UNITED NATIONS UNIVERSITY
UNU-WIDER

DRAFT

This paper is a draft submission to

Inequality—Measurement, trends, impacts, and policies

5–6 September 2014 Helsinki, Finland

This is a draft version of a conference paper submitted for presentation at UNU-WIDER's conference, held in Helsinki on 5–6 September 2014. This is not a formal publication of UNU-WIDER and may reflect work-in-progress.

THIS DRAFT IS NOT TO BE CITED, QUOTED OR ATTRIBUTED WITHOUT PERMISSION FROM AUTHOR(S).

Ordinal bivariate inequality: concepts and application to child deprivation in Mozambique*

Christoffer Sonne-Schmidt[†], Finn Tarp[‡], Lars Peter Østerdal[§]

July 2014

Abstract

This paper introduces a concept of inequality comparisons with ordinal bivariate categorical data. In our model, one population is more unequal than another when they have common arithmetic median outcomes and the first can be obtained from the second by correlation-increasing switches and/or median-preserving spreads. For the canonical 2x2 case (with two binary indicators), we derive a simple operational procedure for checking ordinal inequality relations in practice. As an illustration, we apply the model to childhood deprivation in Mozambique.

JEL classification: D63, I32, O15

Keywords: Qualitative data, two-dimensional first order dominance, multidimensional inequality, ordinal comparison, childhood deprivation.

*An early version of this paper was circulated under the title "Ordinal comparison of multidimensional deprivation: theory and application". We are grateful to conference participants at PET (Galway), New Directions in Welfare (Oxford), and the DASIG Policy Analytics workshop (Paris), Alain Trannoy, and Marc Fleurbaey, for very helpful comments. Financial assistance from the Danish Consultative Research Committee for Development Research (FFU) is appreciated. The usual caveats apply.

[†]The Danish National Centre for Social Research.

[‡]Department of Economics, University of Copenhagen, and UNU-WIDER.

[§]Department of Business and Economics, and Centre of Health Economics Research (COHERE), University of Southern Denmark. Correspondence to: Lars Peter Østerdal, Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. E-mail: lpro@sam.sdu.dk

1 Introduction

It is widely agreed in the literature that a multi-dimensional view of individual well-being along the lines suggested by Sen (1985, 1993) is needed when poverty, social welfare, and inequality comparisons are made. Alkire (2002), Alkire and Foster (2011), and Alkire and Santos (2011) pursue this point and helped form the Multidimensional Poverty Index (MPI), the United Nations Development Programme (UNDP) introduced in its 2010 "Human Development Report". These are welcome innovations in a challenging area of research and policy application.[1]

A persistent methodological challenge in the analysis of (multidimensional) inequality is that outcomes are often ordinal in nature; i.e. the outcomes are (partially) ranked in terms of better or worse, but there is no natural measure for the distances between them. In recent years considerable progress has been made in the development of methods based on stochastic dominance theory for comparisons of multidimensional inequality that are robust over broad classes of "utility" indices and aggregation rules across dimensions of well-being.[2] Gravel and Moyes (2006, 2011, 2012) characterize the elementary redistributive operations that reduce inequality in a two-dimensional model where one of the attributes is cardinally measurable. Their framework provides a method for making inequality comparisons between distributions with common mean for the cardinal attribute and identical marginal distributions for the ordinal attribute. Decancq (2012) and Meyer and Strulovici (2012) study ordinal interdependence concepts for comparisons of discrete multidimensional distributions that have identical marginal distributions. A closely related line of literature deals with robust comparisons of poverty/social welfare (e.g. Atkinson and Bourguignon 1982; Bourguignon 1989; Duclos et al. 2006, 2007, 2011; Gravel et al. 2009; Gravel and Mukhopadhyay 2010; Østerdal 2010; Arndt et al. 2012, among

[1] For general discussion of the case for multidimensional understanding of inequality and poverty, see Grusky and Kanbur (2006) and Sen (2006).

[2] For surveys of traditional (cardinal) multidimensional inequality measures, such as the various multidimensional generalizations of the Gini index and the Atkinson-Kolm-Sen approach, we refer to Maasoumi (1999) and Weymark (2006). See also Savaglio (2006) and Trannoy (2006) for broader discussions.

others) where the latter two papers apply an ordinal multidimensional first order dominance approach. However, ordinal inequality concepts for multidimensional distributions (not necessarily having the same marginal means or distributions) are yet to be developed.[3]

When data are ordinal in nature, use of a conventional income inequality measure, such as the Gini index, is not meaningful since it requires that outcomes are measured on a cardinal scale that reflects the relative desirability of outcomes.[4] Measures of dispersion for one-dimensional ordinal categorical data have been developed since at least the early 1990's, see, e.g., Blair and Lacy (1996, 2000). Allison and Foster (2004) put forward a simple but illuminating and intuitive model for robust comparisons of inequalities when outcomes are categorical and ordinally ranked. The Allison-Foster framework is a median-based dominance approach in which distribution x is more unequal than distribution y whenever the two distributions have common median and x is more spread out relative to the median than y. As discussed in Allison and Foster (2004), the median, rather than for instance the mean, is chosen as the reference point since the median is the natural ordinally invariant center of distribution. For empirical illustration, they provided both first order dominance comparisons and ordinal inequality comparisons of distributions of self-assessed health across states and regions of the United States, and showed that the inequality comparison concept was both meaningful and operational. A number of recent contributions have developed related inequality measures based on dispersions from the median and provided further applications of these methods, e.g. Apouey (2007), Abul Naga and Yalcin (2008), Madden (2010), Kobus and Miłoś (2012), and Dutta and Foster (2013).[5]

The aim of this paper is to introduce a median-based notion of inequality for ordinal two-dimensional categorical data, with emphasis on the case of

[3] Kobus (2012) characterizes a class of multidimensional inequality indices for ordinal data though.

[4] This means that outcomes are measured on a ratio-scale. We refer to Sen (1997) and Lambert (2001) for definitions and discussion of the Gini index.

[5] Zheng (2011) introduces a new approach to rank and measure *socioeconomic* health inequality with ordinal health data.

binary indicators. Our concept is relevant for situations where well-being is measured along two dimensions (attributes), and where only ordinal information about the desirability of outcomes is available. This means that along each dimension outcomes can be ranked according to their desirability, but nothing is assumed about the relative importance of attributes, complementarity/substitutability relationships, and the relative importance of levels within each attribute.[6] Our concept extends the Allison-Foster framework for assessing inequality of one-dimensional categorical distributions to a two-dimensional one. Roughly speaking, in our model, distribution f is more unequal than distribution g, if the two distributions have a common arithmetic median (i.e. they have a common ordinally invariant reference point) and f can be obtained from g by certain "inequality-increasing elementary transformations" in population mass relative to the reference point. Note that the arithmetic median is the vector of marginal medians (e.g. Hayford 1902; Haldane 1948; Barnett 1976). It has been described as the only reasonable multi-attribute generalization of the median concept when the attributes are different in kinds, e.g. Haldane (1948) and Barnett (1976).

As in the Allison-Foster model, when the (arithmetic) medians differ for two distributions they are incomparable inequality-wise. This tends to limit applicability in cases with many dimensions and levels where it happens less frequently that two given distributions have a shared median. Another obstacle for empirical implementation is that it is in general difficult to check if a given distribution is more unequal than another. Therefore, we focus in this paper on the 2x2 case, where common arithmetic median outcome tend to be the rule rather than the exception, and where an easily implementable procedure for detecting inequality relations between empirical subpopulation distributions can be derived.

