

A SOLUTION TO AGGREGATION AND AN APPLICATION TO MULTIDIMENSIONAL ‘WELL-BEING’ FRONTIERS

ESFANDIAR MAASOUMI AND JEFFREY S. RACINE

ABSTRACT. We propose a new technique for identification and estimation of aggregation functions in multidimensional evaluations and multiple indicator settings. These functions may represent “latent” objects. They occur in many different contexts, for instance in propensity scores, multivariate measures of well-being and the related analysis of inequality and poverty, and in equivalence scales. Technical advances allow nonparametric inference on the joint distribution of continuous and discrete indicators of well-being, such as income and health, conditional on joint values of other continuous and discrete attributes, such as education and geographical groupings. In a multi-attribute setting, “quantiles” are “frontiers” that define equivalent sets of covariate values. We identify these frontiers nonparametrically at first. Then we suggest “parametrically equivalent” characterizations of these frontiers that reveal likely weights for, and substitutions between different attributes for different groups, and at different quantiles. These estimated parametric functionals are “ideal” aggregators in a certain sense which we make clear. They correspond directly to measures of aggregate well-being popularized in the earliest multidimensional inequality measures in Maasoumi (1986). This new approach resolves a classic problem of assigning weights to multiple indicators such as dimensions of well-being, as well as empirically incorporating the key component in multidimensional analysis, the relationship between the indicators. It introduces a new way for robust estimation of “quantile frontiers”, allowing “complete” assessments, such as multidimensional poverty measurements. In our substantive application, we discover extensive heterogeneity in individual evaluation functions. This leads us to perform robust, weak uniform rankings as afforded by tests for multivariate stochastic dominance. A demonstration is provided based on the Indonesian data analyzed for multidimensional poverty in Maasoumi & Lugo (2008).

1. INTRODUCTION

The holy grail of multiple indicator settings in many fields of science is to discover credible weights for and relations between the indicators. These questions arise in many settings such as in multidimensional analysis of well-being advocated by Amartya Sen and others,

Date: June 11, 2014.

We would like to thank Steve Koch and Paul Kattuman for extensive comments that helped improve the presentation of this paper, two very helpful referee reports, the co-editor, and Oliver Linton and participants at the conference on Advances in Stochastic Dominance, Trinity College, Cambridge, June 2013.

in multi-variable “matching” as a basis for counterfactual analysis in program evaluations, health status characterizations, and equivalence scales, to name a few.

Traditional solutions have been restricted. As examples of these solutions, we mention principal components, dynamic multiple indicator models, and subjective selection of weights and (implicitly) relations between different indicators. Almost all of these approaches are set within linear spaces and restricted to a priori diffusion processes. A much criticized example is the United Nations index of well-being, which employs subjective weights for several indicators, in a weighted sum/average functional which clearly assumes infinite substitutability between different dimensions. Propensity scores can also be interpreted as probability transforms of aggregation functions of conditioning variables, based on regression and/or maximum likelihood optimization; see Ginindza & Maasoumi (forthcoming).

The primacy of the underlying relations between the dimensions/indicators has been widely recognized since the classic work of Atkinson & Bourguignon (1982); See Duclos, Sahn & Younger (2006). These relations determine compensatory or reinforcing effects of the indicators. The lingering question has been how to discover these relations in settings in which there is no direct observation of the “latent” dependent variable (well-being, say), which would allow for regression type solutions. In this paper we offer a solution based on a mapping from the joint distribution of the multiple attributes, possibly conditional on others, to the aggregator functions that represent well-being, or whatever the latent concept of interest.

Our primary insights are that joint distributions contain all the statistical information there is, and if estimated nonparametrically, they are least contaminated by a priori restrictions. The other insight is that multi-variable quantiles are “frontiers”. These are contours that are equi-probable surfaces that correspond to the very aggregation functions we seek to discover. Our other observation is that there are eminently strong parametric candidates for such aggregation functions which admit estimation and identification based on our nonparametric quantile estimates, not just the ones we employ in this paper.

Our approach provide estimates for the sought after weights and substitution values for different indicators. We demonstrate the feasibility of our approach, and indicate its utility in measurement and uniform ranking of multidimensional well-being, especially focusing on poverty. This is an improvement over subjective assignment of these parameters, as in the United Nations Development Index (UNDI), or the Alkire-Foster measures of deprivation.

Formally, let z be a vector of conditioning information (e.g. group membership, such as education level and/or location), and denote by y the main variable(s) of interest (e.g. income and/or health). Let $f(y|z)$ and $F(y|z)$ denote the conditional PDF and CDF, respectively. Let $x = (y, z)$, $F(y|z) = \int^y f(t|z)dt = \int^y f(x)dt/f(z)$ where $f(x)$ is the joint PDF and $f(z)$ the marginal PDF, respectively.

Consistent nonparametric kernel-based estimators of $f(y|z)$ and $F(y|z)$ are given by

$$(1) \quad \begin{aligned} \hat{f}(y|z) &= \frac{\hat{f}(y, z)}{\hat{f}(z)} = \frac{n^{-1} \sum_{i=1}^n K_{\gamma_y}(Y_i, y) K_{\gamma_z}(Z_i, z)}{n^{-1} \sum_{i=1}^n K_{\gamma_z}(Z_i, z)} \\ \hat{F}(y|z) &= \frac{\int^y \hat{f}(t, z) dt}{\hat{f}(z)} = \frac{n^{-1} \sum_{i=1}^n G_{\gamma_y}(Y_i, y) K_{\gamma_z}(Z_i, z)}{n^{-1} \sum_{i=1}^n K_{\gamma_z}(Z_i, z)} \end{aligned}$$

where $K_{\gamma}(\cdot)$ is the generalized product kernel outlined in Hall, Racine & Li (2004) and $G_{\gamma}(\cdot)$ is the generalised product kernel outlined in Li & Racine (2008), the former suitable for PDF estimation and the latter for CDF estimation. We direct the interested reader to these articles for theoretical underpinnings of the proposed approach. The bandwidths (i.e. the γ s) are obtained via a data-driven cross-validation approach also given in the aforementioned references. Under standard regularity conditions, these are consistent estimators of $f(y|z)$ and $F(y|z)$.

A conditional τ th quantile of a conditional distribution function $F(\cdot|z)$ is defined by ($\tau \in (0, 1)$)

$$(2) \quad q_{\tau}(z) = \inf\{y: F(y|z) \geq \tau\} = F^{-1}(\tau|z).$$

Or equivalently, $F(q_\tau(z)|z) = \tau$. Our approach will be based on nonparametrically estimating $F(y|z)$ via the estimator outlined above and then inverting this estimate to obtain $q_\tau(z)$ using the approach outlined in Li & Racine (2008). We let $\hat{q}_\tau(z)$ denote our resulting quantile (set) estimate (when $y \in \mathbb{R}^2$ this is the locus of (y_1, y_2) for which the conditional CDF equals τ , as in an isoquant).

Each member of a quantile set, at a given level, corresponds to a potential member of the population. An evaluation or aggregation function is required for cardinal representation of that member. Such a function is like a utility function but need not be one. It is a multidimensional representation of the “status” of each individual (unit). All individuals below a certain quantile set will be “poorer” in a multidimensional sense. The principal questions in this paper are, how to obtain and quantify this functional relationship, parametrically, based on unrestricted smooth nonparametric estimates of the quantile graph. Different parameter values may apply at different quantiles. We will find that this is so in our demonstration with Indonesian data. This is anticipated by extended “impossibility” theorems on Social Welfare Functions (SWF); see Fleubaey & Maniquet (2011) for some recent statements on this. It is compelling evidence against homogeneity and homothetic preferences extant in economics.

