\ell, k, \omega, F) = \mathbb{E}_F \left[\mathbb{1} \left(\sum_{j=1}^{d} \omega_j \mathbb{1}(X_j \leq \ell_j) \geq k \right) \right], \tag{1}$$

and

$$P_\alpha(\ell, k, \omega, F) = \frac{1}{d}\mathbb{E}_F\left[\mathbb{1}\left(\sum_{j=1}^{d}\omega_j\mathbb{1}(X_j \leq \ell_j) \geq k\right)\left(\sum_{j=1}^{d_1}\omega_j\mathbb{1}(X_j \leq \ell_j)\right)\right]$$
$$+\frac{1}{d}\mathbb{E}_F\left[\mathbb{1}\left(\sum_{j=1}^{d}\omega_j\mathbb{1}(X_j \leq \ell_j) \geq k\right)\right.$$
$$\left.\times\left(\sum_{j=d_1+1}^{d}\omega_j\left(\frac{\ell_j - X_j}{\ell_j}\right)^{\alpha}\mathbb{1}(X_j \leq \ell_j)\right)\right]. \qquad (2)$$

For a given choice of k, ω, and ℓ, we see that under either measure an individual with observed vector of achievement $X = (X_1, \ldots, X_d)$ is identified as poor only if $\sum_{j=1}^{d}\omega_j\mathbb{1}(X_j \leq \ell_j) \geq k$. Identification thus involves a dual cut-off approach. In the first step, deprivation in dimension j is determined by comparing the level of achievement in dimension j to the corresponding poverty line. In the second stage, an individual is identified as being poor only if the weighted (by ω) sum of the indicators of dimension-specific poverty are at least equal to the multidimensional poverty threshold k.

When the dimensions are given equal weight (i.e., when ω equals the unit vector in $\mathbb{R}^d$), $H(\ell, k, \omega, F)$ is simply the proportion of the population that is deprived in k or more dimensions; or equivalently the probability that a randomly drawn person from population F is deprived in k or more dimensions. Alternatively, the measure $P_\alpha(\ell, k, \omega, F)$, for $\alpha > 0$, is a weighted sum of $H(\ell, k, \omega; F)$ where the individual weights correspond to *FGT*-type measures (Foster et al., 1984) of the individual dimensions, thus allowing for the “depth” of deprivation to enter into the overall assessment of poverty. Greater values of α correspond to greater emphasis being placed on the “depth” of deprivation or equivalently greater emphasis being placed on the poorest of the poor. When $\alpha = 0$, $P_\alpha(\ell, k, \omega, F)$ reduces to a weighted sum of $H(\ell, k, \omega, F)$ where the weights are simply the probabilities of being deprived in each of the dimensions under consideration.[3]

Varying ω away from the unit vector amounts to a rescaling of the importance attributed to the various dimensions of poverty. For instance, if $\omega = (2, 0.5, 0.5, 0.5)$ and $k = 3$, then an individual is identified as poor only if they are deprived in the first dimension along with being deprived in at least two other dimensions. Thus, deprivation in dimension one

[3]In some situations, it may be of interest to allow the value of α to be dimension-dependent. Although we have not formulated $P_\alpha(\cdot)$ to explicitly account for this possibility, we note that such an extension can easily be accommodated.

becomes a necessary condition for identification under this weighting scheme. In contrast, we see that under the equally weighted scheme the same individual would be identified as poor only if they are deprived in at least *any* three of the four dimensions. Thus, the choice of ω (and, of course, k) plays a crucial role in the identification of deprived individuals.

In addition to being intuitive and simple to compute, the AK measure also possesses the desirable properties of both subgroup and dimension-specific decomposability. For example, if Z is a discrete random variable with $Z = i$ denoting membership in subgroup i, then we may write the poverty measure as a weighted sum of the subgroup contributions to overall poverty.

The values of $H(\ell, k, \omega, F)$ and $P_\alpha(\ell, k, \omega, F)$ are clearly influenced by the parameters ℓ, ω, α, and k, about whose values there may be considerable disagreement. Consequently, it may be of interest, for example, to test the robustness of an AK poverty ordering of two populations to changes in these parameter values. The formulation of such hypotheses is the subject of the next section.

3. HYPOTHESES

Let G denote the joint distribution of achievement of a population which is to be compared to that of F. Tests of multidimensional poverty orderings will invariably involve hypotheses that are formulated based on the difference between $H(\ell_F, k, \omega, F)$ and $H(\ell_G, k, \omega, G)$, $P_\alpha(\ell_F, k, \omega, F)$ and $P_\alpha(\ell_G, k, \omega, G)$, or the difference between several such population parameters.[4] In this section, we outline the basic structure of the statistical hypotheses that are treated in this article. We begin with a number of specific examples that are likely to be of particular interest to practitioners.

Example 3.1 (Poverty Component Analysis). Due to the composite nature of the measures, inferring, for example, that $P_\alpha(\ell_G, k, \omega, G) > P_\alpha(\ell_F, k, \omega, F)$ invariably leads to the question: “In which dimensions is the population G worse off?” Consequently, it may be of greater interest to consider both the P_α-ordering and the dimension specific orderings via a simultaneous test of the $d + 1$ hypotheses

$$H_0 : P_\alpha(\ell_G, k, \omega, G) - P_\alpha(\ell_F, k, \omega, F) \le 0$$

and

$$H_s : P_{\alpha,s}(\ell_G, k, \omega, G) - P_{\alpha,s}(\ell_F, k, \omega, F) \le 0 \qquad \text{for } 1 \le s \le d,$$

[4]The subscript on the poverty line vector highlights the fact that we allow for the prespecified (exogenous) poverty lines to differ across any two populations.

where the additional subscript "s" on the poverty measure P_α denotes the sth dimension's contribution to the the poverty measure.

Example 3.2 (Robustness). In empirical work researchers often observe the poverty ordering reverse when the value of α or k is adjusted. When this does not occur and the ordering is consistent for all plausible values of α and k, the ordering is said to be robust. Along the lines of the previous example, robustness over (say) α may be tested via a simultaneous test of

$$H_s : P_{\alpha_s}(\ell_G, k, \omega, G) - P_{\alpha_s}(\ell_F, k, \omega, F) \leq 0 \qquad \text{for } 1 \leq s \leq S.$$

Clearly, testing for robustness over k is analogous, with the test being over various values of k as opposed to various values of α.

Example 3.3 (Poverty Orderings Relative to a Benchmark). For a given poverty measure, say $P_\alpha(\cdot)$, an analyst may wish to identify those populations that have less poverty than a benchmark population F_0. Letting $F_1, \ldots, F_S$ denote the various populations that have been chosen for comparison the testing problem can be formulated as a simultaneous test of the S hypotheses

$$H_s : P_\alpha(\ell_{F_s}, k, \omega, F_s) - P_\alpha(\ell_{F_0}, k, \omega, F_0) \leq 0 \quad \text{for } 1 \leq s \leq S.$$

The theme which is common to these (and many other) examples is that the hypotheses of interest may be written in the general form

$$E_P[m(X; \theta)] \leq 0,$$

where m is a vector-valued function, X is a random vector with distribution P, and θ is a vector of (known) parameter values. This observation suggests that in our discussion of statistical testing we may treat these and other seemingly disparate tests in a unified manner; i.e., as simultaneous tests of multiple inequalities.

4. ESTIMATION AND ASYMPTOTICS

Fundamental to our testing procedures is the estimation of the the multidimensional headcount ratio and *generalized* AF poverty measures for various configurations of the exogenous parameters α, ℓ, ω, and k. In this section, we discuss our estimation strategy, and we also establish the joint asymptotic distribution of the resulting estimators. Since the specific estimators of interest and the associated joint distribution will invariably depend upon the particular hypothesis under consideration, the asymptotic analysis here is most aptly handled by treating the empirical poverty measures as a stochastic process in the exogenous parameters and applying techniques from the empirical process literature for their analysis. We therefore begin this section by introducing an empirical process which nests many statistics, including for instance those pertinent to Examples 1 through 3, as special cases. Then, by establishing the weak convergence of this process, the joint asymptotic normality of the statistics of interest may be obtained as simple corollaries.