The rest of the paper is organized as follows. In Section 2 we motivate, illustrate and provide intuition. Section 3 contains general definitions and a comparison of our approach to that of Allison and Foster (2004). Section 4 addresses the 2x2 case (i.e., the case of two binary outcome variables),

[6] For explorations of multidimensional *poverty* measurement in similar frameworks, see for instance Duclos, Sahn, and Younger (2007) and Alkire and Foster (2008).

and we develop a procedure for detecting inequality relations in practice. Briefly, testing that one distribution is more unequal than another consists of a comparison of medians, and, if a common median exists, verification of a system of inequalities which depends on the location of the median. The test requires straightforward calculations and can be carried out in a spreadsheet. In Section 5 we apply our model to two-dimensional indicators of childhood deprivation in Mozambique. Section 6 concludes.

2 An ordinal approach to bivariate inequality: illustration and intuition

Suppose a person's well-being can be measured using two 0-1 binary variables, so there are four possible outcomes. Let $(0,0)$ denote the outcome where both variables take the value 0, $(1,0)$ the outcome where the first variable takes the value 1 and the second the value 0, and so on. In the figure below arrows point to better outcomes.

$$\begin{array}{ccc} (0,0) & \rightarrow & (1,0) \\ \downarrow & \searrow & \downarrow \\ (0,1) & \rightarrow & (1,1) \end{array}$$

Outcome $(0,0)$ is the worst and $(1,1)$ is the best outcome. We assume it is unknown which of the two intermediate outcomes $(0,1)$ and $(1,0)$ is better. A population is characterized by how people are distributed among the four outcomes. This can be illustrated as follows:

f :

	0	1
0	$\frac{2}{16}$	$\frac{4}{16}$
1	$\frac{4}{16}$	$\frac{6}{16}$

where $\frac{2}{16}$ of the population has $(0,0)$, $\frac{4}{16}$ has $(0,1)$ and $(1,0)$ respectively, and $\frac{6}{16}$ has $(1,1)$. Call this distribution f, and compare with distribution g:

$$g: \quad \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & \frac{4}{16} & \frac{2}{16} \\ 1 & \frac{2}{16} & \frac{8}{16} \end{array}$$

Here g can be obtained from f by moving mass amounting to $\frac{1}{8}$ from outcome $(0,1)$ to outcome $(0,0)$ and by moving a similar amount from $(1,0)$ to $(1,1)$. In other words, g can be obtained from f by a *correlation-increasing switch* (Hamada 1974; Epstein and Tanny 1980; Tchen 1980; Boland and Proschan 1988). As argued by Atkinson and Bourguignon (1982), Tsui (1999), Atkinson (2003), Bourguignon and Chakravarty (2003), Decancq (2011) and others, such a correlation-increasing switch intuitively increases inequality. It provides a balanced movement of mass from two intermediate outcomes to the two extremes that does not change the marginal distributions but increases interdependence. If a person experiences a bad outcome in one of the dimensions of g, the conditional probability that the other outcome is also bad is higher for g than for f, so indeed it seems reasonable to say that g is more unequal than f.[7]

For one population distribution to be obtained from another by a correlation-increasing switch, it is required that the difference in mass between the two distributions for the outcome $(0,0)$ is exactly equal to the corresponding difference for the outcome $(1,1)$. Unless the populations (or number of observations) underlying the two distributions are very small this is only going to happen in exceptional cases.

[7] Note that a correlation-increasing switch may move some mass toward the center of the distribution (the median outcome). However, it will then simultanuously move an equal amount of mass away from the center of the distribution, which overall results in a more dispersed distribution.

However, let us consider a third distribution h:

$$h: \quad \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & \frac{4}{16} & \frac{2}{16} \\ 1 & \frac{3}{16} & \frac{7}{16} \end{array}$$

Obviously, h cannot be obtained from g or f by a correlation-increasing switch. But f, g and h all have the same arithmetic median in $(1,1)$, i.e. a median value of 1 in each of the two dimensions.[8] If we regard the arithmetic median as the natural center of the distributions then intuitively h is more unequal than g. Indeed, distribution h can be obtained from g by moving population mass amounting to $\frac{1}{16}$ from the median outcome $(1,1)$ to $(1,0)$: that is, h can be obtained from g by a *median-preserving spread* (Allison and Foster, 2004).

Accordingly, we will say that a distribution is *ordinally more unequal* than another if it is possible to obtain the distribution from the other through a sequence of correlation-increasing switches and/or median-preserving spreads. In our example, h is ordinally more unequal than f since there exists a distribution g, such that g can be obtained from f through a correlation-increasing switch and h can be obtained from g through a median-preserving spread.

3 General formulation

Suppose that there are 2 attributes (dimensions). An *outcome* is an two-dimensional vector $x = (x_1, x_2)$ where each x_j is defined on an attribute set $X_j = \{0, ..., n_j\}$, $j = 1, 2$. The set of outcomes is the product set $X = X_1 \times X_2$.

The statement $x \leq y$ will mean that $x_j \leq y_j$ for all j, and $x < y$ will

[8] As mentioned in the Introduction, the arithmetic median is the vector of marginal medians and the natural multi-attribute generalization of the median concept when the attributes are different in kind (e.g. Haldane, 1948; Barnett 1976). For geometric problems, like that of defining the geographical median for a population distributed on a plane (or a sphere), a number of other median concepts exists, see Small (1990).

mean that $x_j \leq y_j$ for all j and $x \neq y$. A *distribution* is a real-valued function f on X with $\sum_{x \in X} f(x) = 1$ and $f(x) \geq 0$ for all $x \in X$. Let f_j denote the marginal distribution on X_j.

Let $m_j(f_j)$ denote the median of f_j on X_j.[9] The (arithmetic) median of f is the vector $m(f) = (m_1(f_1), m_2(f_2))$, of coordinate-wise medians.

We say that distribution g can be derived from distribution f by a *bilateral transfer* (of mass between two outcomes), if there are outcomes x, y and a non-negative scalar ε such that $g(x) = f(x) - \varepsilon$, $g(y) = f(y) + \varepsilon$ and $g(z) = f(z)$ otherwise. If $y < x$ the bilateral transfer is *diminishing* (i.e. moves mass from a better to a worse outcome), if for some outcome z that $y < x \leq z$ or $z \leq x < y$ it is *directed away* from z, and if $m(f) = m(g)$ it is *median-preserving.* A median-preserving bilateral transfer directed away from the median is a *median-preserving spread* (see Section 2 for an example).

We say that g is derived from f by a *correlation-increasing switch* if we can choose outcomes x, y, v, w such that $v = (\{\min\{x_1, y_1\}, \min\{x_2, y_2\}\})$ and $w = (\max\{x_1, x_2\}, \max\{x_2, y_2\})$, $f(x) - g(x) = f(y) - g(y) > 0, f(v) - g(v) = f(w) - g(w) < 0$, and $g(z) = f(z)$ otherwise (again, see Section 2 for an example).

In the following, we define an *inequality-increasing elementary transformation* to be a correlation-increasing switch or a median-preserving spread.

If g can be derived from f by a finite sequence of inequality-increasing elementary transformations, we say that g is *ordinally more unequal* than f, or, as an equivalent statement, f is *ordinally more equal* than g. Formally, g is ordinally more unequal than f if there are distributions $f = f^1, f^2, ..., f^k = g$, where f^{i+1} is obtained from f^i by an inequality-increasing elementary transformation, $i = 1, ..., k-1$. Note that the relation 'ordinally more unequal' is a partial order (i.e. reflexive, antisymmetric, and transitive).[10]

[9] To ensure a unique median value, we will define $m_j(f_j)$ as the smallest element x_j in X_j such that $\sum_{i=1}^{x_j} f_j(i) \geq \frac{1}{2}$.