To see this graphically, consider Figure 1 based on the empirical study detailed later in the paper. In this figure we depict the nonparametric joint CDF of two attributes, income and health, for a particular group/conditioning variable, those with “low” education, (2 years) and residing in a region, Jawa. The plane cutting the joint CDF at the point τ identifies the solid curve on the CDF surface. This is the set of all values in (income, health) which correspond to the same quantile (probability) level τ . Let a parametric function $S(\text{income, health, } \theta|\tau)$ represent an “aggregation” of the two dimensions of well-being. Then $\Delta S(\text{income, health, } \theta|\tau) = 0$ corresponds to the solid line “isoquant” we see in Figure 1. We estimate the joint CDF to obtain nonparametric estimates of this isoquant and then fit a desired parametric form to the same, in order to obtain estimates of the unknown parameters

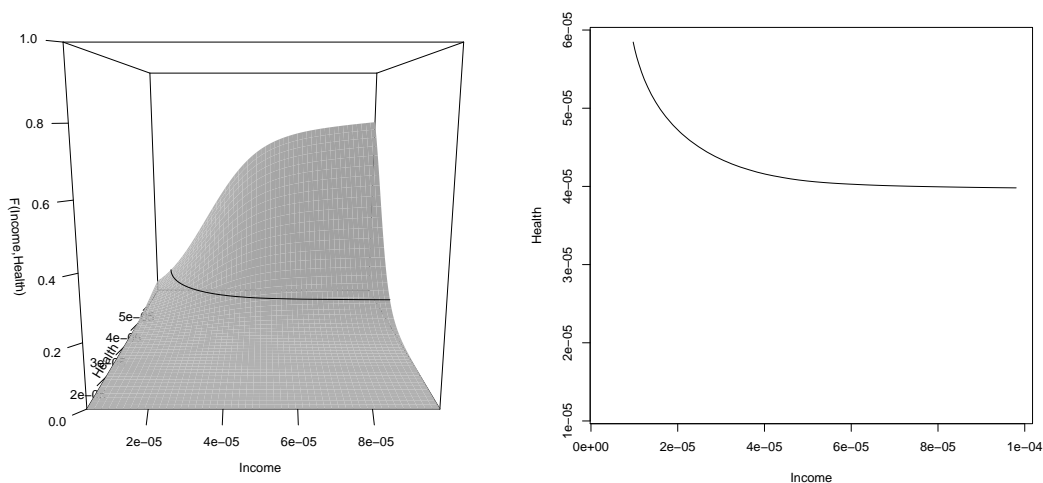


FIGURE 1. The joint CDF of health and income in Jawa along with the $\tau = 0.1$ isoquant for low education household heads.

θ . The nonparametric estimates are consistent at rates which are established in, for instance, Li & Racine (2007). This procedure produces a very well informed data driven estimate of individual “welfare evaluation functions”, represented by the $S(\cdot)$ function described above. These are central objects in multivariate analysis of poverty and inequality. Index based, “complete” ranking of welfare states is driven by these objects.

Ideal Parametric Evaluation functions: Let, $X = (X_{ij})$, $i = 1, \dots, n$, $j = 1, \dots, m$, be a well-being matrix. All index based evaluations involve, inevitably, two types of aggregation: aggregation over the n recipients of the attributes, and aggregation over the m attributes. The latter are preference functions requiring assignment of weights to each attribute, and a measure of their interactions (substitution, complementarity). Note that the choice of $m = 1$ is the univariate analysis that represents radical “zero” weights to excluded dimensions, and the consequent neglect of how they interact with the chosen attribute. Univariate approaches are thus the most subjective!

An appeal to such principles as fundamentalism and impartiality, see Kolm (1977), would suggest that the greater the number of welfare attributes considered, the more reasonable are

the common assumptions of *anonymity* and homogeneity in aggregation functionals. Among other things, and as a practical matter, this requires inclusion of individual characteristics. Otherwise we need to control for some attributes by conditioning. A lucid recent discussion of some of these issues is given in Decancq & Lugo (2012).

Our proposed functions were first derived in Maasoumi (1986) as ideal aggregators, or representation functions in the context of measuring well-being in many dimensions. They form the basis of what has come to be known in that literature as the “two-step” method of assessing multivariate inequality. The basis for referring to our proposed functional form as “ideal” is that it minimizes the “divergence” between its distribution, on the one hand, and the distributions of its constituent components, on the other. Since we wish to do this for the entire distribution of the variables, entropy metrics, particularly “relative entropies” are the optimization measures of choice here. Consider the following weighted average, over dimensions $j = 1, \dots, m$, of a generalized relative entropy between the aggregator function S_i , and each of the X_{ij} , as follows:

$$(3) \quad D_\beta(S, X; \alpha) = \sum_{j=1}^m \alpha_j \left\{ \sum_{i=1}^n S_i [(S_i/X_{ij})^\beta - 1] / \beta(\beta + 1) \right\},$$

where the α_j s are the weights given to the attributes, and β is a choice of entropic divergence between the distribution of an aggregator function, S , and its constituent attributes. As a referee pointed out, these parameters should be indexed at each quantile to anticipate the heterogeneity that will be discovered in them. Minimizing D_β with respect to S_i such that $\sum S_i = 1$, produces the following “optimal” aggregation functions:

$$(4) \quad S_i \propto \left(\sum_j^m \alpha_j X_{ij}^{-\beta} \right)^{-1/\beta}, \quad \beta \neq 0, -1,$$

$$(5) \quad S_i \propto \Pi_j X_{ij}^{\alpha_j}, \quad \beta = 0,$$

$$(6) \quad S_i \propto \sum_j \alpha_j X_{ij}, \quad \beta = -1.$$

These are, respectively, the hyperbolic, the generalized geometric, and the weighted means of the attributes, see Maasoumi (1986). The raw data may be standardized in a number of ways to have some numerical parity between otherwise very different measurement units in different indicators. We focus on “shares” which are unit invariant and have a size distribution interpretation for all attributes.

The “Constant Elasticity of Substitution” (CES) parameter is related to the choice of the entropy metric by $\sigma = 1/(1 + \beta)$. To appreciate the generality of the functional family here, note that it includes the weighted arithmetic mean ($\beta = -1$) which itself subsumes most of the popular composite indicators, such as those based on “Multiple Indicators Multiple Causes” (MIMIC) models, or the Principal Components (PC) of X , when α_j s are the elements of the first eigenvector of the $X'X$ matrix; see Ram (1982) and Maasoumi (1989). Indeed, even the often criticized Human Development Index (HDI) of well-being is a special case of the weighted arithmetic mean, with rather arbitrary weights. See Decancq & Lugo (2012) for a recent survey on these issues. Note that a linear function imposes infinite substitutability between such dimensions of well-being as income, health and education! Relatedly, a linear or nonlinear aggregator with “fixed parameters” implies that a very wealthy individual, for instance, has the same valuation of a dollar of income, in terms of health or education, as a very income poor individual. Our empirical results are able to reveal different parameter estimates of weights and substitution at different quantiles, and for different conditioning sets. Here, the traditionally difficult issue of heterogeneity is addressed empirically.

The additively separable nature of these wellbeing indices (see also the Alkire-Foster and UN indices) has received some criticism (Ravallion (2010)) since it severely limits the nature of complementarity and substitutability between “goods” or “wellbeings” admitted by such indices. This can be rectified by adding cross product terms in the various dimensions of wellbeing as in Anderson, Leo & Anand (2014). Our technique can be used to estimate more complicated indices than ours.

We also note, that biased view of attribute contributions may result due to omission of other related attributes. This reservation applies to all estimation methods.

As noted earlier, other examples of estimation for evaluation functions are PC and factor analysis. These impose strong assumptions, including “linearity” of aggregation, and homogeneity among different groups. These methods are also fundamentally based on “variance” decomposition, thus employing variance as a metric for the distributional information, and for assessing good estimation. This is not satisfactory for attribute distributions which are clearly non-Gaussian.