In our analysis, we treat both the case of mutually dependent samples as well as the case of independent samples, the former being relevant in examining the evolution of poverty of a single group (e.g., changes in poverty over time), whereas the latter is relevant in comparing poverty across any two groups (e.g., cross-national), where sampling is done independently within each group. For the sake of exposition we will assume, without loss of generality, that the number of populations under consideration in any given hypothesis is less than or equal to three. We begin our analysis with the dependent case.

4.1. Dependent Samples

Let $(X_1, Y_1, Z_1), \ldots, (X_n, Y_n, Z_n)$ be i.i.d. copies of a $3d \times 1$ random vector with distribution P and d-dimensional marginal cdfs F, G, and H. We denote by $\mathbb{P}_n$ the empirical measure based on a sample of size n from P, and we introduce the poverty vector functions $m_i : (x, y, z) \in \mathbb{R}^{3d} \to \mathbb{R}^{d+2}$, $i = 1, 2, 3$, which we define by

$$m_1(x, y, z; \ell, k, \omega, \alpha)$$
$$= \begin{pmatrix} \mathbb{1}\left(A(x)\right) \\ \mathbb{1}\left(A(x)\right) \frac{1}{d}\left[\sum_{j=1}^{d_1} \omega_j \mathbb{1}(x_j \leq \ell_j) + \sum_{j=1+d_1}^{d} \omega_j \left(\frac{\ell_j - x_j}{\ell_j}\right)^{\alpha} \mathbb{1}(x_j \leq \ell_j)\right] \\ \mathbb{1}\left(A(x)\right) \omega_1 \mathbb{1}(x_1 \leq \ell_1) \\ \vdots \\ \mathbb{1}\left(A(x)\right) \omega_d \left(\frac{\ell_d - x_d}{\ell_d}\right)^{\alpha} \mathbb{1}(x_d \leq \ell_d) \end{pmatrix}, \tag{3}$$

$$m_2(x, y, z; \ell, k, \omega, \alpha) = m_1(z, x, y; \ell, k, \omega, \alpha),$$

and

$$m_3(x, y, z; \ell, k, \omega, \alpha) = m_1(y, z, x; \ell, k, \omega, \alpha),$$

where $A(x) = \left\{ \sum_{j=1}^{d} \omega_j \mathbb{1}(x_j \leq \ell_j) \geq k \right\}$. Thus, m_i for $i = 2, 3$ is obtained from m_{i-1} through a cyclical permutation of the three $d \times 1$ arguments

x, y, and z. For a fixed choice of parameters $(\ell, k, \omega, \alpha)$ the poverty vectors associated with the F, G, and H distributions are simple population means which may be estimated in a straightforward manner as $\mathbb{P}_n m_1(x, y, z; \ell, k, \omega, \alpha)$, $\mathbb{P}_n m_3(x, y, z; \ell, k, \omega, \alpha)$, and $\mathbb{P}_n m_2(x, y, z; \ell, k, \omega, \alpha)$, respectively.[5]

In each of the examples considered in the previous section, appropriate test statistics of the individual hypotheses may be derived from

$$\sqrt{n}\mathbb{P}_n[m_i(x, y, z; \ell_i, k, \omega, \alpha) - m_j(x, y, z; \ell_j, k, \omega, \alpha)], \tag{4}$$

for some $i, j \in \{1, 2, 3\}$ and some configuration of the parameters (k, ω, α). Consequently, a treatment of the asymptotic behavior of the seemingly disparate cases may be handled in a uniform manner by viewing (4) as a stochastic process in the parameters and applying to it results from the empirical process literature. To this end, we begin by introducing the class of real-valued functions

$$\mathcal{F}_i = \Big\{\langle m_i(x, y, z; \ell, k, \omega, \alpha), h\rangle : \ell \in [0, \bar{\ell}]^d, \\ k \in [\underline{k}, \bar{k}], \sum \omega_i = d, \omega_i \geq 0, h \in [0, 1]^{d+2}\Big\} \tag{5}$$

where i is a fixed integer belonging to the set $\{1, 2, 3\}$ and $\langle \cdot, \cdot \rangle$ denotes the scalar product of two vectors. Our goal is to establish that $\mathcal{F}_i$ is a Donsker class and hence that the empirical process $\{\sqrt{n}(\mathbb{P}_n - P)f : f \in \mathcal{F}_i\}$ converges weakly to a mean-zero Gaussian process in $\ell^\infty(\mathcal{F}_i)$. Establishing this result, which we state formally as Theorem 4.1 below, will enable us to obtain as corollaries a number of convergence results which will prove particularly useful in the development of various statistical tests of interest.

Theorem 4.1. *Suppose* $(X_1, Y_1, Z_1), \ldots, (X_n, Y_n, Z_n)$ *are i.i.d. copies of a* $3d \times 1$ *random vector with distribution* P. *Then, the class of functions* $\mathcal{F}_i$ *defined in (5) is* P*-Donsker for* $i \in \{1, 2, 3\}$.

Theorem 4.1 can be used to derive several important results. First, by defining the class of functions

$$\mathcal{F}_i' = \Big\{ - \langle m_i(y, x, z; \ell, k, \omega, \alpha), h\rangle : z \in [0, \bar{z}]^d, \\ k \in [\underline{k}, \bar{k}], \sum \omega_i = d, \omega_i \geq 0, h \in [0, 1]^{d+2}, 1 \leq \alpha \leq 3\Big\} \tag{6}$$

[5]For a given probability measure P, the notation Pf denotes the expectation with repect to P, i.e., $Pf := \int f dP$.

we obtain via Theorem 4.1 and Donsker preservation under addition (Kosorok, 2008, p. 173) that the empirical process

$$\{\sqrt{n}(\mathbb{P}_n - P)f : f \in \mathscr{F}_1 + \mathscr{F}_2'\}$$

converges weakly to a tight Gaussian process in $\ell^\infty(\mathscr{F}_1 + \mathscr{F}_2')$. Since finite dimensional convergence is necessary for weak convergence of the empirical process, we immediately obtain, for example, the convergence of $\{\sqrt{n}(\mathbb{P}_n - P)(f_1, \ldots, f_S)\}$ to a S-dimensional mean-zero normal distribution provided $f_s \in \mathscr{F}_1 + \mathscr{F}_2'$ for $s = 1, \ldots, S$. The connection to our testing problem is made upon noticing that an element, say f, of $\mathscr{F}_1 + \mathscr{F}_2'$ is of the form

$$f = \langle m_1(x, y, z; \ell_F, k, \omega, \alpha), h\rangle - \langle m_2(x, y, z; \ell_G, k, \omega, \alpha), h'\rangle,$$

and hence, for $h = h' = (1, 0, \ldots, 0)$ or $h = h' = (0, 1, 0, \ldots, 0)$, the scaled and centered random quantity $\sqrt{n}(\mathbb{P}_n - P)f$ is nothing other than the scaled and recentered difference between the estimates of $H(\ell_F, k, \omega, F)$ and $H(\ell_G, k, \omega, G)$, or $P_\alpha(\ell_F, k, \omega, F)$ and $P_\alpha(\ell_G, k, \omega, G)$, respectively.