[10] A distribution with 50% mass at one extreme outcome and 50% mass at the other extreme outcome is the unique maximal element with respect to this relation (i.e. no other distribution is more unequal). A distribution having all mass concentrated on one outcome is clearly a minimal element (i.e. no other distribution is more equal). However,

As illustrated by Allison and Foster (2004), it is often of interest to complement (ordinal) comparisons of inequality with (ordinal) comparisons of *social welfare,* see also Zheng (2008). In our ordinal framework, the natural criterion for comparison of social welfare is first order dominance.[11] A population distribution f first order dominates population distribution g whenever g can be obtained from f by iteratively moving population mass from better to worse outcomes, i.e., if there are distributions $f = f^1, f^2, ..., f^k = g$, where f^{i+1} is obtained from f^i by a diminishing bilateral transfer $i = 1, ..., k-1$. Equivalently, any additive non-decreasing social welfare function would give as least as much social welfare to f than to g.[12]

Before proceeding, we compare these definitions and concepts with the one-dimensional case put forward by Allison and Foster (2004). With $X = X_1 = \{1, ..., n_1\}$, $f = f_1$ and $g = g_1$, define $F(k) = \sum_{j=1,..,k} f(j)$ and G in a similar way. Allison and Foster (2004) say that g has a *greater spread* than f whenever $m(g) = m(f)$ and $G(k) \geq F(k)$ for all $k < m(f)$ and $G(k) \leq F(k)$ for all $k \geq m(f)$. For the one-dimensional case, g has greater spread than f precisely if g is ordinally more unequal than f (as defined here). Also, the general definition of first order dominance given here is equivalent to the standard definition in the one-dimensional case. Thus, the definitions presented here generalize those of Allison and Foster's one-dimensional case.

4 Implementation of the 2x2 case

A central question is how to test if one distribution is ordinally more unequal than another (i.e. has greater spread in an ordinally meaningful sense). For two *one-dimensional* distributions f and g, such testing is a straightforward

for $N \geq 2$ it is possible to find minimal elements with mass at more than one outcome.

[11] First order dominance is also known as the *usual (stochastic) order*. For general references on stochastic ordering theory, see Müller and Stoyan (2002) or Shaked and Shanthikumar (2007).

[12] It is worth mentioning that in the multidimensional context the term "first order dominance" has been used with other meanings in the economics literature. In particular, Atkinson and Bourguignon (1982) and subsequent literature have used this term for a less restrictive stochastic dominance concept which corresponds to additional restrictions on the social welfare function (also known as an *orthant stochastic order* in the stochastic orderings literature).

matter of checking whether n_1 inequalities hold.[13] For the multidimensional case (even the two-dimensional), checking if one distribution is more unequal than another is more complicated. We focus in our empirical implementation on the 2x2 case which can be dealt with in a tractable manner.

In this section, we assume that an outcome is a vector $x = (x_1, x_2)$ where each attribute x_j is defined on an attribute set $X_j = \{0, 1\}$, $j = 1, 2$. Thus, the outcome set is $X = \{0, 1\} \times \{0, 1\}$. For an outcome $x = (x_1, x_2)$ we use the notation $f(x_1, x_2)$ for $f(x)$.

4.1 Checking first order dominance relations

Let f and g denote distributions on X. By application of Strassen's Theorem (Strassen 1965), it follows that f first order dominates g if and only if the cumulative probability mass at f is smaller than or equal to that at g for every *lower comprehensive subset* of outcomes. A lower comprehensive subset $Y \subseteq X$ holds the property that if an outcome is in the subset, then all smaller outcomes are also included in that subset. That is, if $x \in Y$, $y \in X$ and $y \leq x$ then $y \in Y$.

Thus, in the 2x2 case, f first order dominates g if and only if the following four inequalities are satisfied: $g(0,0) \geq f(0,0), g(0,0) + g(0,1) \geq f(0,0) + f(0,1), g(0,0) + g(1,0) \geq f(0,0) + f(1,0)$, and $g(0,0) + g(1,0) + g(0,1) \geq f(0,0) + f(1,0) + f(0,1)$.[14]

4.2 Checking ordinal inequality relations

We proceed next to present necessary and sufficient conditions for f being ordinally more equal than g, as defined in Section 2.

Correlation-increasing switches are median-preserving, so a necessary condition g to be ordinally more unequal than f is that the two distributions

[13] See Allison and Foster (2004) for a detailed discussion of how this test can be nicely visualized.

[14] Note that the Atkinson and Bourguignon (1982) definition of first order dominance requires only that there is at least as much mass at g than at f for all lower rectangular sets of outcomes. Thus, for the 2x2 case, it is not required that $g(0,0)+g(1,0)+g(0,1) \geq f(0,0)+f(1,0)+f(0,1)$. In general, the concepts differ greatly in the number of inequality restrictions that are imposed.

have common median.[15] We can therefore rely on considering in turn each of four possible cases of common median, and proceed as described below.

Proposition 1 *(Ordinal inequality check for the 2x2 case) Let* $X = \{0,1\} \times \{0,1\}$ *and let* f *and* g *be two distributions on* X*. Then* g *is ordinally more unequal than* f *if and only if one of the following six cases holds:*

A1. f *and* g *have common median* $(1,1)$*, and* f *first order dominates* g*.*

A2. f *and* g *have common median* $(0,0)$*, and* g *first order dominates* f*.*

B1. f *and* g *have common median* $(1,1)$*, and* $f(1,0) \geq g(1,0)$*,* $f(0,1) \geq g(0,1)$*,* $f(1,1) \leq g(1,1)$*,* $g(1,1) - f(1,1) \leq \min\{f(1,0) - g(1,0), f(0,1) - g(0,1)\}$*.*

B2. f *and* g *have common median* $(0,0)$*, and* $f(1,0) \geq g(1,0)$*,* $f(0,1) \geq g(0,1)$*,* $f(0,0) \leq g(0,0), g(0,0) - f(0,0) \leq \min\{f(1,0) - g(1,0), f(0,1) - g(0,1)\}$*.*

C1. f *and* g *have common median* $(1,0)$*, and* $g(1,0) \leq f(1,0)$*,* $g(0,1) \leq f(0,1)$*,* $g(1,1) \geq f(1,1)$*,* $g(0,0) \geq f(0,0)$*,* $f(1,0) - g(1,0) \geq f(0,1) - g(0,1)$*.*

C2. f *and* g *have common median* $(0,1)$*, and* $g(0,1) \leq f(0,1)$*,* $g(1,0) \leq f(1,0)$*,* $g(1,1) \geq f(1,1)$*,* $g(0,0) \geq f(0,0)$*,* $f(0,1) - g(0,1) \geq f(1,0) - g(1,0)$*.*

The proof of Proposition 1 is given in Appendix A. The intuition behind the conditions is discussed below.

The cases A1 and A2 are symmetric so we will only discuss A1. As mentioned in Section 4.1, f first order dominates g if and only if it is possible to go from f to g by a finite sequence of diminishing bilateral transfers. Each such bilateral transfer is a median-preserving spread (as shown formally in the Appendix) and thereby an inequality-increasing elementary transformation.

The cases B1 and B2 are symmetric so we will only discuss B1. To provide some intuition for the inequalities in B1, suppose that f does not first order dominate g, and g does not first order dominate f. Then, if g

[15] Suppose that g is derived from f by some correlation-increasing switch. For the case $X = \{0,1\} \times \{0,1\}$, any correlation-increasing switch can be conducted by means of two bilateral transfers (of the same amount of mass) from $(0,1)$ and $(1,0)$ to the extreme outcomes $(0,0)$ and $(1,1)$.

is ordinally more unequal than f, it is impossible to go from f to g via a finite sequence of inequality-increasing elementary transformations without making use of at least one correlation-increasing switch (because we would then have first order dominance since the median is an extreme outcome $(1,1)$). Thus, if g is ordinally more unequal than f, then $f(1,0) > g(1,0)$ and $f(0,1) > g(0,1)$, since otherwise it would be possible to go from f to g without any correlation-increasing switches (because if a correlation-increasing switch is involved, one of the intermediate outcome would receive at least as much probability mass from $(1,1)$ as is moved to $(1,1)$ and hence only diminishing bilateral transfers are needed, a contradiction). However, these two conditions are not sufficient for g being ordinally more unequal than f. Roughly speaking, we need a condition ensuring that all mass transferred to $(1,1)$ in the process of moving from f to g can be transferred from the intermediate outcomes $(0,1)$ or $(1,0)$ in connection with a correlation-increasing switch. This is precisely the condition $g(1,1) - f(1,1) \leq \min\{f(1,0) - g(1,0), f(0,1) - g(0,1)\}$.