What is now abundantly clear is that the “relation”, or dependence, between attributes is the essence of the “multidimensional” assessments. Placing the joint distribution of the attributes at the center of all inferences ensures consistency with this fundamental feature.

Dominance rankings. Robust uniform ranking of distributions of well-being is possible, in principle, over classes of welfare functions that commonly characterize various orders of stochastic dominance. This type of uniform/incomplete ranking is even more desirable in multidimensional settings since, in principle, one can avoid two sets of cardinal aggregations, one over the population and another over the attribute dimensions. The first influential writing in the area of multidimensional dominance is Atkinson & Bourguignon (1982), where the complexity of the question is laid bare. In particular, the complex role of “statistical dependence” between attributes is explicated through a discussion of “correlations”. Groups with different “needs” and other characteristics cannot credibly be represented with the same “utility functions” based on other covariates. Allowing for transfers between groups as well as within groups demands a clarification of how different groups value one attribute (income, say) in terms of other attributes (health and education, say). This heterogeneity is very evident in our empirical example. This invites weak uniform ranking based on tests for first and second order dominance. We will report tests based on generalized Kolmogorov-Smirnov

statistics. When distribution functions cross, the approach is able to identify quantiles below which poverty rankings are sustained and robust to the choice of a “poverty line”.

Note, however, that our evaluation function estimation also opens the door to multi-variable definitions of “poverty lines” by reference to the distribution of the aggregate functions; see below. This approach is referred to as the “intermediate” definition of poverty sets, subsuming the “intersection” and “union” definitions. The intersection convention considers a unit as “poor” if it is poor in all dimensions; see Maasoumi & Lugo (2008), and Bourguignon & Chakravarty (2003).

With respect to dominance rankings, our work is generally in the same spirit as the ambitious work of Duclos et al. (2006) and Duclos, Sahn & Younger (2010). The distinction here is estimation of new parametrically fitted nonparametric poverty frontiers, and tests of stochastic dominance as proposed by Linton, Maasoumi & Whang (2005), Linton, Maasoumi & Whang (2008).

The rest of this paper is organized as follows: Section 2 describes the estimation of the aggregator functions. Section 3 presents the dominance tests and relevant background. Section 4 illustrates our approach via profiles of distributions and quantiles and the optimal multi-variable parametric frontiers based on Indonesian data. Section 5 concludes.

2. NONPARAMETRIC ISO-WELL-BEING SETS AND EQUIVALENT PARAMETRIC CHARACTERIZATIONS

For the analysis at hand we shall first need to estimate a family of fundamental statistical objects, namely, the conditional density, conditional distribution, and conditional quantile functions. These objects are frequently defined over a mix of continuous and categorical datatypes (i.e. the variable ‘group’ defined in the application that follows is categorical). Modeling distributions defined over mixed datatypes is known to be “parametrically awkward” (Aitchison & Aitken (1976, page 419)). Using kernel smoothed nonparametric methods allows us to consistently estimate the PDF, CDF, and quantile functions of income and

health, conditional on, say, group membership. Our estimates are determined purely by the data at hand and, aside from smoothness, place minimal prior structure on the resulting estimates.

After nonparametrically estimating $q_\tau(z)$, we condition on desired values of z (e.g. a particular group of interest).

Without loss of generality, let $y \in \mathbb{R}^2$. The CES aggregation function is given by

$$(7) \quad S(y_1, y_2) = A \left(\alpha y_1^{-\beta} + (1 - \alpha) y_2^{-\beta} \right)^{-1/\beta}$$

where $A > 0$ and $0 < \alpha < 1$. The partial derivatives are as follows,

$$(8) \quad S_{y_1} = \frac{\partial S(y_1, y_2)}{\partial y_1} = A \alpha \left(\alpha y_1^{-\beta} + (1 - \alpha) y_2^{-\beta} \right)^{(-1/\beta)-1} y_1^{-\beta-1}$$

and,

$$(9) \quad S_{y_2} = \frac{\partial S(y_1, y_2)}{\partial y_2} = A(1 - \alpha) \left(\alpha y_1^{-\beta} + (1 - \alpha) y_2^{-\beta} \right)^{(-1/\beta)-1} y_2^{-\beta-1}.$$

Along an iso-well-being quantile $\Delta S(\cdot) = 0$ (i.e. $S_{y_1} \partial y_1 + S_{y_2} \partial y_2 = 0$) hence,

$$(10) \quad -\frac{S_{y_1}}{S_{y_2}} = \frac{\partial y_2}{\partial y_1} = \frac{\alpha}{\alpha - 1} \left(\frac{y_2}{y_1} \right)^{\beta+1}.$$

We exploit the fact that, for $y = (y_1, y_2)$, conditional on z , we can obtain estimates of $\partial y_2 / \partial y_1$ directly from the estimated quantile $q_\tau(z)$ (i.e. for a given value of τ we can compute $\partial y_2 / \partial y_1$ since the level of multidimensional well-being is constant).

It is a simple matter to obtain estimates of α and β via (nonlinear) regression of our nonparametrically estimated $\partial y_2 / \partial y_1$ on y_2 / y_1 using (10) (or take logs and run a log linear regression). The parameter estimates, and fitted values of the $S(\cdot)$ functions, are generally consistent as smooth functions of consistent nonparametric estimates.

Of course, the parameter values may change at different quantiles and different values of the conditioning variables z . This is an added benefit of the approach (i.e. we naturally allow

for a varying coefficient representation where the values of α and β can freely vary with the quantile).

Note that this approach provides measures of how the different attributes contribute to welfare (and the substitution among them) at each chosen welfare level. With an optimisation interpretation, optimality would require that the ratio of marginal products of attributes equals the ratio of their implicit “prices”, as the elasticity of substitution is the relative change in the attributes ratio due to the relative change in attribute “prices”. If the estimated elasticities of substitution differ between welfare levels (implicit attribute price ratios may differ between welfare levels), then the same changes in “prices” will lead to different “optimising” investment responses at different quantiles. This would be a policy relevant extension, for example, in the multi-dimensional evaluation of schools, universities, business schools, hospitals, firms etc. At quantiles where the market price ratio and implicit price ratio do not agree, there is probably potential for improving the investment choices and thus welfare. ¹

3. WEAK UNIFORM RANKINGS (STOCHASTIC DOMINANCE)

The second approach alluded to earlier is based on the desire to avoid full cardinalization required by the index/aggregation approach.

3.1. Definitions and Tests in the Univariate Case. Let W and V be two income variables at either two different points in time, before and after taxes, or for different regions or countries. Let W_1, W_2, \dots, W_{N_1} be N_1 not necessarily i.i.d observations on W , and V_1, V_2, \dots, V_{N_2} be similar observations on V . Let U_1 denote the class of all utility functions u such that $u' \geq 0$, (increasing). Also, let U_2 denote the class of all utility functions in U_1 for which $u'' \leq 0$ (strict concavity). Let $W_{(i)}$ and $V_{(i)}$ denote the i -th order statistics, and

¹we thank a referee for pointing out this potential for optimality analysis. The case where some attributes are discrete poses no issues when estimating the joint distribution nonparametrically. The interpretation of relative changes and substitution must be carried out with respect to discrete measures. Kannai (1980) provides a helpful description of the Auspitz-Lieben-Edgeworth-Parato (ALEP) concepts of substitution and complementarity for discrete variates.

assume $F_1(w)$ and $F_2(w)$ are continuous and monotonic cumulative distribution functions (CDF's) of W and V , respectively.

Quantiles $q_\tau(w)$ and $q_\tau(v)$ are implicitly defined by, for example, $F[W \leq q_\tau(w)] = \tau$.

Definition 3.1. *W First Order Stochastic Dominates V , denoted W FSD V , if and only if any one of the following equivalent conditions holds:*

- (1) $E[u(W)] \geq E[u(V)]$ for all $u \in U_1$, with strict inequality for some u .
- (2) $F_1(w) \leq F_2(w)$ for all w in the support of W , with strict inequality for some w .
- (3) $q_\tau(w) \geq q_\tau(v)$ for all $0 \leq \tau \leq 1$.