Notice that Example 3.3 is a slight variation on the above themes in that it involves a comparison between several populations. In order to subsume Example 3.3, we introduce the class of functions $\mathscr{G}_j = \mathscr{F}_1 + \mathscr{F}_j'$ and denote by $\mathscr{H}$ the class of functions

$$\{\langle f, \lambda\rangle : f \in \mathscr{G}_2 \times \mathscr{G}_3, \lambda \in [-1, 1]^2\}.$$

which is also P-Donsker under the conditions of Theorem 4.1. The application of these results to our testing problems are now made explicit by revisiting our earlier examples:

Example 4.1 (Example 3.1 Continued). Let h_i denote the ith standard basis vector in $\mathbb{R}^{d+2}$, $\omega \in \mathbb{R}^d_+$ satisfy $\sum \omega_i = d$, $\ell_G, \ell_F \in (0, \bar{\ell}]^d$, and α be a fixed positive integer. Then each member of the finite collection

$$\{\langle m_1(x, y, z; \ell_F, k, \omega, \alpha), h_i\rangle - \langle m_2(x, y, z; \ell_G, k, \omega, \alpha), h_i\rangle : 2 \le i \le d + 2\} \quad (7)$$

belongs to $\mathscr{G}_2$. We therefore obtain the convergence of

$$\sqrt{n}(\mathbb{P}_n - P)\begin{pmatrix} m_{1,2}(x, y, z; \ell_G, k, \omega, \alpha) - m_{2,2}(x, y, z; \ell_F, k, \omega, \alpha) \\ m_{1,2}(x, y, z; \ell_G, k, \omega, \alpha) - m_{2,3}(x, y, z; \ell_F, k, \omega, \alpha) \\ \vdots \\ m_{1,d+2}(x, y, z; \ell_G, k, \omega, \alpha) - m_{2,d+2}(x, y, z; \ell_F, k, \omega, \alpha) \end{pmatrix}$$

to a mean-zero multivariate normal distribution.

Example 4.2 (Example 3.2 Continued). Let $h = (0, 1, 0, \ldots, 0) \in \mathbb{R}^{d+2}$, $\omega \in \mathbb{R}_+^d$ satisfy $\sum \omega_i = d$, $\ell_G, \ell_F \in (0, \bar{\ell}]^d$, and $\alpha(i) = i$ for $i = 1, 2, 3$. Then each member of the finite collection $\{\langle m_2(x, y, z; \ell_G, k, \omega, \alpha(i)), h\rangle - \langle m_1(x, y, z; \ell_F, k, \omega, \alpha(i)), h\rangle : 1 \leq i \leq 3\}$ belongs to $\mathscr{G}_2$. We therefore obtain the convergence of

$$\sqrt{n}(\mathbb{P}_n - P)\begin{pmatrix} m_{1,2}(x, y, z; \ell_F, k, \omega, 1) - m_{2,2}(x, y, z; \ell_G, k, \omega, 1) \\ m_{1,2}(x, y, z; \ell_F, k, \omega, 2) - m_{2,2}(x, y, z; \ell_G, k, \omega, 2) \\ m_{1,2}(x, y, z; \ell_F, k, \omega, 3) - m_{2,2}(x, y, z; \ell_G, k, \omega, 3) \end{pmatrix}$$

to a mean-zero multivariate normal distribution.

Example 4.3 (Example 3.3 continued). Let $h = (0, 1, 0, \ldots, 0) \in \mathbb{R}^{d+2}$, $\omega \in \mathbb{R}_+^d$ be a fixed vector satisfying $\sum \omega_i = d$, $\ell_G, \ell_F, \ell_H \in (0, \bar{\ell}]^d$, and α be a fixed positive integer. Then each member of the finite collection $\{\langle f, \lambda\rangle : f \in \mathscr{G}_2 \times \mathscr{G}_3, \lambda \in \{(1, 0), (0, 1)\}\}$ belongs to $\mathscr{H}$. We therefore obtain the convergence of

$$\sqrt{n}(\mathbb{P}_n - P)\begin{pmatrix} m_{1,2}(x, y, z; \ell_F, k, \omega, \alpha) - m_{3,2}(x, y, z; \ell_H, k, \omega, \alpha) \\ m_{1,2}(x, y, z; \ell_F, k, \omega, \alpha) - m_{2,2}(x, y, z; \ell_G, k, \omega, \alpha) \end{pmatrix}$$

to a mean-zero bivariate normal distribution.

4.2. Independent Samples

We now specialize the above results to the case where $\mathscr{X} = (X_1, \ldots, X_{n_1})$, $\mathscr{Y} = (Y_1, \ldots, Y_{n_2})$, and $\mathscr{Z} = (Z_1, \ldots, Z_{n_3})$ are independent random samples with respective distributions P_X, P_Y, and P_Z. To this end, let $\mathscr{F}$ denote the class of functions

$$\{\langle m(x; \ell, k, \omega, \alpha), h\rangle : \ell \in [0, \bar{\ell}]^d, k \in [\underline{k}, \bar{k}], \sum \omega_i = d, \omega_i \geq 0, h \in [0, 1]^{d+2}\},$$

where $m : \mathbb{R}^d \to \mathbb{R}^{d+2}$. Further, denote by $\mathbb{G}_{n_1, P_X}$ the signed measure $\sqrt{n_1}(\mathbb{P}_{n_1, X} - P_X)$ with analogous definitions for $\mathbb{G}_{n_2, P_Y}$ and $\mathbb{G}_{n_3, P_Z}$. For analyzing cases such as those presented in Examples 3.1 and 3.2 our interest centers on the asymptotic behavior of an empirical process of the form

$$\left\{\left(\frac{n_1 n_2}{n_1 + n_2}\right)^{1/2}\left[n_1^{-1/2}\mathbb{G}_{n_1, P_X} f_1 - n_2^{-1/2}\mathbb{G}_{n_2, P_Y} f_2\right] : (f_1, f_2) \in \mathscr{F} \times \mathscr{F}\right\}. \qquad (8)$$

In order to establish the asymptotic behavior of the empirical process in (8) we will require the following assumption:

Assumption 4.1 (Sampling Process). $\inf_{i \neq j}\{n_i/n_j\} \to (0, 1)$ as $n \to \infty$.

From the independence assumption together with Assumption 4.1, we obtain the following.

Theorem 4.2. *Suppose Assumption* 4.1 *holds, then the empirical process in (8) converges to the limit process*

$$\left\{\lambda_1^{1/2}\mathbb{G}_{P_X}f_1 - (1-\lambda_1)^{1/2}\mathbb{G}_{P_Y}f_2 : (f_1, f_2) \in \mathcal{F} \times \mathcal{F}\right\}$$

for some $\lambda \in (0,1)$, where $\{\mathbb{G}_{P_X}f : f \in \mathcal{F}\}$ and $\{\mathbb{G}_{P_Y}f : f \in \mathcal{F}\}$ are independent zero-mean Gaussian processes.

The applications to our Examples 1 and 2 are immediate.

Example 4.4 (Example 3.1 Continued). Let h_i denote the ith standard basis vector in $\mathbb{R}^{d+2}$, $\omega \in \mathbb{R}^d_+$ satisfy $\sum \omega_i = d$, $\ell_G, \ell_F \in (0, \bar{\ell}]^d$, and α be a fixed positive integer. Then each member of the finite collection $\{(\langle m(x;\ell_F,k,\omega,\alpha),h_i\rangle, \langle m(x;\ell_G,k,\omega,\alpha),h_i\rangle) : 2 \le i \le d+2\}$ belongs to $\mathcal{F} \times \mathcal{F}$. We therefore obtain form Theorem 4.2 the convergence of

$$\left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\left[\mathbb{G}_{n_2,P_Y}\begin{pmatrix} m_2(x;\ell_G,k,\omega,\alpha) \\ m_3(x;\ell_G,k,\omega,\alpha) \\ \vdots \\ m_{(d+2)}(x;\ell_G,k,\omega,\alpha)\end{pmatrix} - \mathbb{G}_{n_1,P_X}\begin{pmatrix} m_2(x;\ell_F,k,\omega,\alpha) \\ m_3(x;\ell_F,k,\omega,\alpha) \\ \vdots \\ m_{d+2}(x;\ell_F,k,\omega,\alpha)\end{pmatrix}\right]$$

to a zero-mean multivariate normal distribution.