The cases C1 and C2 are symmetric so we will only discuss C1.[16] The first inequalities ensure that f has at least as much mass at the intermediate outcomes, and not more mass than at the extreme outcomes, than g. The last condition ensures that the difference in mass at the median outcome $(1,0)$ is no less than the difference in mass at the other intermediate outcome $(0,1)$. As verified formally in the Appendix, the conditions imply (and are implied by) that g can be obtained from f by a correlation-increasing switch and bilateral transfers of mass from $(1,0)$ to $(0,0)$ and $(1,1)$ respectively.

The following illustrates how a concrete data set can be analyzed in the present framework. For illustrative purposes, we highlight examples of all the basic types of ordinal inequality relations that can occur in the 2x2 case

[16] It is possible to show that A1 or B1 holds if and only if the distributions have common median in $(1,1)$ and f dominates g according to the first order dominance concept of Atkinson and Bourguignon (1982). Although the foundations of the Atkinson-Bourguignon concept of first order dominance in terms of elementary operations has never been fully established (see Moyes 2012, Footnote 13), it shows that our concept and the operation involved is closely linked with the Atkinson-Bourguignon concept of first order dominance for cases where the median takes an extreme value.

(see Section 5.3).

5 Empirical illustration

In Mozambique, investment in schooling, health, and sanitation has increased the level of human capital and indices of human development. While this development has influenced living standards of both adults and children, its impact on children is of particular interest. The acquisition of human capital in early childhood is imperative for future learning, earnings and health status (UNICEF 2006). Large gaps in basic welfare goods during childhood tend to persist, if not widen, the variation in human capital, productivity and living standards throughout adulthood, see Strauss and Thomas (1995), and Orazem and King (2007).

To address the above challenges voucher or cash transfer programmes targeted at disadvantaged children have in recent year become more common.[17] A general problem with such government transfer programmes is to make sure that transfers are directed at the most disadvantaged children. Efficient targeting of government resources require that administrators can detect the most vulnerable groups. We illustrate how our model can be used for examining inequalities within and between groups of Mozambican children, concentrating on three key characteristics, rural-urban area of residence, gender of head of household, and gender of the child.[18] This results in a total of eight categories of children that we compare with each other.

[17] The most famous of these initiatives is probably Mexico's PROGRESA/Oportunidades programme, which aims at increasing children's school attendance among poor families, by awarding grants to mothers conditional on school enrolment. See Parker, Rubalcava and Teruel (2007) for further discussion and examples.

[18] Urban-rural area of residence is likely to have a significant impact on living standards mainly due to the low population density of rural areas, which makes supply of high quality public services more costly. Children living in female headed households are more likely than other children to fall below the poverty line primarily because women's wages and education tend to be lower than men's. Buvinić and Gupta (1997) review literature relating female headship and poverty. However, as Handa (1996) observes, female headed households are also likely to spend a larger share of their income on improving children's human capital. Finally, households may discriminate based on gender of the child. For example in Mozambique, it is not uncommon for especially rural families to invest more in the education of boys as compared to girls (UNICEF, 2006).

5.1 Data and summary statistics

We apply the model to the Mozambican Demographics and Health Survey from 2003 (DHS 2003).[19] This is a nationally representative data set that includes detailed information on childhood poverty. We focus on three indicators for severe deprivation in sanitation, health, and education respectively (cf. Gordon et al., 2003). Sanitation deprivation indicates lack of access to a toilet of any kind, including communal toilets or latrines. Health deprivation is an indicator for pre-school-aged children (under five years) who have never been immunized or who have recently been ill with diarrhea but did not receive medical attention. Education deprivation is an indicator for school-aged children (between seven and eighteen years) who have never been to school. We combine these into two 2x2 indicators of childhood poverty for school-aged and pre-school-aged children respectively. A detailed description of the survey is given in UNICEF (2006).[20]

Table 1 summarizes how indicators of childhood poverty are distributed among the four possible outcomes. The top panel lists the distribution of sanitation and education (for school-aged children), and the lower panel lists the distribution of sanitation and health (for pre-school-aged children), each by area of residence, gender of head of household, and gender of the child. For example, the first row of the lower panel shows that among pre-school-aged girls in rural, male-headed households 18.8% live with poor sanitation and under poor health conditions, 44.4% have poor sanitation but adequate health, 4.8% have good sanitation but poor health, and the rest, 32%, have both good sanitation and good health conditions.[21]

[19] Recently, Arndt et al. (2012) provide an alternative implementation of the multidimensional first order dominance approach with an application to comparisons of child poverty in Vietnam and Mozambique between groups and over time.

[20] Lindelow (2006) studies socioeconomic health inequalities in Mozambique using the concentration index. His study is based on income and health data from the 1996-1997 household survey.

[21] We have weighted these shares by survey sample weights.

Area, Gender of household head, Gender of child	*(0,0)*	*(0,1)*	*(1,0)*	*(1,1)*	*Median*	*# of obs.*
(Sanitation deprivation, Education deprivation)						
Rural, Male, Girl	27.7	34.5	10.3	27.6	*(0,1)*	3716
Rural, Male, Boy	16.3	41.3	9.4	33.0	*(0,1)*	4010
Rural, Female, Girl	21.6	38.4	8.7	31.2	*(0,1)*	1223
Rural, Female, Boy	19.2	41.0	8.1	31.7	*(0,1)*	1348
Urban, Male, Girl	6.0	9.9	7.2	76.9	*(1,1)*	2858
Urban, Male, Boy	5.0	13.1	5.3	76.6	*(1,1)*	2912
Urban, Female, Girl	8.2	9.0	5.3	77.5	*(1,1)*	1140
Urban, Female, Boy	7.2	11.2	4.2	77.4	*(1,1)*	1025
(Sanitation deprivation, Health deprivation)						
Rural, Male, Girl	18.8	44.4	4.8	32.0	*(0,1)*	2262
Rural, Male, Boy	19.2	44.8	4.7	31.3	*(0,1)*	2288
Rural, Female, Girl	13.9	47.9	4.7	33.6	*(0,1)*	580
Rural, Female, Boy	15.7	44.9	3.7	35.6	*(0,1)*	598
Urban, Male, Girl	2.6	18.8	7.6	71.0	*(1,1)*	1215
Urban, Male, Boy	2.9	18.2	8.1	70.9	*(1,1)*	1156
Urban, Female, Girl	2.0	19.5	3.2	75.3	*(1,1)*	382
Urban, Female, Boy	3.7	16.7	8.5	71.1	*(1,1)*	341

Note: The first element, i, in vector (i,j) indicates sanitation deprivation. The second element, j, indicates education deprivation in the top panel and health deprivation in the bottom panel. $i,j = 0$ is deprivation, $i,j = 1$ is no deprivation.
Source: Authors' calculations from DHS 2003.

Table 1: Percentages of children's two-dimensional living standards.

5.2 Results from pairwise comparisons

Table 2 shows first order dominance and ordinal inequality relations between all distributions (within each panel) in Table 1. A number 1 in an entry indicates that the row distribution first order dominates the column distribution, while 0 indicates no first order dominance. Capital letters indicate ordinal inequality relations of the types specified in Proposition 1. An asterisk means that a bootstrapping test indicates that the groups are significantly different (for more details on this bootstrap procedure see Appendix A.2). Note that first order dominance might be compatible with ordinal inequality relations (of the type A), yet these cases do not occur in the data shown, while ordinal inequality relations of type B and C are present. We highlight these cases below.