Definition 3.2. *W Second Order Stochastic Dominates V , denoted W SSD V , if and only if any of the following equivalent conditions holds:*

- (1) $E[u(W)] \geq E[u(V)]$ for all $u \in U_2$, with strict inequality for some s .
- (2) $\int_{-\infty}^w F_1(t)dt \leq \int_{-\infty}^w F_2(t)dt$ for all w in the support of W and V , with strict inequality for some w .
- (3) $\int_0^\tau q_t(w)dt \geq \int_0^\tau q_t(v)dt$, for all $0 \leq \tau \leq 1$, with strict inequality for some value(s) τ .

Our preferred tests of FSD and SSD are based on various empirical evaluations of the CDF-based conditions (2) or (3) in the above definitions. Evidently, smoothed nonparametric estimates of both the CDFs and quantiles present alternative tests.

For stochastic dominance tests, we generally follow Linton et al. (2005). But they employ the simplest nonparametric estimator, the Empirical CDF (ECDF), whereas we employ both the ECDF and the smoothed nonparametric estimates. A two-way Kolmogorov-Smirnov test of a null hypothesis of dominance, such as in condition (2), requires comparisons at a finite number of points on the joint support of the multidimensional sample. Since higher order dominance implies lower orders, finding FSD to a statistical degree of confidence implies SSD, etc.

Theorem 3.1. *Given the mathematical regularity conditions:*

(1) *The variables i, j are first-order stochastically ranked; i.e.*

$$d_1 = \min_{i \neq j} \sup_w [F_i(w) - F_j(w)] < 0,$$

if and only if for each prospect i and j , there exists a continuous increasing function u such that $E u(W_i) > E u(W_j)$. i and j are from a countable set of K “prospects” that are being ranked. In our paper all rankings are binary, so $K=2$.

(2) *The variables are second order stochastically ranked; i.e.*

$$d_2 = \min_{i \neq j} \sup_w \int_{-\infty}^w [F_i(\mu) - F_j(\mu)] d\mu < 0,$$

if and only if for each i and j , there exists a continuous increasing and strictly concave function u such that $E u(W_i) > E u(W_j)$.

(3) *Linton et al. (2005) deal with stochastic processes that are strictly stationary and α – mixing with $\alpha(j) = O(j^{-\delta})$, for some $\delta > 1$. Our data are iid cross-sections which trivially satisfy their assumptions. When ECDFs are used, we have:*

$$d_{1,N} \rightarrow d_1, \text{ and } d_{2,N} \rightarrow d_2,$$

where $d_{1,N}$ and $d_{2,N}$ are the empirical test statistics defined as:

$$d_{1,N} = \sqrt{\frac{N_i N_j}{N_i + N_j}} \min_{i \neq j} \max_w [F_{iN_i}(w) - F_{jN_j}(w)], \text{ and}$$

$$d_{2,N} = \sqrt{\frac{N_i N_j}{N_i + N_j}} \min_{i \neq j} \max_w \int_0^w [F_{iN_i}(\mu) - F_{jN_j}(\mu)] d\mu$$

where the Empirical CDFs are given by

$$(11) \quad F_{jN_j}(w) = \frac{1}{N_j} \sum_{i=1}^{N_j} I_{W_j \leq w}, \quad j = 0, 1.$$

In our empirical application, we don't have dependence within sample observations, and do not expect dependence between samples/groups that are being ranked. Thus subsampling

is not used to conduct the tests. We report the tests based on nonparametric smoothed CDFs in the empirical section (Tests based on ECDFs were also conducted which agreed with those reported).

3.2. Multivariate Stochastic Dominance. Atkinson & Bourguignon (1982), and Atkinson & Bourguignon (1987) developed some conditions for ranking multi-dimensioned distributions of welfare attributes. SWFs are taken to be individualistic and (for convenience) separable. But anonymity may be dropped in recognition of the fact that households (individuals) must be distinguished according to their distinct needs or other characteristics, such as the conditioning variables in our analysis. Here we wish to rank the “joint” distributions of income and health, within groups, and between two states (populations), conditional on characteristics such as ethnicity and education.

3.2.1. Testing for Multivariate conditional Dominance Relations. Formally, we consider testing for dominance relationships of

$$F(\text{income, health}|\text{group, education=low})$$

versus

$$F(\text{income, health}|\text{group, education=high}).$$

This will be a “within-group” test by education. In our empirical section, we also keep education fixed, and test across (ethnic) groups. A lucid discussion of tests for multivariate SD is given in Duclos et al. (2006) and Duclos et al. (2010). They employ different tests that are based on functionals of the CDFs.

We consider testing for FSD using the following, two way, Kolmogorov-Smirnov (KS) statistic:

$$(12) \quad D = \min \{ \max(F_1 - F_2), \max(F_2 - F_1) \},$$

where F_1 and F_2 are two multivariate conditional CDFs that differ according to the values of their conditioning covariates, in this case, education (high, low). $D \leq 0$ indicates FSD, and $D > 0$ indicates no dominance. We therefore consider testing the hypothesis $H_0 : D > 0$ versus $H_1 : D \leq 0$. Note that, because this is a McFadden-type generalization of the KS test, when we reject H_0 we are able to infer which distribution statistically dominates the other.

We elect to use a nonparametric bootstrap method whereby we impose the “least favorable” member of the null, that is $F_1 = F_2$. We do this by first drawing a bootstrap sample pairwise, and then once more bootstrapping e.g. the education variable only (i.e. shuffling in place leaving the remaining variables unchanged), thereby removing any systematic relationship between education and other variables (health, income, and geography/ethnicity) for the bootstrap sample. To construct the KS test, we evaluate F_1 and F_2 on a 25×25 grid, lying in the 0.025th to the 0.975th quantiles of the data for health and income. We conduct $B = 999$ bootstrap replications and report the statistic D along with its nonparametric P -value given by

$$(13) \quad \hat{P} = B^{-1} \sum_{i=1}^B I_{D < D^*}$$

where $I_{D < D^*}$ is an indicator function equal to one when the sample statistic D is less than the bootstrap statistic computed under the null (D^*) and zero otherwise. In other words, it is the proportion of bootstrap statistics more extreme than the sample statistic (i.e. more negative)².

In this paper we consider both the smooth version of the test based upon Li & Racine (2008, Equation (7)) and the frequency-based counterpart that does not smooth the conditioning variables (the frequency-based counterpart uses the empirical distribution function kernel

²Re-centering methods are also available that avoid imposing the least favorable null. See Linton et al. (2005) and more recently Donald & Hsu (2012)

for Y when constructing $\hat{F}(y|x)$ but uses classical smoothing for X , i.e. Li & Racine (2008, Equation (4))).

In order to interpret our findings for different groups and educational levels, it would be useful to recount some of the conditions described by Atkinson & Bourguignon (1982).

Let there be G groups which are characterized in terms of their “health” and “incomes”. It is assumed that all members within a group $g \in G$ have the same valuation and marginal valuation of income. If there were no income transfers between groups, the necessary and sufficient conditions for FSD and SSD, given above, must hold for *all* groups, for FSD and SSD to hold overall. If there is any transfer between groups, however, one must deal with each group’s evaluation as well as between-group valuations of the trade-offs between income and “health”.

Interpersonal comparisons of well-being are inevitable whenever heterogeneous populations are involved. This can give rise to an “impossibility” of unambiguous or *consensus* rankings. But, *majority* rankings are possible with plausible restrictions. To see this, it is worthwhile to formally describe the conditions of Atkinson & Bourguignon (1987) here as they combine the desirable elements of “decomposability” axioms and partial ordering. Although this avoids full cardinalization, it shows the directions in which an analyst may wish to make increasingly normative assumptions to *approach cardinality*; see Basu (1980).