Example 4.5 (Example 3.2 Continued). Let $h = (0,1,0,\ldots,0) \in \mathbb{R}^{d+2}$, $\omega \in \mathbb{R}^d_+$ satisfy $\sum \omega_i = d$, $\ell_G, \ell_F \in (0,\bar{\ell}]^d$, and $\alpha(i) = i$ for $i = 1,2,3$. Then each member of the finite collection $\{(\langle m(x;\ell_F,k,\omega,\alpha(i)),h\rangle, \langle m(x;\ell_G,k,\omega,\alpha(i)),h\rangle) : 1 \le i \le 3\}$ belongs to $\mathcal{F} \times \mathcal{F}$. We therefore obtain the convergence of

$$\left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\left[\mathbb{G}_{n_2,P_Y}\begin{pmatrix} m_2(x;\ell_G,k,\omega,1) \\ m_2(x;\ell_G,k,\omega,2) \\ m_2(x;\ell_G,k,\omega,3)\end{pmatrix} - \mathbb{G}_{n_1,P_X}\begin{pmatrix} m_2(x;\ell_F,k,\omega,1) \\ m_2(x;\ell_F,k,\omega,2) \\ m_2(x;\ell_F,k,\omega,3)\end{pmatrix}\right]$$

to a zero-mean multivariate normal distribution.

Again, as in the dependent case, testing problems such as those encountered in Example 3.3 require a slight modification; namely, consider the process

$$\Big\{\eta^{1/2}\Big[n_1^{-1/2}\mathbb{G}_{n_1,P_X}f_1 - n_2^{-1/2}\mathbb{G}_{n_2,P_Y}f_2 + (n_1^{-1/2}\mathbb{G}_{n_1,P_X}f_3 - n_3^{-1/2}\mathbb{G}_{n_3,P_Z}f_4)\Big] : \\ (f_1,f_2,f_3,f_4) \in \mathcal{F}^4\Big\}, \tag{9}$$

where $\eta = \left(\frac{n_1 n_2 n_3}{n_1 n_2 + n_1 n_3 + n_2 n_3}\right)$. In order to establish the asymptotic behavior of the empirical process in (9), we require the following assumption.

Assumption 4.2 (Sampling Process). $\inf_{(i,j)\neq(k,l)}\{(n_i n_j)/(n_k n_l)\} \to (0,1)$ as $n \to \infty$ whenever $i \neq j$ and $k \neq l$.

From the independence assumption together with Assumption 4.2, we are able to establish the following result.

Theorem 4.3. *Suppose Assumption* 4.2 *holds, then (9) converges to the limit process*

$$\Big\{\lambda_1^{1/2}\mathbb{G}_{P_X}f_1 - \lambda_2^{1/2}\mathbb{G}_{P_Y}f_2 + \lambda_1^{1/2}\mathbb{G}_{P_X}f_3 - (1-\lambda_3)^{1/2}\mathbb{G}_{P_Z}f_4 : (f_1,f_2,f_3,f_4) \in \mathcal{F}^4\Big\}$$

for some $\lambda_1, \lambda_2, \lambda_3 \in (0,1)$ *with* $\sum \lambda_i = 1$, *where* $\{\mathbb{G}_{P_X}f : f \in \mathcal{F}\}$, $\{\mathbb{G}_{P_Y}f : f \in \mathcal{F}\}$, *and* $\{\mathbb{G}_{P_Z}f : f \in \mathcal{F}\}$ *are independent zero-mean Gaussian processes.*

We are now in a position to obtain the convergence result relevant to Example 3.3.

Example 4.6 (Example 3.3 continued). Let $h = (0,1,0,\ldots,0) \in \mathbb{R}^{d+2}$, $\omega \in \mathbb{R}_+^d$ be a fixed vector satisfying $\sum \omega_i = d$, $\ell_G, \ell_F, \ell_H \in (0,\bar{\ell}]^d$, and α be a fixed positive integer. Then each member of the finite collection $\{\langle f,\lambda\rangle : f \in \mathcal{G}_{(2)} \times \mathcal{G}_{(3)}, \lambda \in \{(1,0),(0,1)\}\}$ belongs to $\mathcal{H}$. We therefore obtain the convergence of

$$\sqrt{\eta}(\mathbb{P}_{n_1,n_2,n_3} - P)\begin{pmatrix} m_{3,2}(x,y,z;\ell_H,k,\omega,\alpha) - m_{1,2}(x,y,z;\ell_F,k,\omega,\alpha) \\ m_{2,2}(x,y,z;\ell_G,k,\omega,\alpha) - m_{1,2}(x,y,z;\ell_F,k,\omega,\alpha) \end{pmatrix}$$

to a mean-zero bivariate normal distribution as an immediate consequence of Theorem 4.3.

5. TESTING METHODOLOGY

For a given collection $(f_1,\ldots,f_S)$ with f_s, $1 \leq s \leq S$, a member of the P-Donsker class $\mathcal{F}$ (c.f. Examples 4.1 through 4.6), our interest centers on a simultaneous test of the hypotheses

$$H_s : Pf_s \leq 0 \qquad \text{against} \qquad H_s' : Pf_s > 0 \quad 1 \leq s \leq S.$$

It is well known that the classical Wald-type tests of Wolak (1989) and Kodde and Palm (1986), for example, can be applied here to test the joint intersection hypothesis

$$H_0 : Pf_s \leq 0 \quad \text{for all } 1 \leq s \leq S \qquad \text{against}$$
$$H_A : Pf_s > 0 \quad \text{for some } 1 \leq s \leq S.$$

Unfortunately, a rejection of H_0 based on the Wald-type test does not necessarily imply that H_s is rejected for some $1 \leq s \leq S$; indeed, we may reject the joint intersection hypothesis H_0 without finding compelling evidence against any individual hypothesis H_s. Thus, in the context of Example 3.1, for instance, policy makers who adopt a Wald-type procedure may infer that a country or region is underachieving and yet be unable to infer the specific dimensions (e.g., income, health, education, etc.) which are responsible for the finding. Clearly, this is undesirable if policy makers wish to obtain compelling evidence regarding dimension-specific underachievement and design targeted efforts accordingly.

In contrast to Wald-type tests, minimum p-value (MinP) tests are designed specifically to allow one to identify the source(s) of rejection when rejection occurs.[6] In order to provide some background on the MinP methodology, we begin first by describing a suitable procedure for the computation of bootstrap p-values. Towards this end, it is well known (Kosorok, 2008, p. 20) that the Donsker property of $\mathscr{F}$ implies not only that

$$\sqrt{n}(\mathbb{P}_n - P)(f_1, \ldots, f_S) \Rightarrow N_S(0, \Omega(P)), \tag{10}$$

but also that

$$\sqrt{n}(\widehat{\mathbb{P}}_n - \mathbb{P}_n)(f_1, \ldots, f_S) \Rightarrow N_S(0, \Omega(P)), \tag{11}$$

in probability, where $\widehat{\mathbb{P}}_n$ denotes the bootstrap empirical measure and $N_S(0, \Omega(P))$ denotes an S-dimensional normal distribution with covariance matrix $\Omega(P)$ (the notation here reflects the dependence of Ω on the underlying probability mechanism P). Letting $J_n(\cdot, \mathbb{P}_n)$ denote the bootstrap approximation (c.f. Eq. (11)) to the sampling distribution in (10) and denoting by $J_{n,s}(\cdot, \mathbb{P}_n)$ the sth marginal distribution, it is straightforward that the bootstrap p-values associated with each of the component statistics may be obtained from

$$\hat{p}_s = 1 - J_{n,s}(\sqrt{n}\mathbb{P}_n f_s, \mathbb{P}_n). \tag{12}$$

[6]See, for example, Westfall and Young (1993) on the use of MinP tests. Also, see Romano and Wolf (2005) and Hso et al. (2010) for a discussion of related procedures.