Type B inequalities are those with extreme common medians, in $(0, 0)$ or $(1, 1)$, but where none of the distributions first order dominates the other. For illustration, compare the distribution for urban boys in female-headed households (last row in upper panel of Table 1) with the distribution for urban boys in male-headed households (third to last row in upper panel of Table 1). None is first order dominating the other, but the latter is more equal. To see this, starting with the distribution for urban boys in male-headed households, use a correlation-increasing switch of 1.1 and then a median-preserving spread of 0.3 and 0.8 from $(1, 1)$ and $(0, 1)$ to $(0, 0)$. This results in the distribution for urban boys in female-headed households .

Type C inequalities are those where the median is non-extreme and where there is no first order dominance. An illustration of type C ordinal inequality can be seen from comparing the distribution for girls in rural female-headed households (third row in lower panel of Table 1) to the distribution of boys in similar households (fourth row in lower panel of Table 1). Here, the girls are more equally distributed than the boys. To see this, starting with the distribution for the girls, apply first a correlation-increasing transfer of 0.9 and then a median-preserving spread of 0.9 and 1.1 from $(0, 1)$ to $(0, 0)$ and $(1, 1)$, which gives the distribution for boys. Note that because of rounding, the numbers do not match exactly.

(Sanitation deprivation, Education deprivation)

Area, Gender of household head, Gender of child	*Median*	Rural, Male, Girl	Rural, Male, Boy	Rural, Female, Girl	Rural, Female, Boy	Urban, Male, Girl	Urban, Male, Boy	Urban, Female, Girl	Urban, Female, Boy
Rural, Male, Girl	*(0,1)*		0	0	0	0	0	0	0
Rural, Male, Boy	*(0,1)*	1*		1*	1*	0	0	0	0
Rural, Female, Girl	*(0,1)*	1*	0		0	0	0	0	0
Rural, Female, Boy	*(0,1)*	1*	0	0		0	0	0	0
Urban, Male, Girl	*(1,1)*	1*	1*	1*	1*		0	0B*	0
Urban, Male, Boy	*(1,1)*	1*	1*	1*	1*	0		0	0B*
Urban, Female, Girl	*(1,1)*	1*	1*	1*	1*	0	0		0
Urban, Female, Boy	*(1,1)*	1*	1*	1*	1*	0	0	0	

(Sanitation deprivation, Health deprivation)

Rural, Male, Girl	*(0,1)*		1	0	0	0	0	0	0
Rural, Male, Boy	*(0,1)*	0		0	0	0	0	0	0
Rural, Female, Girl	*(0,1)*	1	1*		0C*	0	0	0	0
Rural, Female, Boy	*(0,1)*	1*	1*	0		0	0	0	0
Urban, Male, Girl	*(1,1)*	1*	1*	1*	1*		0	0	0
Urban, Male, Boy	*(1,1)*	1*	1*	1*	1*	0		0	0
Urban, Female, Girl	*(1,1)*	1*	1*	1*	1*	0	0		0
Urban, Female, Boy	*(1,1)*	1*	1*	1*	1*	0	0	0	

Note: The number 1 indicates that the row distribution first order dominates the column distribution. The letters B and C indicate that the row distribution is ordinally more equal of type B or C respectively cf. Proposition 1. We conducted tests of significance for first order dominance and ordinal inequality by using the permutation bootstrap method. * indicates a significant test statistic at the 5% level.
Source: Authors' calculations from DHS 2003.

Table 2: First order dominance and ordinal inequality relations among groups of children

For the eight groups of pre-school-aged children, there are 21 first order dominances and 2 ordinal inequality relations among the 28 different pairs of groups compared. For school-aged children, we observe 21 first order dominances and 1 ordinal inequality relation. From Table 2 it emerges that urban groups are better off than rural groups. This is not so surprising. However, it also emerges that there are more first order dominances between rural groups than between urban groups, indicating more between-group inequality in the rural areas than in the urban areas. In particular, school-age boys in rural male headed households are better off than any other rural group. Moreover, there is more within-group inequality for school-age children in urban female headed households than in the corresponding male

headed households. These findings deserve attention in policy debates.

6 Conclusion

In this paper we have developed an ordinal concept of multi-dimensional inequality, building on the Allison and Foster (2004) framework for comparing inequalities with one-dimensional categorical data. To illustrate how our model can be applied in the 2x2 case we compared poverty distributions of pre-school- and school-aged children from the DHS data in Mozambique. Such data is available for a large number of countries across the developing world, meaning that potentially interesting comparisons are possible.

For these indicators, we find that first order dominance occur relatively frequently while ordinal bivariate inequality relations are less frequent. Whether this is because ordinal inequality relations generally are "rare" empirically or whether it is due to the chosen indicators of child poverty cannot be established with the data in hand. However, the example shows that while instances of ordinal bivariate inequality relations may be relatively uncommon, they do exist empirically. Moreover, our indicators of sanitation, health and education by area of residence, gender of the household health, and gender of the child provide insights into how targeting of for example cash transfer programmes presently under consideration by the Mozambican government should be pursued.

In sum, we have shown that it is possible to develop a meaningful and intuitive concept of ordinal bivariate inequality. We have also demonstrated how it can be applied in the 2x2 case. Future research will be required to explore how to deal with variations of the concept and more general cases. In particular, an important generalization would be to provide an ordinal inequality check procedure that applies to general bivariate problems. Providing such a general procedure will however not be straightforward since the many possibilites of combining correlation-increasing switches and median-preserving spreads in various sequences will be to complex to analyze directly as in the proof of Proposition 1, and thus a deeper understanding of what can be obtained from these inequality-increasing elementary operations is

needed. It is also possible to generalize the definitions and concepts to an arbitrary finite number of dimensions, although the correlation-increasing switch concept does not generalize in a straightforward manner to more than two dimensions (Decancq 2012), and checking inequality would be even more challenging. Finally, the restriction that f and g have common medians for ordinal inequality relations to be viable could possible be relaxed. See Abul Naga and Yalcin (2010) for an exploration along these lines for the one-dimensional case.

A Appendix

A.1 Proof of Proposition 1

We will make use of the following lemma which applies to the general case.

Lemma A *Suppose that g is obtained from f by a sequence of bilateral transfers directed away from $m(f)$ and $m(f) = m(g)$. Then each of these bilateral transfers is median-preserving (i.e. is a median-preserving spread).*

Proof of Lemma A: Define the sets $L(m(f)) = \{x \in X | x \leq m(f)\}$ and $U(m(f)) = \{x \in X | m(f) \leq x\}$. Then g is obtained from f by a sequence of bilateral transfers of the following four kinds: from $m(f)$ to outcomes in the sets $L(m(f))\backslash\{m(f)\}$ and $U(m(f))\backslash\{m(f)\}$ respectively, and bilateral transfers directed away from $m(f)$ within the sets $L(m(f))\backslash\{m(f)\}$ and $U(m(f))\backslash\{m(f)\}$ respectively.

Order the bilateral transfers in the sequence with the numbers $1, 2, ...$ etc., such that we first have all the bilateral transfers directed away from $m(f)$ within $L(m(f))$ and second all the bilateral transfers directed away from $m(f)$ within $U(m(f))$.

Suppose that the bilateral transfers $1, 2, ..., h-1$ are median-preserving, but bilateral transfer h fails to be median-preserving. Let $\overline{m}$ denote the new median following bilateral transfer h.

If bilateral transfer h is within $L(m(f))$ we have $\overline{m} < m(f)$, since each bilateral transfers in $L(m(f))$ is diminishing. In particular, for the median $\widetilde{m}$

resulting after the last bilateral transfer within $L(m(f))$ we have $\widetilde{m} \leq \overline{m} < m(f)$. However, the remaining bilateral transfers cannot move the median back to $m(f)$ since they are all within $U(m(f))$, contradicting $m(f) = m(g)$.