Let $S(Y, H)$ denote private valuations of income Y and “health” (or needs). $w(S(Y, H))$ or just $w(Y, H)$ represents the social welfare (or decision) function, and p_g , $g = 1, 2, \dots, G$, the marginal frequency in group g . The cumulative function is $P_g = \sum_{j=1}^g p_j$, $P_G = 1$. Social valuation of income received by household g is $S^g(Y)$ which is assumed continuously differentiable as needed. It is assumed that the first partial derivatives $S_Y^g \geq 0$, and $S_{Y_Y}^g \leq 0$.

If no assumptions are made about how S^g varies with g the conditions of FSD and SSD must hold for all groups g for FSD and SSD to hold. These are strong conditions. Among other things, they require that the mean income of all groups must be no lower in the dominant distribution. This would rule out equalizing redistributions between groups with

different needs. To resolve this situation one must specify some aspects of the trade-off between incomes and health.

The traditional univariate/homogenous analysis is implicitly based on the extreme assumption that $S_Y^g(Y) = S_Y(Y), \forall g$. The *level* of welfare can vary with health, but no more. This assumption is sufficient to allow a consideration only of the marginal distribution of income. But suppose one were to follow Sen in weakening the “Equity Axiom” by assuming that groups can be ranked by their *marginal valuation of income*. For instance, if the least healthy group has the highest marginal valuation of income, the next healthiest group has the second highest marginal valuation, and so on, then the necessary and sufficient condition for FSD of F_1 over F_2 is:

$$(14) \quad \sum_{g=1}^j p_g [F_1^g - F_2^g] \leq 0, \text{ for all } Y \text{ and all } j = 1, \dots, G,$$

where superscript indicates the income distribution for the g -th group. Note that the final condition here is the FSD of the entire marginal distribution of incomes. This is testable, of course. But as Atkinson & Bourguignon (1987) point out, marginal valuation by *society* can take into account the *level* of individual welfare. Therefore it is possible that the assumed negativity of S_{YH} may be offset by sufficient degree of concavity ($-w''/w'$) of the additive social valuation function $w(\cdot)$. Thus the ranking of groups assumed by Atkinson & Bourguignon (1987) coincides with a ranking of levels of welfare, where lower health status increases marginal valuations of income, or the social welfare function has a sufficiently large degree of concavity. This latter property is generally not testable and is subjective.

The above FSD condition may be weakened further for SSD if we are willing to assume “the differences in the social marginal valuation of income between groups become smaller as we move to higher income levels”; see Atkinson & Bourguignon (1987). That is, $-S_{YY}$ decreases for healthier groups reflecting less social concern with “differences” in health for higher income groups. If this assumption is adopted, a necessary and sufficient condition for

SSD is:

$$(15) \quad \sum_{g=1}^j p_g \left[\int_0^x (F_1^g - F_2^g) dY \right] \leq 0 \text{ for all } Y, \text{ and } j = 1, \dots, G.$$

This includes the usual SSD condition for the marginal distribution of incomes.

Atkinson & Bourguignon (1987) consider weaker SSD conditions by exploring further assumptions toward cardinality. One such assumption allows further comparability between the *differences* of S_Y and S_{YY} . Thus, if the rate of decline of social marginal valuation of income across groups is positive, *and declines with g , and* the same property holds for the degree of concavity ($-S_{YY}$), the necessary and sufficient condition for SSD is given as follows:

$$\sum_{j=1}^k \sum_{g=1}^j p_g \left[\int_0^x (F_1^g - F_2^g) dY \right] \leq 0 \text{ for all } Y \text{ and } k = 1, \dots, G - 1$$

and

$$(16) \quad \sum_{g=1}^G p_g \left[\int_0^x (F_1^g - F_2^g) dY \right] \leq 0 \text{ for all } Y.$$

It is worth noting that all the above conditions are testable using the tests outlined above, even though some of the predicates/assumptions are not.

Consistent with a philosophy of “partial comparability” developed by Sen (1970), Atkinson & Bourguignon (1987) suggested that nihilism may be avoidable if certain plausible assumptions are made about the trade-offs between incomes and other variables, such as health or “needs”, and at different levels of health, should we agree that groups can be ranked by such “other” characteristics as “health”. Our empirical results, based on Indonesian data suggest there may be too much heterogeneity at different levels of income, health, education, and/or ethnicity for the above assumptions to be reasonable across the board. For this case, we do not anticipate many findings of statistically significant multidimensional rankings.

4. THE INDONESIAN CASE

4.1. **Data.** A family of multidimensional poverty measures were analyzed in Maasoumi & Lugo (2008), and demonstrated with application to data from several regions in Indonesia. Based on the same data set, we consider the following variables:

- (i) income: per capita expenditures³
- (ii) health: levels of hemoglobin adjusted by gender and age
- (iii) education: level of education of the head of the household (years of schooling)
- (iv) group: group variable, based on assigned residence in Betawi, Jawa, and Sunda

There are $N = 19,602$ individual records. We treat level of education as an ordered factor (i.e. discrete), and “group” as an unordered factor (“Jawa”, “Sunda”, “Betawi”). Table 1 presents a summary of the data along with P -values from pairwise t -tests for equality of means between group.

TABLE 1. Data summary by group (leftmost table) and P -values from pairwise t -tests for equality of means for income, health, and education (rightmost table).

	income	health	education	Group	P-value
Jawa				income	
N	13557.00	13557.00	13557.00	Jawa–Sunda	0.0028
mean	280437.40	13.96	6.40	Jawa–Betawi	< 0.0001
Std.Dev.	303552.52	1.71	4.48	Sunda–Betawi	0.0008
Sunda				health	
N	4254.00	4254.00	4254.00	Jawa–Sunda	0.0004
mean	297689.63	13.85	6.77	Jawa–Betawi	0.2633
Std.Dev.	336723.27	1.72	4.30	Sunda–Betawi	0.0014
Betawi				education	
N	1791.00	1791.00	1791.00	Jawa–Sunda	< 0.0001
mean	329907.48	14.01	6.72	Jawa–Betawi	0.0051
Std.Dev.	342959.99	1.69	4.51	Sunda–Betawi	0.6464

The largest sample is from Jawa, the largest population mass, with the lowest “mean” income. There is considerable variability in incomes in all three regions. Mean “health”

³10,000 Indonesian Rupiah = \$1.12 USD (November 9, 2010).

level, and its relatively low “variability”, is similar in all three regions. “mean” education and its high variability is also similar in all the regions. Based on a mere mean-variance analysis, a sum-of-one dimensional ranking of well-being would seem to be: Betawi > Sunda > Jawa. The mean differences are all statistically significant, except for the mean health between Betawi and Jawa, and the mean education between Betawi and Sunda.

Our multivariate dominance tests below present a more complex, non-uniform ranking, with few exceptions.

4.2. Nonparametric estimation. We use the methods outlined in Section 2 and estimate $f(y_1, y_2|z_1, z_2) = f(\text{income, health}|\text{education, group})$, $F(y_1, y_2|z_1, z_2) = F(\text{income, health}|\text{education, group})$, and $q_\tau(z_1, z_2) = q_\tau(\text{education, group})$.

All computations are carried out in the R environment (R Core Team (2013)) and make use of the np package (Hayfield & Racine (2008)). Table 5 in the appendix presents a summary of the bandwidths and kernel functions. All code is available upon request from the authors.

Note that, for what follows, the entire sample is always used for nonparametric estimation purposes. However, in order to inspect features of distributions and quantiles, we make presentations based on two resolutions, i) no ‘cropping’ (min/max of all variables used for axes range), and ii) the min/90th quantiles of the marginals of each variables appearing on the axes, respectively. Similarly, we provide derived parameter estimates for the two parameters of interest within several finite “bands” that limit the consideration of very extreme points in the grid of values for (income, health). The presence of a small number of extreme tail values would otherwise obscure interesting features of the whole sample. For the Kolmogorov-Smirnov (KS) tests of dominance, we perform the tests on the full sample CDF estimates, but trim the tails of a 25×25 evaluation grid beyond the 0.025th to the 0.975th multidimensional quantiles of the data for health and income. This is for power considerations when computing the KS statistic.