The bootstrap p-value $\hat{p}_s$ in (12) provides a measure of the strength of evidence against H_s, and it is tempting to reject H_s at the nominal level α if $\hat{p}_s < \alpha$. This testing strategy, however, ignores the multiplicity of the hypotheses under test and will tend to reject true hypotheses too often in the sense that

$$Prob_P\left\{\text{Reject at least one } H_s, s \in I(P)\right\} > \alpha \tag{13}$$

whenever the collection of true hypotheses $I(P)$ contains two or more elements. For instance, if $S = 5$, $Pf_s = 0$ for every s (all H_s are true), and all tests are mutually independent, then, at the 5% level of significance

$$\begin{aligned} Prob_P\left\{\text{Reject at least one } H_s, s \in I(P)\right\} &= Prob_P\left\{\min_{1\leq s\leq S} \hat{p}_s < 0.05\right\} \\ &\xrightarrow{n\to\infty} Prob_P\left\{\min_{1\leq s\leq S} U_s < 0.05\right\} \\ &= 1-(1-0.05)^5 = 0.226, \end{aligned} \tag{14}$$

where we have used the fact that the estimated p-values converge to mutually independent uniform random variates under the assumed conditions. If the number of hypotheses S is increased to 10 or the significance level of the test is increased to 10%, the corresponding error rates jump to 0.401 and 0.409, respectively.[7]

The essence behind the classical MinP procedure lies in appropriately adjusting the standard p-values so as to ensure, at least asymptotically, that

$$Prob_P\left\{\text{Reject at least one } H_s, s \in I(P)\right\} \leq \alpha \tag{15}$$

With the bootstrap distribution $J_n(\cdot, \mathbb{P})$ already in hand, obtaining adjusted p-values satisfying (15) is rather straightforward. Indeed, for a random draw Y from the known distribution $J_n(\cdot, \mathbb{P}_n)$ we may compute

$$\hat{p}_{\min} = \min_{1\leq s\leq S}[1 - J_{n,s}(Y_s, \mathbb{P}_n)]. \tag{16}$$

The corresponding empirical distribution from B such draws, which we denote by $Q_n(\cdot, \mathbb{P}_n)$, constitutes an approximation to the distribution of the

[7]The assumption of mutual independence is made here for illustrative purposes. In practice, we can generally expect some degree of dependence among the hypotheses under test; however, it is only in the case of perfect dependence that we can be guaranteed of appropriate error rate control if we adopt the strategy of independently testing several hypotheses on the basis of individual (unadjusted) p-values.

minimum p-values and hence may be used to obtain the MinP adjusted p-values

$$\hat{p}_s^{adj} = Q_n(\hat{p}_s, \mathbb{P}_n). \tag{17}$$

In contrast to the liberal procedure in which the individual hypotheses are rejected if their unadjusted p-values fall below the nominal level α, it may be shown (Bennett, 2010) that testing the individual hypotheses based on the modified decision rule

$$\text{Reject } H_s \text{ if } \hat{p}_s^{adj} < \alpha$$

guarantees control of the error rate in (15), at least asymptotically. Bennett (2010) also demonstrates that the ability of the MinP test to identify false hypotheses can be greatly enhanced by replacing the random draw $Y \sim J_n(\cdot, \mathbb{P}_n)$ which is subsequently evaluated in (16) with a random draw from the bootstrap distribution $J_n^{PC}(\cdot, \mathbb{P}_n)$, which is defined according to

$$\sqrt{n}(\widehat{\mathbb{P}}_n - \mathbb{P}_n)(f_1, \ldots, f_S) - \sqrt{n}(|\mathbb{P}_n f_1 \mathbb{1}_{\{\mathbb{P}_n f_1 > \delta_{1,n}\}}, \ldots, \mathbb{P}_n f_S \mathbb{1}_{\{\mathbb{P}_n f_S > \delta_{S,,n}\}}), \tag{18}$$

where the $S \times 1$ vector δ_n is selected by the practitioner in accordance with Assumption 5.1 below.

Assumption 5.1. i. $\|\delta_n\| = o_P(1)$;[8] ii. $\operatorname{plim}_{n\to\infty} \inf_{1\le s\le S} n^{1/2}\delta_{n,s} \to \infty$.

Remark 5.1. An example of a sequence δ_n satisfying the conditions of Assumption 5.1 above is given by

$$\delta_{n,s} = \sqrt{\frac{2\hat{\sigma}_{n,s}^2 \log\log n}{n}},$$

where $\hat{\sigma}_{n,s}^2$ denotes a consistent estimator of the asymptotic variance of $\sqrt{n}\mathbb{P}_n f_s$.

To gain some intuition for the mechanics of this procedure first consider the case where all of the hypotheses are on the boundary, i.e., $Pf_s = 0$ for every s. In this case $J_n^{PC}(\cdot, \mathbb{P}_n)$ and $J_n(\cdot, \mathbb{P}_n)$ both converge to $N_S(0, \Omega(P))$, and consequently, $[1 - J_{n,s}(Y_s, \mathbb{P}_n)]$ where $Y \sim J_n^{PC}(\cdot, \mathbb{P}_n)$ converges to a uniform random variable for every $s \in \{1, \ldots, S\}$. Thus, asymptotically, the minimum is over an $S \times 1$ vector random variable with uniform (univariate) marginals, as should be expected when all of the $Pf_s = 0$. In contrast, when $Pf_s > 0$ the sth marginal distribution

[8] $\|\cdot\|$ denotes the standard Euclidean norm.

$J_{n,s}^{PC}(\cdot, \mathbb{P}_n)$ converges in probability to a degenerate distribution at $-\infty$ (the term $\sqrt{n}(\mathbb{P}_n f_s \mathbb{1}_{\{\mathbb{P}_n f_s > \delta_{n,s}\}}$ in (18) tends to ∞ with probability tending to 1 provided δ_n is chosen in accordance with Assumption 5.1), in which case

$$[1 - J_{n,s}(Y_s, \mathbb{P}_n)] \to 1$$

in probability as $n \to \infty$, and the index set over which the minimum is computed is effectively reduced. Since the minimum p-value is generally decreasing in the number of indices over which the minimum is computed, the elimination of any index for which $Pf_s > 0$ generally reduces the adjusted p-values and ultimately enhances the test's ability to detect false hypotheses while still allowing us to maintain appropriate control over the error rate (c.f. Eq. (15)).

The implementation of the MinP testing procedure as described above in the specific context of Example 3.1 and the case of dependent samples is conveniently summarized in Algorithm 1 below.

Algorithm 5.1 (Example 1 Cont'd: The Dependent Case).