If bilateral transfer h is within $U(m(f))$, for the new median $\overline{m}$ we have $m(f) < \overline{m}$, since each bilateral transfer in $U(m(f))$ is the reverse of a diminishing transfer (i.e. moving mass from worse to better outcomes). Hence, for the median $\widetilde{m}$ resulting after the last bilateral transfer within $U(m(f))$ we have $m(f) < \overline{m} \leq \widetilde{m}$, contradicting $m(f) = m(g)$. □

We are now ready to prove Proposition 1. As mentioned prior to the statement of Proposition 1, a shared median is a necessary condition for one distribution to be ordinally more unequal than another. We proceed by showing that for each case of common median, the relevant sets of inequalities stated in Proposition 1 are indeed necessary and sufficient for an ordinal inequality relation to hold. We focus on the case $m(f) = m(g) = (1, 1)$ (Case 1) and the case $m(f) = m(g) = (1, 0)$ (Case 2). The case $m(f) = m(g) = (0, 0)$ is symmetric to Case 1 and the case $m(f) = m(g) = (0, 1)$ is symmetric to Case 2.

In the proof of Proposition 1, a *pseudo-distribution* is defined as a real-valued function f on X with $\sum_{x \in X} f(x) = 1$. A pseudo-distribution f is thus a distribution if $f(x) \geq 0$ for all $x \in X$.

Case 1: $m(f) = m(g) = (1, 1)$.

For the inequalities in case A1 in Proposition 1 recall that f first order dominates g if and only if it is possible to go from f to g by a finite sequence of diminishing bilateral transfers. By Lemma A, each such bilateral transfer is a median-preserving spread and thereby an inequality-increasing elementary transformation. Thus, A1 implies that g is ordinally more unequal than f. Conversely, if g is ordinally more unequal than f and we can go from f to g by a finite sequence of diminishing bilateral transfers using no correlation-increasing switches then A1 is satisfied.

Now, we claim that if g is ordinally more unequal than f and if g is not first order dominated by f then it is possible to obtain g from f from a

sequence of inequality-increasing elementary transformations that involves only a single correlation-increasing switch and no bilateral transfers from the outcome $(1,1)$ to other outcomes.

We first verify the last part of the claim, i.e. we show that no bilateral transfers from $(1,1)$ to other outcomes are required. For this, consider a given sequence of inequality-increasing switches (leading from f to g) that contains a bilateral transfer of the amount β from $(1,1)$ to another outcome z. We assume, without loss of generality, that $z = (0,1)$. (If $z = (0,0)$ then we can split the bilateral transfer up into two nested bilateral transfers, one from $(1,1)$ to $(0,1)$ and one from $(0,1)$ to $(0,0)$; the case $z = (1,0)$ is symmetric to the one treated here and hence can be omitted).

As noted earlier, we know that the sequence contains at least one correlation-increasing switch (since if otherwise f would first order dominate g). Now, pick an arbitrary correlation-increasing switch from the sequence, and let α denote the amount of mass moved from each of the outcomes $(0,1)$ and $(1,0)$ to $(0,0)$ and $(1,1)$ respectively. We can then decompose this correlation-increasing switch into two bilateral transfers: a bilateral transfer of the amount α from $(0,1)$ to $(1,1)$ and a bilateral transfer of the amount α from $(1,0)$ to $(0,0)$. We consider two cases: (a) $\alpha \geq \beta$, and (b) $\alpha < \beta$.

(a) Replace the bilateral transfer from $(1,1)$ to $(0,1)$ of the amount β with a bilateral transfer of the amount β from $(1,0)$ to $(0,0)$, and reduce the amount of mass transferred between each pair of outcomes from α to $\alpha - \beta$. Note that the amount of mass eventually allocated to each outcome remains the same.

(b) Replace the correlation-increasing switch (which moves the amount α between each pair of outcomes) with a bilateral transfer of the amount α from $(1,0)$ to $(0,0)$, and reduce the size of the bilateral transfer from $(1,1)$ to $(0,1)$ to $\beta - \alpha$. Again, note that the amount of mass eventually allocated to each outcome remains the same.

Proceeding in this way until no bilateral transfers from $(1,1)$ to other outcomes remain, we can eliminate all bilateral transfers from $(1,1)$ to other outcomes. Note that we have not shown (and it is not needed for our argument) that after each elimination of some bilateral transfer from $(1,1)$

to another outcome, the resulting sequence of pseudo-distributions consists entirely of distributions. It is sufficient to observe that when all bilateral transfers from $(1,1)$ to other outcomes have been eliminated, what remains is a sequence of correlation-increasing switches and/or bilateral transfers from $(0,1)$ and/or $(1,0)$ to $(0,0)$. For this sequence, it is clear that each intermediate pseudo-distribution is a distribution. Moreover, the transformations (i.e. correlation-increasing switches and/or bilateral transfers from $(0,1)$ to $(0,0)$ and from $(1,0)$ to $(0,0)$) can be arranged in an arbitrary order and we can obtain g from f by a single operation of each type. This proves our claim.

From these observations we get the following: Suppose that f does not first order dominate g. Then g is ordinally more unequal than f if and only if the following 3 inequalities are satisfied: $f(1,0) - g(1,0) \geq 0, f(0,1) - g(0,1) \geq 0$ and $g(1,1) - f(1,1) \leq \min\{f(1,0) - g(1,0), f(0,1) - g(0,1)\}$. Note that in conjunction with the assumption that f does not first order dominate g, the 3 inequalities imply that $g(0,0) > f(0,0)$ (i.e. with strict inequality). From this observation it follows that the 3 inequalities are both necessary and sufficient: The three inequalities are necessary, since if one of them were violated, clearly we could not get g from f by a single correlation-increasing switch and/or bilateral transfers from $(0,1)$ and/or $(1,0)$ to $(0,0)$. To verify that the conditions are sufficient, we give the following constructive argument: Suppose that the conditions are satisfied. Let $\alpha = g(1,1) - f(1,1)$. Given f, let $\widehat{f}$ be the distribution obtained from a correlation-increasing switch of the amount α (where α is transferred from $(0,1)$ to $(0,0)$ and α is transferred from $(1,0)$ to $(1,1)$). Thus, $\widehat{f}(1,1) = g(1,1)$, $\widehat{f}(0,1) \geq g(0,1), \widehat{f}(1,0) \geq g(1,0)$. This means that g can be obtained from $\widehat{f}$ by diminishing bilateral transfers from $(0,1)$ and/or $(0,1)$ to $(0,0)$, and we are done.

Case 2: $m(f) = m(g) = (1,0)$.

Note that if $m(f) = m(g) = (1,0)$ and if g is ordinally more unequal than f, then g can be obtained from f by a finite number of correlation-increasing switches (from $(1,0)$ and $(0,1)$ to $(1,1)$ and $(0,0)$) and bilateral

transfers from $(1,0)$ to the extreme outcomes $(1,1)$ and $(0,0)$. Regardless of how these correlation-increasing switches and bilateral transfers are ordered, each intermediate pseudo-distribution is a distribution. Thus, a single correlation-increasing switch is enough (since all correlation-increasing switches can be amalgamated into a single correlation-increasing switch and still each intermediate pseudo-distribution is a distribution). In particular, we can obtain g from f in three steps, ordered as follows: (1) A correlation-increasing switch, (2) A bilateral transfer from $(1,0)$ to $(0,0)$, and (3) A bilateral transfer from $(1,0)$ to $(1,1)$.

From these observations, we can show that g is ordinally more unequal than f if and only if the 5 inequalities of case C1 hold: $g(1,0) \leq f(1,0)$, $g(0,1) \leq f(0,1)$, $g(1,1) \geq f(1,1)$, $g(0,0) \geq f(0,0)$, and $f(1,0) - g(1,0) \geq f(0,1) - g(0,1)$.