In the plots that follow we report the estimated conditional PDF, CDF, and quantile sets for only specific group-education levels (i.e. we condition on values deemed to be of particular interest). We choose five quantile values, two education levels (2, 12), and consider all three regions. As was noted earlier, the ‘substitutability’ of income and health, conditional on education and group for a given quantile, is simply the derivative of the equivalence set (quantile set) boundary. It measures the rate at which income and health can be ‘traded-off’ holding the quantile, group, and education level constant. The ‘elasticity’ (conditional income elasticity of health) is the standard definition, i.e. ‘substitutability’ multiplied by income/health (i.e. ratio of percentage changes), again holding the quantile, group, and education level constant.

4.3. Jawa. It is evident from Table 1 that the Jawa ethnic group is by far the largest group in the data. We therefore consider this group by way of illustration. We consider the conditional CDF, holding $(z_1, z_2) = (\text{education, group}) = (2, \text{Jawa})$. This represents low education, compared with $(z_1, z_2) = (12, \text{Jawa})$ which represents high education. This allows us to focus on the joint distribution of $(y_1, y_2) = (\text{income, health})$, conditional on the covariate values. We present two graphs in Figure 2. The left is the estimated conditional CDFs for low and high education, and the right is the difference between the estimated CDFs, indicating a dominance relationship (i.e. the difference is non-negative uniformly). It is evident from these figures that there is a higher probability of low income and health for the low education level.

Figure 3 presents a number of iso-well-being contours (quantiles) taken directly from the CDF estimates plotted in the upper figure in Figure 2 for $\tau = (0.1, 0.2, \dots)$. A graphical dominance relationship is clear from Figure 3. For instance, considering the $\tau = 0.1$ quantile sets for the high and low education groups, we observe that the iso-well-being quantile for the low education group lies everywhere below that for the high education group; in particular, there is no crossing. Other than the joint tests of quantile differences implied by

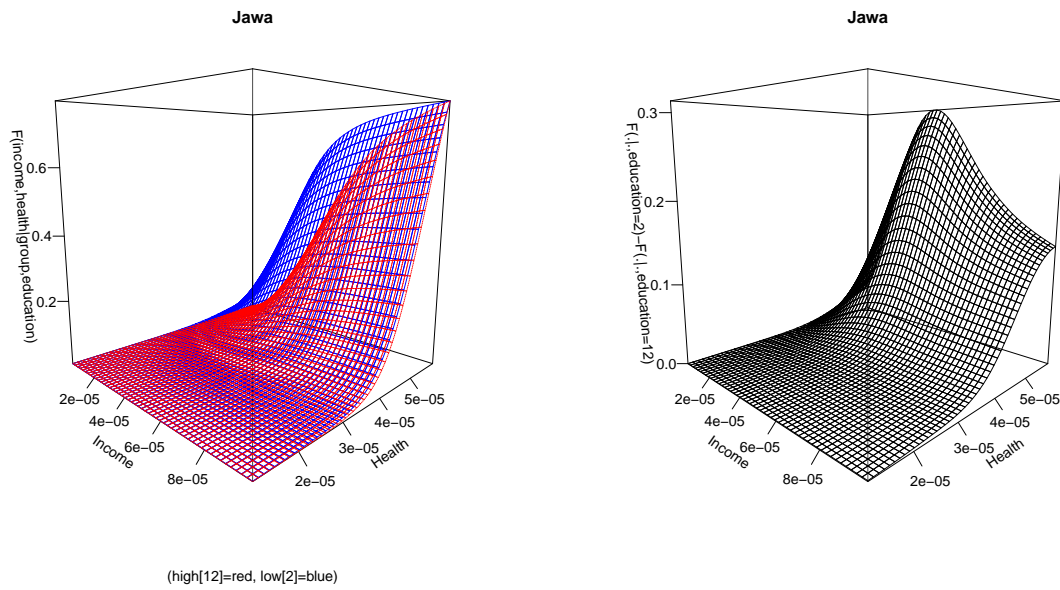


FIGURE 2. CDF and CDF difference plots, Jawa, education=2,12, cropping=0.90

our stochastic dominance tests, we do not pursue the statistical tests of the specific quantile differences in this paper.

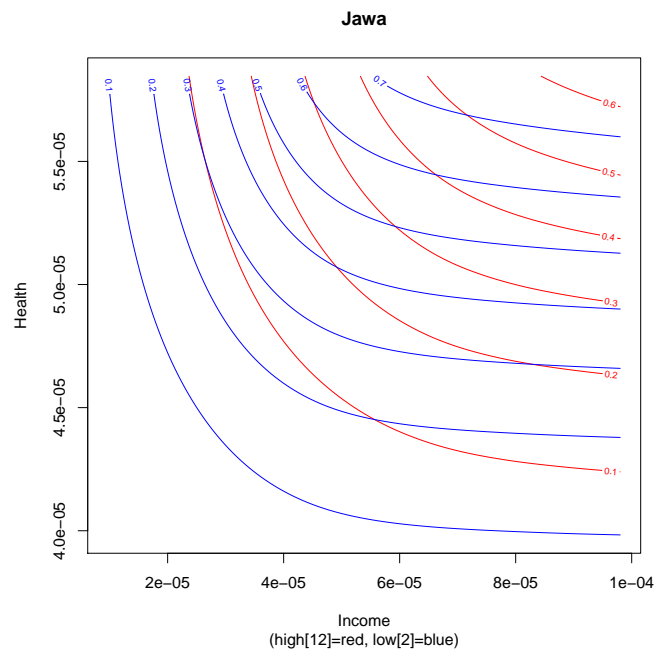


FIGURE 3. quantile plots, Jawa, education=2,12, cropping=0.90

Next we focus on a particular iso-well-being quantile, that for $\tau = 0.1$. From the iso-well-being quantile set we compute $\partial y_2/\partial y_1$, and then fit the implied parameters for the CES function (i.e. α and β). These three figures are plotted in Figure 4.