1. *Draw a random sample of size n, i.e.,* $\{(X_1^*, Y_1^*, Z_1^*), \ldots, (X_n^*, Y_n^*, Z_n^*)\}$, *from* $\{(X_1, Y_1, Z_1), \ldots, (X_n, Y_n, Z_n)\}$ *and compute the difference*

$$n^{-1} \sum_{i=1}^{n} [m_{(1)}(X_i^*, Y_i^*, Z_i^* : \ell, k, \omega, \alpha) - m_{(2)}(X_i^*, Y_i^*, Z_i^*; \ell, k, \omega, \alpha)] \qquad (19)$$

2. *Repeat Step 1 B times and compute the empirical bootstrap distribution* $J_n(\cdot, \mathbb{P}_n)$ *and the* $B \times S$ *matrix of partially recentered bootstrap statistics using equation* (18).
3. *Compute the p-values of the S original and* $B \times S$ *partially recentered bootstrap statistics by evaluating them in the appropriate marginal distributions* $J_{n,s}(\cdot, \mathbb{P}_n)$ *of* $J_n(\cdot, \mathbb{P}_n)$.
4. *Compute the empirical distribution of row-minimums from the* $B \times S$ *matrix of p-values obtained in Step* 3.
5. *Compute the adjusted p-values corresponding to each test by evaluating the p-values of the S original statistics (obtained in Step* 2*) in the empirical distribution obtained via Step* 4.

Aside from substituting for the appropriate statistics (i.e., in Eq. (19) of Step 1), the algorithms for Examples 3.2 and 3.3 are identical, and are thus omitted. Similarly, the modifications necessary for treating the case of independent samples are also straightforward, and we omit the details of the respective bootstrap algorithms.

6. EMPIRICAL ILLUSTRATION

In this section, we apply our proposed testing methodology to data from India's NSS. We are particularly interested in examining the relative state of poverty across two ethnic groups, namely, Hindus and Muslims. India has a predominantly Hindu population however it has a sizeable proportion of Muslims as well. Traditional income poverty analysis has shown that a lesser proportion of Hindus are poor than the corresponding numbers for Muslims. However, it is of interest to examine whether these findings persist when relevant dimensions or indicators of poverty other than income (or consumption) are included in the analysis.

Our data source is the NSS's 60th round health and morbidity survey. This survey was conducted in the last six months of 2004. For the purposes of this illustration, we restrict attention to urban poverty, for which there are 26,566 households included. Since we are looking only at Hindu and Muslim poverty, all other households are dropped. In India, these two religious groups together account for more than 95% of the total population, and so the resulting sample of 20,243 Hindu households and 3,715 Muslim households consists of the majority of all urban households.

While the NSS is a multistage stratified random sample, for the purpose of this illustration we ignore the complications introduced by this particular sampling design and instead assume the observations to be generated through the process of simple random sampling. While ignoring the specific sampling design is likely to bias our findings,[9] a thorough consideration of the sampling design issue (e.g., providing a detailed discussion of the NSS sampling design, modifying the bootstrap accordingly, etc.) is beyond the scope of the current article.

As for the dimensions of deprivation used in our analysis, we include the following: per capita monthly expenditure (PCME), level of educational attainment, source of drinking water, type of housing structure, type of sanitation, drainage facilities available, and main cooking medium. Since we are measuring household poverty and not individual level poverty, we take for the education level the highest level of education earned by any member of the household. Except per capita expenditure and education, all variables used in our analysis are ordinal. We also implicitly treat all households equally in terms of size since the NSS weighs all households in a village/block equally and therefore does not explicitly account for household size.[10]

The dimensions are chosen to represent the standard of living and the capabilities of the households to improve their position. A notable

[9]Bhattacharya (2007), for example, discusses in detail the effect of ignoring the sampling design of the NSS on inequality measurement.

[10]As pointed out by one of the referees, this is likely to bias the results since households typically have different sizes and household size is likely correlated with both poverty and religion.

TABLE 1 Correlation between the dimensions

	Income	Housing	Sanitation	Drainage	Water	Cooking medium
Income	1					
Housing	0.2042	1				
Sanitation	0.1684	0.3012	1			
Drainage	0.3129	0.3841	0.2908	1		
Water	0.1160	0.1218	0.0187	0.2243	1	
Cooking medium	0.3031	0.3754	0.3339	0.4270	0.1831	1
Education	0.3786	0.3063	0.2777	0.3330	0.0839	0.4748

TABLE 2 Dimension specific poverty lines

Dimension	Poverty line
Income	As given by the Planning Commission
Education	Having not obtained at least a primary education
Sanitation	No sanitation facility available
Drainage	No accessible drainage system
Housing	Person does not reside in a pucca* structure
Source of drinking water	Person used a river, canal, pond, or well
Primary cooking medium	Person had no cooking arrangement or used firewood or dung cakes

*Pucca refers to brick and mortar structures.

omission is health. Unfortunately, reliable sources of data for health of individuals and households are not easily available for India. One source for data on health for India is the National Family Health Survey, however this survey does not ask about income or per capita expenditure. Researchers have used this data after computing an asset index. However, for the purpose of this analysis, we have chosen to use more standard measures of income at the cost of omitting the dimension of health.

Of the seven dimensions used, one might be concerned about a high degree of correlation and the inclusion of "redundant" dimensions. Surprisingly, we find correlations between the various dimensions to be rather low. Indeed, as can be seen in Table 1, no correlation coefficient exceeds 0.5.[11] Therefore, by incorporating all of these dimensions we are able to capture different forms of deprivation in urban India.

Table 2 gives the dimension specific poverty lines used. For PCME, we use the poverty line as established by the Planning Commission of India. The remaining cut-offs are chosen as to describe a minimum standard of living.

[11]We thank an anonymous referee for suggesting that we investigate the correlation among dimensions.

TABLE 3 Incidence of deprivation expressed as a percentage

Dimension	Incidence (Hindu %)	Incidence (Muslim %)
Income	18.9	30.8
Housing	17.9	19.5
Sanitation	19.7	15.8
Drainage	16.8	15.3
Water	5.3	8.8
Cooking medium	26.2	35.9
Education	10.0	17.1

Table 3 summarizes the incidence of deprivation in each of the seven dimensions for Hindus and Muslims, respectively. We note that in every dimension, except sanitation and drainage, the incidence of poverty among Muslims is greater than among Hindus. Further, we note that the largest disparities appear to be in the dimensions of income, main cooking medium, and education level.

In addition to the dimension-specific poverty lines, the P_0 and H measures (or more generally the AF methodogy) require us to set a second cut-off.[12] The second cutoff is the dimension cutoff k which in our analysis can take any value between 1 and 7. The value of k may be set before the analysis is undertaken by governments or by the investigator given the objectives of the exercise. Once k is fixed we may compute the associated level of poverty. When k equals 5, for instance, we see that Hindus are poorer than Muslims under both the H P_0 measures of multidimensional poverty (see Tables 4 and 5, respectively). This conclusion depends on both the dimension specific poverty lines (which we assume here to be exogenously determined) and the value of k (which may be set by the investigator). So a natural robustness check would entail checking the levels of poverty for various values for k. For example if we see that Hindus remain poorer than Muslims for k values ranging from, say, 3 through 6, then we may infer that the poverty ordering is robust to the choice of k. This robustness check corresponds to a multiple inequality test where the null hypotheses are given by

$$H_k : H(\ell_F, k, \omega, F) \leq H(\ell_G, k, \omega, G), \qquad k = 3, 4, 5, 6,$$

and

$$H_k : P_\alpha(\ell_F, k, \omega, F) \leq P_\alpha(\ell_G, k, \omega, G), \qquad k = 3, 4, 5, 6.$$

[12]For the sake of brevity we consider only the P_0 measure and multidimensional headcount H.

TABLE 4 Level of poverty: Multidimensional headcount

	$K = 3$	$K = 4$	$K = 5$	$K = 6$
H for Hindus	0.187	0.113	0.054	0.016
H for Muslims	0.226	0.109	0.044	0.015
Adjusted p-values (Null: M-H ≥ 0)	1.000	0.902	0.061	**0.016**
Adjusted p-values (Null: H-M ≥ 0)	**0.002**	0.302	0.987	0.998

Boldface values are statistically significant at the 5% nominal level.