Necessity of the first four inequalities is straightforward. The fifth inequality $f(1,0) - g(1,0) \geq f(0,1) - g(0,1)$ must hold since the only way that mass can be transferred away from $(0,1)$ is by means of a correlation-increasing switch and thus at least the same amount is going to be transferred away from $(1,0)$. For sufficiency, we give the following constructive argument: Suppose that the 5 inequalities hold. Let $\lambda = f(0,1) - g(0,1)$. Define the distribution $\widehat{f}$ by $\widehat{f}(0,1) = f(0,1) - \lambda$, $\widehat{f}(1,0) = f(1,0) - \lambda$, $\widehat{f}(0,0) = f(0,0) + \lambda$, and $\widehat{f}(1,1) = f(1,1) + \lambda$. Thus, $\widehat{f}$ is obtained from f by a correlation-increasing switch. Then $\widehat{f}(0,1) = g(0,1)$, $\widehat{f}(0,0) \geq g(0,0)$, $\widehat{f}(0,0) \geq g(0,0)$ and $\widehat{f}(1,0) \geq g(1,0)$. Thus, g can be obtained from $\widehat{f}$ by bilateral transfers from $(1,0)$ to $(0,0)$ and/or $(1,1)$ and we are done.

A.2 Bootstrapping

Data is a sample of a larger population, so testing whether two sample groups are genuinely distinct, accounting for sample uncertainty, is of interest. We employ a bootstrap procedure which can be interpreted as testing a null-hypothesis of equality of two distributions against an alternative hypothesis of arbitrary distributions.

Define $D(f,g) = \min\{g(0,0) - f(0,0),\ g(0,0) + g(0,1) - f(0,0) - f(0,1),$

$g(0,0)+g(1,0)-f(0,0)-f(1,0)$, $g(0,0)+g(1,0)+g(0,1)-f(0,0)-f(1,0)-f(0,1)\}$. Then, $D(f,g) \geq 0$ if and only if f first order dominates g. The function D is useful for testing statistical significance of first order dominance relations. Following common convention, the null-distribution is generated by merging the observations from the two groups. From the null-distribution, two new samples are generated (drawing randomly with replacement) corresponding in size to the original two samples, and the test statistic D is calculated. Repeating this procedure 1000 times, we obtain a distribution over the test statistic consistent with the null-hypothesis, which we then compare with the test statistic of the original sample.[22] Asterisks in Table 2 indicate significance at the five percent level, meaning that the observed value of D is larger than the 95th percentile of its bootstrapped distribution (indicating that the two groups are genuinely distinct).[23] In the case of ordinal inequality (without the presence of first order dominance) as the bootstrapped test statistic we use the minimum function over the appropriate differences induced by the inequalities as specified in type A, B, or C in Proposition 1.

An alternative null-hypothesis, discussed by Howes (1993), Kaur et al. (1994) and Dardanoni and Forcina (1999) and more recently by Davidson and Duclos (2006) in the context of one-dimensional dominance of first and higher order, is non-dominance (including exact equality of distributions) against the alternative of strict dominance. This means that first order dominance is rejected unless there is strong evidence in its favor. A similar kind of test could be envisioned for the ordinal inequality relations. In order to perform such tests in a multidimensional framework, we would have to determine a "least favorable case"; that is, a null-distribution consistent with the null-hypothesis that makes the observed distributions as plausible

[22] We refer to Efron and Tibshirani (1993, ch. 16) for a general discussion of the bootstrap approach to hypothesis testing.

[23] Robertson, Wright and Dykstra (1988) and Bhattacharya and Dykstra (1994) develop a test for equality of multivariate distributions against an alternative of first order dominance. We do not discuss this approach here. For continuous-variable models, methods for testing multidimensional first and higher order dominance have been developed by Crawford (2005), Duclos et al. (2006), McCaig and Yatchew (2007) and Anderson (2008).

as possible. We conjecture that the least favorable case in this situation is, in fact, a case of equal distributions, and that it is valid to interpret our bootstrap procedure along this line of reasoning. A detailed exploration of these subtle econometric issues is beyond the scope of the present paper.

References

[1] Abul Naga RH, Yalcin T, 2008. Inequality measurement for ordered response health data. Journal of Health Economics 27, 1614-1625.

[2] Abul Naga RH, Yalcin T, 2010. Median independent inequality orderings. SIRE Discussion Paper 2010-118.

[3] Alkire S, 2002. Dimensions of human development. World Development 30, 2, 181–205.

[4] Alkire S, Foster JE, 2011. Counting and multidimensional poverty measurement. Journal of Public Economics 95, 7, 476-487.

[5] Alkire S, Santos ES, 2011. Acute multidimensional poverty: a new index for developing countries. Proceedings of the German Development Economics Conference, Berlin 2011, No. 3.

[6] Allison RA, Foster JE, 2004. Measuring health inequality using qualitative data. Journal of Health Economics 23, 505-524.

[7] Anderson G, 2008. The empirical assessment of multidimensional welfare, inequality and poverty: sample weighted multivariate generalizations of the Kolmogorov-Smirnov two sample test for stochastic dominance. Journal of Economic Inequality 6, 73-87.

[8] Arndt C, Distante R, Hussain MA, Østerdal LP, Huong PL, Ibraimo M, 2012. Ordinal welfare comparisons with multiple discrete indicators: a first order dominance approach and application to child poverty. World Development 40, 11, 2290–2301.

[9] Apouey B, 2007. Measuring health polarization with self-assessed health data. Health Economics 16, 875–894.

[10] Atkinson AB, Bourguignon F, 1982. The comparison of multi-dimensioned distributions of economic status. Review of Economic Studies 12, 183-201.

[11] Atkinson AB, 2003. Multidimensional deprivation: contrasting social welfare and counting approaches. Journal of Economic Inequality 1, 53-65.

[12] Barnett V, 1976. The ordering of multivariate data. Journal of the Royal Statistical Society. Series A (General) 139, 318-355.

[13] Bhattacharya B, Dykstra RL, 1994. Statistical inference for stochastic ordering. In: M Shaked M, JG Shanthikumar (Eds.), Stochastic orders and their applications, Academic Press.

[14] Blair J, Lacy, MG, 1996). Measures of variation for ordinal data as functions of the cumulative distribution. Perceptual and Motor Skills, 82, 411-418.

[15] Blair J, Lacy MG, 2000. Statistics of ordinal variation. Sociological Methods & Research 28, 251-280.

[16] Boland PJ, Proschan F, 1988. Multivariate arrangement increasing functions with applications in probability and statistics. Journal of Multivariate Analysis, 25, 286-298.

[17] Bourguignon F, 1989. Family size and social utility: Income distribution dominance criteria. Journal of Econometrics 42, 67-80.

[18] Bourguignon F, Chakravarty S, 2003. The measurement of multidimensional poverty. Journal of Economic Inequality 1, 25-49.

[19] Buvinic M, Gupta GR, 1997. Female-headed households and female-maintained families: are they worth targeting to reduce poverty in de-

veloping countries. Economic Development and Cultural Change 45, 2, 259-280.

[20] Crawford I, 2005. A nonparametric test of stochastic dominance in multivariate distributions. DP 12/05, Department of Economics, University of Surrey.

[21] Davidson R, Duclos Jy, 2006. Testing for Restricted Stochastic Dominance. IZA Discussion Paper No. 2047.

[22] Dardanoni V and Forcina A, 1999. Inference for Lorenz curve orderings. Econometrics Journal 2, 49-75.

[23] DHS. Demographic and Health Survey. Macro International Inc. United States. Email: archive@orcmacro.com. Web: http://www.measuredhs.com.

[24] Decancq K, 2012. Elementary multivariate rearrangements and stochastic dominance on a Fréchet class. Journal of Economic Theory 147, 4, 1450-1459.

[25] Duclos J-Y, Sahn DE, Younger SD, 2006. Robust multidimensional poverty comparisons. Economic Journal 116, 943-968.

[26] Duclos J-Y, Sahn DE, Younger SD, 2007. Robust multidimensional poverty comparisons with discrete indicators of well-being. In: Inequality and Poverty Re-examined. Editors: SP Jenkins, J Micklewright. Oxford University Press.

[27] Duclos J-Y, Sahn DE, Younger SD, 2011. Partial multidimensional inequality orderings. Journal of Public Economics 95, 225-238.