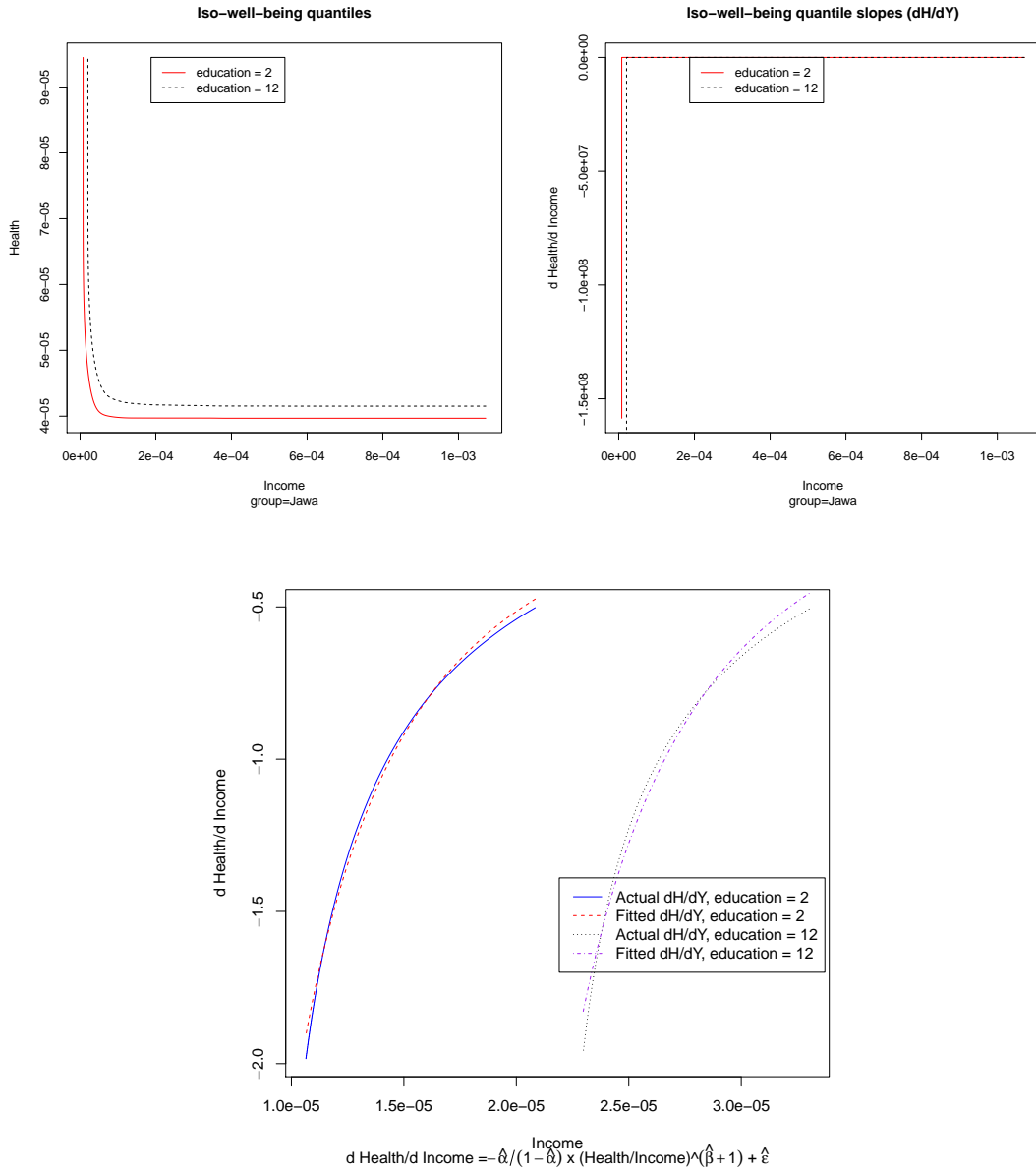


FIGURE 4. Estimated quantiles (top left), $\partial y_2/\partial y_1$ (top right), and CES estimates (bottom) for which $\gamma \leq |\partial y_2/\partial y_1| \leq 1/\gamma$, Jawa, 2/12, $\tau = 0.10$, $\gamma = 1/2$.

Both the $\tau = 0.1$ quantile planes and the corresponding S function estimates indicate a dominance of high education level over the low education level groups. This is generally

consistent across all quantiles. But the distances are not as clear cut at all quantiles, when the distance between low and high educational levels is smaller (i.e. smaller than the 10 years of schooling difference depicted in Figure 4 above).

On the other hand, the fitted values of the weights and substitution parameters are quite varied across quantiles and/or education levels. This can be seen in Table 2 which presents values for the CES parameter estimates for a range of quantiles.

The high level of heterogeneity is unmistakable. This is a significant finding as it has major implications for measurement of poverty, inequality, equivalence scales, and other summary measures. All of these objects depend, to various degrees, on presumption of relatively homogenous evaluation functions across individuals and even groups. We elaborate on this challenging issue below.

TABLE 2. Values of $\hat{\beta}$ and $\hat{\delta} = \hat{\alpha}/(1 - \hat{\alpha})$ for Jawa for various quantile values τ and levels of education (low = 2 years, high = 12 years), $\gamma = 1/2$.

τ	High		Low	
	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$
0.05	1.05	0.12	0.07	0.15
0.10	1.61	0.13	0.62	0.11
0.15	2.24	0.13	1.03	0.10
0.20	2.79	0.13	1.39	0.09
0.25	3.27	0.14	1.72	0.08
0.30	3.80	0.16	2.09	0.08
0.35	4.35	0.18	2.46	0.07
0.40	4.98	0.21	2.87	0.07
0.45	5.71	0.24	3.30	0.06
0.50	6.51	0.29	3.82	0.06

Consider the challenging choice of a multi-attribute “poverty line”. For instance, let us consider

$$S(z) \text{ vs. } q_\tau.$$

q_τ is a chosen quantile of the joint conditional CDF. $S(z)$, however, is the same S function of the $1 \times m$ vector z of poverty lines for each j attribute, $z_j \geq X_{ij}$. These thresholds ignore relations between attributes, and their different weights. Both of these choices commonly ignore the heterogeneity issues between groups and quantiles that is highlighted in Table 2 above.

To see the implication of this, consider the family of Foster-Greer-Thorbeck (FGT) scalar poverty indices. These can be calculated based on different aggregation functions and “poverty lines”. A rendition of the FGT is as follows:

$$(17) \quad P(X; z) = \frac{1}{n} \sum_{i=1}^n (S_i)^\epsilon = FGT_\epsilon.$$

Aggregation issues that are implicit in any choice of the S_i are resolved in several ways, including the information efficient method adopted in this paper, or by incorporating popular axioms. But, it can be seen that a measure for the entire population or group seemingly requires the same functional S for all the included units. Our results make clear that this is empirically unsupported.

In our approach we rely on a different definition of “poor”: the q_τ quantile of the joint conditional distribution of health and income. This corresponds to a set, S^τ , which we are able to consistently estimate in this paper, with specific reference to group characteristics such as education and ethnicity.

Note that we do not conduct SD tests based on the distribution of the estimated S functions in this paper. Such a distribution also has its own quantiles which are generally different from the two discussed above.

Our empirical findings support the notion that it is difficult to conceive of multidimensional poverty indices for the “whole” population. Different weights and substitution values apply to different subgroups, as anticipated in the discussion of Atkinson and Bourguignon above.

Our approach to presentation of the heterogeneous intra-group comparisons exposes the high level of heterogeneity in valuation functions for different groups. This makes clear that, in the multidimensional setting, statistical uniform weak rankings are more than usually called for to support robust statements about poverty. The robustness derives from the ranking over classes of functions such as S , as well as certain social welfare functions over individual units. This may be achieved by stochastic dominance tests to which we now turn.

4.4. KS Tests. It is evident that there is a very large number of KS tests that could be undertaken, each having different combinations of the conditioning covariates. To avoid overwhelming the reader we provide a set of representative results that we hope are of interest. In particular, we consider pairwise comparisons involving low and high parental education households for each ethnic group. Here we report the nonparametric tests described in Section 3.2, and the bootstrap method described there (for brevity we only report tests based on the smooth CDF estimates as non-smooth variants of the test produce qualitatively identical results).

TABLE 3. KS test, the test statistic D , P -value, and critical values of the test statistic at the 0.01, 0.05, and 0.10 level (Education low/high = (2,12)).

Group	D	P	$q_{0.01}$	$q_{0.05}$	$q_{0.10}$
Jawa	-0.0093	0.0000	-0.0036	-0.0012	0.0000
Sunda	-0.0193	0.0000	-0.0070	-0.0030	-0.0008
Betawi	0.0021	0.1391	-0.0043	-0.0012	0.0011

TABLE 4. KS test, the test statistic D , P -value, and critical values of the test statistic at the 0.01, 0.05, and 0.10 level (Education = 6).

Group	D	P	$q_{0.01}$	$q_{0.05}$	$q_{0.10}$
Betawi-Jawa	0.0072	0.6637	-0.0016	-0.0002	0.0007
Betawi-Sunda	0.0089	0.7037	-0.0026	-0.0006	0.0010
Sunda-Jawa	0.0160	0.9980	-0.0015	0.0002	0.0008

As can be seen in Table 3, within the groups Jawa and Sunda, we have clear rejection of the null of no first order stochastic dominance. Higher education levels FSD lower ones in both Jawa and Sunda. In Betawi there is no FSD as we fail to reject at all conventional

levels. But looking at Table 1 the small sample for Betawi may lead to a loss of power. Were more data available we might deduce a dominance relationship is at work for this group as well. The graph of the corresponding two CDFs (not presented here), indicates a crossing of the high and low education CDFs at the extremely high quantiles of (health, income). This supports the possibility of second order stochastic dominance. Thus any aversion to inequality in health and income would rank the higher education group above the lower one, for all increasing and concave preference functions.