TABLE 5 Level of poverty: Multidimensional poverty

	$K = 3$	$K = 4$	$K = 5$	$K = 6$
M0 for Hindus	0.107	0.075	0.041	0.014
M0 for Muslims	0.121	0.071	0.034	0.013
Adjusted p-values (Null: M-H ≥ 0)	1.000	0.709	**0.036**	**0.008**
Adjusted p-values (Null: H-M ≥ 0)	**0.000**	0.508	0.991	0.998

Boldface values are statistically significant at the 5% nominal level.

The p-values from these tests, which are presented in Tables 4 and 5, in fact suggest a reversal in the levels of poverty for Hindus and Muslims as k is varied. When k equals 3, for example, we are able to infer that poverty among Muslims is higher than poverty among Hindus. However, for the higher k values of 5 and 6, we reach the opposite conclusion. At k equals 4 there is no significant difference between the levels of poverty for the two groups. It is important to emphasize that the reported p-values are adjusted for multiplicity and thus permit us to draw valid inferences concerning the individual hypotheses under test. Consulting unadjusted p-values, on the other hand, would not protect against the multiplicity problem and generally lead one to find "too" false positives.

The observed reversal in the poverty ordering raises questions about a plausible explanation. Perhaps this reversal is the result of the fact that Hindus can be divided further on the basis of the caste to which they belong. Traditionally, the lower castes have been found to be more deprived, for instance, being made to do menial labor for low wages, and at the expense of receiving education. Even in modern times these castes have lagged behind the rest of the population and constitute some of the poorest individuals in the society. We therefore offer the following plausible explanation for the observed reversal: at higher levels of k, we are primarily capturing the lower castes within the Hindu population. Perhaps what we are observing then is low caste Hindus facing greater hardships, on average, than the Indian Muslim population.[13] Another

[13]Regional and religious disaggregation of poverty in India is explored in greater depth in Mitra (2011).

plausible explanation for the reversal is that for lower values of k, it may be the case that income contributes relatively more to multidimensional poverty than it does for higher values of k. In such a case, we will see that for lower k, we have Hindus less poor, simply because they are less poor by any measure of income poverty. But as k increases the other dimensions become increasingly important in which case, we may see a reversal. A test of this second conjecture is pursued next.

For a given value of k, for example k equals 3, we decompose the P_0 measure into the contributions of each of the dimensions. We then test whether there is significant difference between the contribution of each dimension to Hindu and Muslim poverty. More precisely, we perform a simultaneous test of the $d + 1$ hypotheses

$$H_0 : P_0(\ell_G, k, \omega, G) - P_0(\ell_F, k, \omega, F) \leq 0$$

and

$$H_s : P_{0,s}(\ell_G, k, \omega, G) - P_{0,s}(\ell_F, k, \omega, F) \leq 0 \qquad \text{for } 1 \leq s \leq d,$$

where the additional subscript "s" on the poverty measure P_α $(\alpha = 0)$ denotes the sth dimension's contribution to the the poverty measure.

The results of the above test for $k \in \{3, 4, 5, 6\}$ are presented in Table 6. We have observed that poverty is higher among Muslims at $k = 3$. We now see from the decomposition that incidence in income, housing, water, cooking medium, and education are all lower for Hindus than for Muslims with k equal to 3. For k equals 4, there is no significant difference between Hindus and Muslim poverty and we also see that most of the dimensions do not have significantly different contributions among Hindus and Muslims. For k equal to 5, we find that Muslims are less poor than Hindus, *and* that there is no significant difference in the contribution of income to Hindu and Muslim poverty levels. The difference in overall poverty can be explained only by differences in the levels of deprivation in the other dimensions, namely housing, sanitation and drainage for which we may infer that there is more deprivation among Hindus than among Muslims. For k equal to 6 we find stronger evidence of higher poverty among Hindus than among Muslims. We find at this level of k we have that Hindu households are significantly more deprived in *all* dimensions.

In summary, we find that as k increases beyond 4, income is no longer enough to differentiate between Hindu and Muslim poverty, and that only by including other dimensions are we able to distinguish between Hindu and Muslim households in extreme poverty. This is an interesting finding which lends empirical support to arguments advocating the use of a multidimensional approach to poverty analysis.

TABLE 6 Dimensional decomposition of the multidimensional poverty estimates

	M_0	Income	Housing	Sanitation	Drainage	Water	Cooking	Education
$M_0(k=3)$ Hindu	0.107	0.110	0.118	0.141	0.120	0.027	0.159	0.069
$M_0(k=3)$ Muslim	0.121	0.139	0.136	0.113	0.117	0.047	0.207	0.089
Adj. p-value ($Null: M_0^M - M_0^H \geq 0$)	1.000	1.000	0.992	**0.000**	0.535	1.000	1.000	1.000
Adj. p-value ($Null: M_0^H - M_0^M \geq 0$)	**0.000**	**0.000**	**0.008**	1.000	0.462	**0.000**	**0.000**	**0.000**
$M_0(k=4)$ Hindu	0.075	0.077	0.086	0.098	0.087	0.018	0.103	0.053
$M_0(k=4)$ Muslim	0.071	0.083	0.086	0.077	0.072	0.017	0.101	0.059
Adj. p-value ($Null: M_0^M - M_0^H \geq 0$)	0.906	1.000	0.952	**0.009**	0.081	0.973	0.973	1.000
Adj. p-value ($Null: M_0^H - M_0^M \geq 0$)	0.712	**0.028**	0.637	1.000	1.000	0.548	0.537	0.076
$M_0(k=5)$ Hindu	0.041	0.043	0.048	0.052	0.046	0.010	0.052	0.034
$M_0(k=5)$ Muslim	0.034	0.042	0.038	0.040	0.036	0.007	0.043	0.029
Adj. p-value ($Null: M_0^M - M_0^H \geq 0$)	**0.027**	0.621	**0.029**	**0.007**	**0.015**	0.292	0.123	0.285
Adj. p-value ($Null: M_0^H - M_0^M \geq 0$)	1.000	0.876	1.000	1.000	1.000	0.987	0.998	0.994
$M_0(k=6)$ Hindu	0.014	0.015	0.017	0.016	0.016	0.006	0.016	0.014
$M_0(k=6)$ Muslim	0.013	0.015	0.015	0.014	0.015	0.004	0.015	0.012
Adj. p-value ($Null: M_0^M - M_0^H \geq 0$)	**0.000**	**0.019**	**0.008**	**0.001**	**0.003**	**0.006**	**0.017**	**0.001**
Adj. p-value ($Null: M_0^H - M_0^M \geq 0$)	1.000	0.978	0.993	1.000	0.997	0.997	0.988	1.000

Boldface values are statistically significant at the 5% nominal level.

7. CONCLUDING REMARKS

We have shown that the Alkire and Foster (2008) (Alkire–Foster) multidimensional approach to poverty naturally gives rise to the consideration of multiple hypotheses. Specific examples include examining the robustness of the AF ordering to the choice of poverty lines and/or the number of dimensions of deprivation before one is considered poor, inferring poverty orderings of various populations relative to a benchmark population, and inferring the specific dimensions in which a population is underachieving. Additionally, we have shown how such hypotheses can be treated in a unified manner and also tested using the minimum p-value (MinP) methodology of Bennett (2010).

In applying our proposed methodology to study Hindu and Muslim poverty in India, we have illustrated the tremendous scope for examining a wide range of hypotheses and for revealing insights into the plight of the poor not otherwise captured by traditional univariate approaches to poverty analysis. Our use of India's National Sample Survey in this illustrative example, however, motivates a thorough consideration of issues raised by the application of our methodology under various sampling designs. While beyond the scope of the current article, research into sampling design related issues in currently in progress.