[28] Efron B, Tibshirani RJ, 1993. An Introduction to the Bootstrap. Chapman & Hall.

[29] Epstein L, Tanny SM, 1980. Increasing generalized correlation: a definition and some economic consequences. Canadian Journal of Economics 13, 16-34.

[30] Grusky D, Kanbur R. 2006. The conceptual foundations of poverty and inequality measurement. In: D Grusky and R Kanbur (Eds.) Poverty and Inequality, Stanford University Press.

[31] Gordon D, Nandy S, Pantazis C, Pemberton S, Townsend, P, 2003. Child Poverty in the Developing World. Bristol, UK: Policy Press.

[32] Gravel N, Moyes P, 2006. Ethically robust comparisons of distributions of two individual attributes. Working paper 0605, IDEP.

[33] Gravel N, Moyes P, 2011. Bidimensional inequalities with an ordinal Variable. In: M Fleurbaey, M Salles and JA Weymark (Eds.). Social Ethics and Normative Economics, Social Ethics and Normative Economics: Essays in Honour of Serge-Christophe Kolm, Studies in Choice and Welfare, 101-127.

[34] Gravel N, Moyes P, 2012. Ethically robust comparisons of bidimensional distributions with an ordinal attribute. Journal of Economic Theory 1384-1426.

[35] Gravel N, Moyes P, Tarroux B, 2009. Robust International Comparisons of Distributions of Disposable Income and Regional Public Goods. Economica 76, 432-461.

[36] Gravel N, Mukhopadhyay A, 2010. Is India better off today than 15 years ago? A robust multidimensional answer. Journal of Economic Inequality 8, 173-195.

[37] Hamada K, 1974. Comment on Hadar and Russel (1974). In M Balch, D McFadden and S Wu (Eds.), Essays on Economic Behavior under Uncertainty, p. 150-153, North-Holland, Amsterdam.

[38] Haldane JBS, 1948. Note on the median of a multivariate distribution, Biometrica 35, 414-415.

[39] Hayford JF, 1902. What is the center of an area, or the center of a population? Journal of American Statistical Organization 8, 47-58.

[40] Handa S, 1996. Expenditure behavior and children's welfare: An analysis of female headed households in Jamaica. Journal of Development Economics 50, 165-187.

[41] Howes S, 1993. Inferring population rankings from sample data. A new test for inferring dominance from sample data. Discussion paper, STICERD, London School of Economics.

[42] IAF, 2002. Inquérito aos Agregados Familiares (Household Survey). Web: http://www.ine.gov.mz/inqueritos_dir/iaf/.

[43] INE. Instituto Nacional de Estatística (National Institut of Statistics). Maputo, Mozambique. Web: www.ine.gov.mz.

[44] Kaur A, Prakasa Rao BLS, Singh H, 1994. Testing for second-order stochastic dominance of two distributions. Econometric Theory 10, 849-866.

[45] Kobus M, 2012. Multidimensional inequality for ordinal data. Draft available at http://coin.wne.uw.edu.pl/mkobus/.

[46] Kobus M, Miłoś P, 2012. Inequality decomposition by population subgroups for ordinal data. Journal of Health Economics 31, 15-21.

[47] Lambert PJ, 2001. The distribution and redistribution of income. Third edition. Manchester University Press, Manchester.

[48] Lindelow M, 2006. Sometimes more equal than others: how health inequalities depend on the choice of welfare indicator. Health Economics 15, 263-279.

[49] Maasoumi E, 1999. Multidimensioned approaches to welfare analysis. In: J Silber (Ed.), Handbook of Income Inequality Measurement, Kluwer.

[50] Madden D, 2010. Ordinal and cardinal measures of health inequality: an empirical comparison. Health Economics 19, 243–250.

[51] McCaig B, Yatchew A, 2007. International welfare comparisons and nonparametric testing of multivariate stochastic dominance. Journal of Applied Econometrics 22, 951-969.

[52] Meyer M, Strulovici B, 2012. Increasing interdependence of multivariate distributions. Journal of Economic Theory 147, 4, 1460-1489.

[53] Moyes P, 2012. Comparisons of heterogeneous distributions and dominance criteria. Journal of Economic Theory, 147, 4, 1351-1383.

[54] MISAU. Ministério da Saúde (Ministry of Health). Maputo, Mozambique. Web: http://www.misau.gov.mz/.

[55] Müller A, Stoyan D, 2002. Comparison methods for stochastic models and risks. John Wiley and Sons, New York, NY.

[56] Muller C, Trannoy A, 2011. A dominance approach to the appraisal of the distribution of well-being across countries. Journal of Public Economics 95, 239-246.

[57] Orazem PF, King OM, 2007. Schooling in developing countries: The roles of supply, demand and government policy. In: Handbook of Development Economics, vol. 4. Elsevier Science B.V., Amsterdam.

[58] Parker S, Rubalcava L, Teruel G, 2007. Evaluating conditional schooling and health programs. In: Handbook of Development Economics, vol. 4. Elsevier Science B.V., Amsterdam.

[59] Quirk JD, Saposnik R, 1962. Admissibility and measurable utility functions. Review of Economic Studies 29, 140-146.

[60] Robertson T, Wright F, Dykstra RL, 1988. Order Restricted Statistical Inference. John Wiley, New York, NY.

[61] Savaglio E, 2006. Three approaches to the analysis of multidimensional inequality. In: Inequality and Economic Integration. Editors: F Farina, E Savaglio, Routledge.

[62] Sen A, 1985. Commodities and Capabilities, Elsevier, Amsterdam.

[63] Sen A, 1993. Capability and well-being, in M. Nassbaum and A. Sen (eds), The Quality of Life, Clarendon Press, Oxford.

[64] Sen A, 1997. On Economic Inequality. Expanded Edition with a substantial annexe by James E. Foster and Amartya Sen. Clarendon Press, Oxford.

[65] Sen A, 2006. Conceptualizing and measuring poverty. In: D Grusky and R Kanbur (Eds.) Poverty and Inequality, Stanford University Press.

[66] Shaked M, Shanthikumar JG, 2007. Stochastic Orders. Springer.

[67] Small CG, 1990. A survey of multidimensional medians. International Statistical Review 38, 263-277.

[68] Strassen V, 1965. The existence of probability measures with given marginals. The Annals of Mathematical Statistics 36, 423–439.

[69] Strauss J, Thomas D, 1995. Human resources: Empirical modeling of household and family decisions. In: Srinivasan, T.N., Behrman, J.R. (Eds.), Handbook of Development Economics, vol. 3A. North-Holland, Amsterdam, pp. 1883–2023.

[70] Tchen A, 1980. Inequalities for distributions with given marginals. The Annals of Probability 8, 814-827.

[71] Trannoy A, 2006. Multidimensional egalitarianism and the dominance approach: a lost paradise? In: Inequality and Economic Integration, F Farina and E Savaglio (eds.). Routledge, London.

[72] Tsui K, 1999. Multidimensional inequality and multidimensional generalized entropy measures: an axiomatic derivation. Social Choice and Welfare 16, 145-157.

[73] UNICEF, 2006. Childhood Poverty in Mozambique: A Situation and Trends Analysis. United Nations, Mozambique. Available at: http://www.eldis.org/fulltext/childpoverty_Moz.pdf.

[74] Weymark JA, 2006. The normative approach to measurement of multidimensional inequality. In: Inequality and Economic Integration. Editors: F Farina, E Savaglio, Routledge.

[75] Zheng B, 2008. Measuring inequality with ordinal data: a note. Research on Economic Inequality 16, 177-188.

[76] Zheng B, 2011. A new approach to measure socioeconomic inequality in health. Journal of Economic Inequality 9, 555-577.

[77] Østerdal LP, 2010. The mass transfer approach to multivariate discrete first order stochastic dominance: direct proof and implications. Journal of Mathematical Economics 46, 1222-1228.