As can be seen from Table 4, FSD ranking by ethnicity is not statistically supported for the education level set at 6. This is representative of all education levels. The crossing of the estimated multivariate CDFs occurs at low levels of income and/or health. This indicates poor prospects for statistical finding of second and third order ranking by ethnicity.

Many other comparisons, including results for ranking across location/ethnicity are also available from the authors.

We infer that education is the best indicator of uniform rankings within and between subgroups. Substantially extreme, ordered, valuations of income and/or health, and how they substitute for each other, would be required for joint income-health ranking by ethnicity. Referencing our discussion above on conditions for uniform ranking, weighted by group relative sizes, some or all of the intermediate conditions for stochastic dominance at the subgroup level do not seem satisfied.

5. CONCLUSIONS

In this paper we have been able to show how unrestricted nonparametric estimates of joint conditional distributions of attributes may be obtained. We have further shown how these can be utilized to obtain estimates of group valuation functions/aggregates. Our results for the cases examined indicate strong heterogeneity in weights for, and relations between attributes among different groups and for different group characteristics. This reinforces the notion that robust analysis may have to avoid index based cardinal methods, as

afforded by summary poverty and inequality indices for entire populations and groups. Our stochastic dominance tests identify some strong rankings to a degree of statistical confidence. Groups may be ordered stochastically by education and ethnicity. But quantifying the differences based on multivariate indices of well-being is challenging, since there is very strong heterogeneity among groups with respect to the weight they place on health and income, and different substitution levels between such dimensions.

REFERENCES

- Aitchison, J. & Aitken, C. G. G. (1976), 'Multivariate binary discrimination by the kernel method', *Biometrika* **63**(3), 413–420.
- Anderson, G., Leo, T. & Anand, P. (2014), Dealing with increasing dimensionality in wellbeing and poverty measurement, some problems and solutions, Technical report, University of Toronto Economics Department Mimeo.
- Atkinson, A. & Bourguignon, F. (1982), 'The comparison of multi-dimensioned distributions of economic status', *Review of Economic Studies* **49**, 183–201.
- Atkinson, A. & Bourguignon, F. (1987), Income distribution and differences in needs, in G. Feiwel, ed., 'Arrow and the Foundations of the Theory of Economic Policy', Macmillan.
- Basu, K. (1980), *Revealed preference of government*, Cambridge University Press.
- Bourguignon, F. & Chakravarty, S. R. (2003), 'The measurement of multidimensional poverty', *Journal of Economic Inequality* **1**, 25–49.
- Decancq, K. & Lugo, M. A. (2012), 'Weights in multidimensional indices of well-being: An overview', *Econometric Reviews* **32**, 7–34.
- Donald, S. & Hsu, Y. C. (2012), Improving the power of stochastic dominance tests, Technical report, University of Texas, Austin.
- Duclos, J. Y., Sahn, D. E. & Younger, S. D. (2006), 'Robust multidimensional poverty comparisons', *The Economic Journal* **116**, 943–968.
- Duclos, J. Y., Sahn, D. E. & Younger, S. D. (2010), Partial multidimensional inequality orderings, Technical report, CIRPEE Working Paper 10-03.
URL: <http://ssrn.com/abstract=1550163>
- Fleubaey, M. & Maniquet, F. (2011), *A Theory of Fairness and Social Welfare*, Cambridge University Press.
- Ginindza, M. & Maasoumi, E. (forthcoming), 'Evaluating inflation targeting based on the distribution of inflation and inflation volatility', *North American Journal of Economics and Finance* .
- Hall, P., Racine, J. S. & Li, Q. (2004), 'Cross-validation and the estimation of conditional probability densities', *Journal of the American Statistical Association* **99**(468), 1015–1026.
- Hayfield, T. & Racine, J. (2008), 'Nonparametric econometrics: The np package', *Journal of Statistical Software* **27**(5), 1–32.
- Kannai, Y. (1980), 'The ALEP definition of complementarity and least concave utility functions', *Journal of Economic Theory* **22**, 115–117.
- Kolm, S. C. (1977), 'Multidimensional egalitarianism', *Quarterly Journal of Economics* **91**, 1–13.
- Li, Q. & Racine, J. (2007), *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.
- Li, Q. & Racine, J. S. (2008), 'Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data', *Journal of Business and Economic Statistics* **26**(4), 423–434.
- Linton, O., Maasoumi, E. & Whang, Y. J. (2005), 'Consistent testing for stochastic dominance under general sampling schemes', *Review of Economic Studies* **72**, 735–765.
- Linton, O., Maasoumi, E. & Whang, Y. J. (2008), 'Corrigendum: Consistent testing for stochastic dominance under general sampling schemes', *Review of Economic Studies* **75**, 333–337.
- Maasoumi, E. (1986), 'The measurement and decomposition of multi-dimensional inequality', *Econometrica* **54**(4), 991–997.
- Maasoumi, E. (1989), 'Continuously distributed attributes and measures of multivariate inequality', *Journal of Econometrics* **42**, 131–144.
- Maasoumi, E. & Lugo, M. A. (2008), The information basis of multivariate poverty assessments, in 'Quantitative Approaches to Multidimensional Poverty Measurement', Palgrave-MacMillan.
- R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: <http://www.R-project.org/>
- Ram, R. (1982), 'Composite indices of physical quality of life, basic needs fulfillment, and income: A principal component representation', *Journal of Development Economics* **11**, 227–247.

- Ravallion, M. (2010), Mashup indices of development, Technical report, Policy Research Working Paper 5432, World Bank, Washington DC.
- Sen, A. (1970), 'Degrees of cardinality and aggregate partial orderings', *Econometrica* **43**, 393–409.

APPENDIX A. TABLES

In the following table ‘pce00’ denotes income (per capita expenditures), ‘hb.ema’ denotes health (levels of hemoglobin adjusted by gender and age), ‘hdeduc00’ denotes education (level of education of the head of the household), and ‘ethnic’ denotes group (group variable used).

TABLE 5. Bandwidth Selection Summary

Conditional density data (19602 observations, 4 variable(s))
(2 dependent variable(s), and 2 explanatory variable(s))

Bandwidth Selection Method: Maximum Likelihood Cross-Validation

Formula: pce00 + hb.ema ~ factor(ethnic) + ordered(hdeduc00)

Bandwidth Type: Fixed

Objective Function Value: 15.42470 (achieved on multistart 4)

Exp. Var. Name: factor(ethnic) Bandwidth: 0.03271572 Lambda Max: 1

Exp. Var. Name: ordered(hdeduc00) Bandwidth: 0.06460369 Lambda Max: 1

Dep. Var. Name: pce00 Bandwidth: 67676.56 Scale Factor: 2.498477

Dep. Var. Name: hb.ema Bandwidth: 1.572197 Scale Factor: 5.152807

Continuous Kernel Type (Dep. Var.): Second-Order Epanechnikov

No. Continuous Dependent Vars.: 2

Unordered Categorical Kernel Type (Exp. Var.): Li and Racine

No. Unordered Categorical Explanatory Vars.: 1

Ordered Categorical Kernel Type (Exp. Var.): Li and Racine

No. Ordered Categorical Explanatory Vars.: 1

EMORY UNIVERSITY, ATLANTA, GA USA 30322-2240 (CORRESPONDING AUTHOR),, MCMASTER UNIVERSITY, HAMILTON, ON CANADA L8S 4M4