Finally, our focus in this article has been on how to formulate and test rather general hypotheses in the specific context of the Alkire–Foster (2008) multidimensional poverty measure. We note, however, that our proposed tests can be extended to test hypotheses that arise from alternative multidimensional poverty or inequality orderings. Obvious examples include the multidimensional orderings of Maasoumi and Lugo (2008) and Duclos et al. (2006). Further, we note that our proposed testing procedures can be extended to allow for sample-dependent measurement parameters—e.g., estimated poverty lines—as opposed to the simpler case of exogenous parameters as treated herein.

APPENDIX A: PROOFS

Proof of Theorem 4.1. The proof turns out to be rather straightforward once we combine the fact that $\mathcal{F}$ can be built up from simple Donsker classes together with well established results on Donsker preservation. Thus, first introduce the classes

$$\mathcal{G}_1 = \left\{ \mathbb{1}\left(\sum_{j=1}^{d} \omega_j \mathbb{1}(x_j \le \ell_j) \ge k \right) : \ell \in \mathbb{R}^d_{++}, k \in [\underline{k}, \bar{k}], \sum \omega_i = d, \omega_i \ge 0 \right\}, \tag{20}$$

$$\mathcal{G}_2 = \left\{ \omega \left(\frac{\ell - x}{\ell} \right)^{\alpha} \mathbb{1}(x \leq \ell) : \ell \in [0, \bar{\ell}], 1 \leq \alpha \leq 3, \omega \in \mathbb{R} \right\},$$

and

$$\mathcal{G}_3 = \left\{ \omega \mathbb{1}(x \leq \ell) : \ell \in [0, \bar{\ell}], \omega \in \mathbb{R} \right\}.$$

That $\mathcal{G}_2$ and $\mathcal{G}_3$ are Donsker follows trivially from Theorem 9.23 and Lemma 9.8 of Kosorok (2008), respectively. By appealing again to Lemma 9.8 Kosorok (2008), it follows directly that $\mathcal{G}_1$ is Donsker if the collection

$$\mathcal{A} = \left\{ A(\ell, \omega, k) : k \in [\underline{k}, \bar{k}] \subset \mathbb{R}_{++}, \ell \in \mathbb{R}^d_{++}, \sum \omega_j = d, \omega \in \mathbb{R}^d_+ \right\}, \quad (21)$$

where $A(\ell, \omega, k) = \{x \in \mathbb{R}^d : \sum_{j=1}^d \omega_j \mathbb{1}(x_j \leq \ell_j) \geq k\}$, forms a Vapnik–Chervonenkis (VC) class of sets. Letting $D = \{1, \ldots, d\}$ and recognizing that $A \in \mathcal{A}$ is always of the form

$$A = \prod_{j \in S \subseteq D} (-\infty, \ell_j] \times \prod_{j \in D \setminus S} (-\infty, \infty),$$

it follows that $\mathcal{A}$ is a subset of the collection of cells in $\mathbb{R}^d$, and thus is VC with VC-index less than or equal to $d + 1$.

Given that $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_3$ are (uniformly bounded) Donsker classes, the proof is completed upon repeated application of Corollary 9.32 together with Theorem 9.31 of Kosorok (2008). □

Proof of Theorem 4.2. That $\mathcal{F}$ is a uniformly bounded Donsker class follows from our proof of Theorem 4.1 above. As an immediate consequence, we obtain

$$\sqrt{n_1}(P_{n_1,X} - P_X)f \rightsquigarrow \mathbb{G}_{P_X} f, \quad (22)$$

and

$$\sqrt{n_2}(P_{n_2,Y} - P_Y)f \rightsquigarrow \mathbb{G}_{P_Y} f, \quad (23)$$

in $\ell^\infty(\mathcal{F})$, where $\rightsquigarrow$ denotes weak convergence. Then, noting that

$$\left(\frac{n_1 n_2}{n_1 + n_2} \right)^{1/2} \left[n_1^{-1/2} \mathbb{G}_{n_1,P_X} f_1 - n_2^{-1/2} \mathbb{G}_{n_2,P_Y} f_2 \right] \quad (24)$$

may be written as

$$\left[\left(\frac{n_2}{n_1+n_2}\right)^{1/2}\sqrt{n_1}(P_{n_1,X}-P_X)f_1-\left(\frac{n_1}{n_1+n_2}\right)^{1/2}\sqrt{n_2}(P_{n_2,Y}-P_Y)f_2\right], \tag{25}$$

where $f_1, f_2 \in \mathcal{F}$, we obtain the desired result as a direct consequence of (22), (23), the assumed independence of the processes, and the convergence of the pre-multiplicative ratios as implied by Assumption 4.1. □

Proof of Theorem 4.3. The proof is analogous to that of Theorem 4.2 and is therefore omitted. □

ACKNOWLEDGMENTS

We wish to thank James Foster, Sabina Alkire, and participants at the Oxford Poverty & Human Development Initiative workshop on Robustness Methods for Multidimensional Welfare Analysis for their comments. We also wish to thank Chrysanthi Hatzimasoura for alerting us to an error in our original R code.

REFERENCES

Alkire, S., Foster, J. (2008). Counting and multidimensional poverty measurement. OPHI Working Paper No. 7, University of Oxford.

Alkire, S., Seth, S. (2008). Multidimensional poverty measurement and BPL methods in India: A comparison of methods. OPHI Working Paper No. 15, University of Oxford.

Anderson, G. (1996). Nonparametric tests of stochastic dominance in income distributions. *Econometrica* 64(5):1183–1193.

Barrett, G. F., Donald, S. G. (2003). Consistent tests for stochastic dominance. *Econometrica* 71(1): 71–104.

Batana, Y. (2008). Multidimensional measurement of poverty in Sub-Saharan Africa. OPHI Working Paper No. 13, University of Oxford.

Bennett, C. J. (2010). Improving the power of multiple testing procedures via single-step methods. Working Paper No. 1654062, SSRN.

Bhattacharya, D. (2007). Inference on inequality from household survey data. *Journal of Econometrics* 137(2):674–707.

Bourguingnon, F., Chakravarty, S. R. (2005). The measurement of multidimensional poverty. *Journal of Economic Inequality* 1:25–49.

Davidson, R., Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica* 68(6):1435–1464.

Davidson, R., Duclos, J.-Y. (2006). Testing for restricted stochastic dominance. IZA Discussion Paper No. 2047.

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

Foster, J., Greer, J., Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica* 52(3):761–66.

Hsu, P.-H., Hsu, Y.-C., Kuan, C.-M. (2010). Testing the predictive ability of technical analysis using a new stepwise test without data snooping bias. *Journal of Empirical Finance* 17(3):471–484.

Kakwani, N., Silber, J. (2008). *Quantitative Approaches to Multidimensional Poverty Measurement.* Basingstoke, UK: Palgrave Macmillan.

Kodde, D. A., Palm, F. C. (1986). Wald criteria for jointly testing equality and inequality restrictions. *Econometrica* 54(5):1243.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference.* New York: Springer.

Linton, O., Maasoumi, E., Whang, Y.-J. (2005). Consistent testing for stochastic dominance under general sampling schemes. *Review of Economic Studies* 72(3):735–765.

Maasoumi, E., Lugo, M. A. (2008). The information basis of multivariate poverty assessments. In: *Quantitative Approaches to Multidimensional Poverty Measurement.* Oxford, UK: Palgrave-MacMillan, pp. 1–29.

Mitra, S. (2011). Topics in Measurement: Multidimensional Poverty and Polarization, PhD Thesis, Vanderbilt University.

Romano, J. P., Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica* 73(4):1237–1282.

Santos, M. E., Ura, K. (2008). Multidimensional poverty in Bhutan: Estimates and policy implications. OPHI Working Paper No. 14, University of Oxford.

Sen, A. (1976). Poverty: An ordinal approach to measurement. *Econometrica* 44(2):219–231.

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

Wolak, F. (1989). Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models. *Econometric Theory* 5:1.