Oxford Poverty & Human Development Initiative (OPHI)
Oxford Department of International Development
Queen Elizabeth House (QEH), University of Oxford

OPHI

Research in Progress series 2009

This paper is part of the Oxford Poverty and Human Development Initiative's Research in Progress (RP) series. These are preliminary documents posted online to stimulate discussion and critical comment. The series number and letter identify each version (i.e. paper RP1a after revision will be posted as RP1b) for citation.

Please cite these papers as Author Last name, First name, "Title" (Year) *OPHI Research in Progress ##a.* (For example: Alkire, Sabina and Foster, James "Counting and Multidimensional Poverty" (2007) *OPHI Research in Progress 1a.*)

For more information, see **www.ophi.org.uk**.

Oxford Poverty & Human Development Initiative (OPHI), Oxford Department of International Development, Queen Elizabeth House (QEH), University of Oxford, 3 Mansfield Road, Oxford OX1 3TB, UK
Tel. +44 (0)1865 271915, Fax +44 (0)1865 281801, ophi@qeh.ox.ac.uk, http://www.ophi.org.uk

OPHI gratefully acknowledges support from the UK Economic and Social Research Council (ESRC)/(DFID) Joint Scheme, Robertson Foundation, Praus, UNICEF N'Djamena Chad Country Office, German Federal Ministry for Economic Cooperation and Development (GIZ), Georg-August-Universität Göttingen, International Food Policy Research Institute (IFPRI), John Fell Oxford University Press (OUP) Research Fund, United Nations Development Programme (UNDP) Human Development Report Office, national UNDP and UNICEF offices, and private benefactors. International Development Research Council (IDRC) of Canada, Canadian International Development Agency (CIDA), UK Department of International Development (DFID), and AusAID are also recognised for their past support.

Testing for stochastic dominance among additive, multivariate welfare functions with discrete variables[1]

Gaston Yalonetzky

OPHI, University of Oxford
gaston.yalonetzky@qeh.ox.ac.uk

ABSTRACT

A flourishing literature on robustness in multidimensional welfare and poverty comparisons has aroused the interest on multidimensional stochastic dominance. By generalizing the dominance conditions of Atkinson and Bourguignon (1982) this paper offers complete conditions, alternative to those proposed by Duclos *et al.* (2006a,b). We also show how to test these conditions for discrete variables extending the non-parametric test by Anderson (1996) to multiple dimensions. An empirical application illustrates these tests.

Key Words: Multidimensional welfare comparisons, stochastic dominance, nonparametric tests

JEL Classification: C14, D30

This version: June 2009.

1. Introduction

A recent literature on multidimensional wellbeing emphasizes comparing individuals or groups with welfare functions which aggregate separate dimensions of well-being into a single indicator. Such dimensions can include consumption items, health attributes and, in general, any other aspect of wellbeing that "people have reason to value" (Sen, 2001). Since several

[1] I am very grateful to Ian Crawford for his comments and support. I would also like to thank Jean-Yves Duclos, Sabina Alkire and participants at the OPHI workshop on Robustness in Multdimensional Welfare Analysis for their valuable comments. I am very grateful to Lorenzo Oimas for helping me with securing the dataset.

such evaluative functions can be proposed, the choices of functional forms, weights attached to dimensions and other details may affect the consistency of the rankings of societies or individuals. In the twin literature of multidimensional poverty an additional source of ranking inconsistency is the choice of multidimensional poverty line and criterion for the identification of the poor (Bourguignon and Chakravarty, 2002; Duclos *et al.*, 2006). Two approaches have been proposed to provide robust welfare comparisons in the presence of such range of choices: one is to conduct sensitivity analysis of a given indicator to gauge the impact of a change in an aspect of the indicator (e.g. the weights) on its ranking performance. Another approach is to provide multidimensional stochastic dominance conditions which state the conditions under which a broad class of welfare functions consistently rank multivariate distributions of individuals or societies. Such exercise has two components: the dominance conditions themselves and the corresponding statistical tests to probe them empirically.

To date there is a broad literature both on stochastic dominance conditions for comparisons of welfare indicators[2], poverty indicators[3] , Lorenz curves[4] , even economic mobility (Fields et al., 2002). Similarly statistical tests abound (e.g. Anderson, 1996; Xu and Osberg, 1998; Davidson and Duclos, 2000; Barrett and Donald, 2003; Thuysbaert and Zitikis, 2005; Linton *et al.*, 2005; Bennett, 2008). They all work with univariate distributions. For multidimensional comparisons the first package of dominance conditions and corresponding statistical test was provided by Duclos *et al.* (2006,2007), in turn based on the theory laid out by Davidson and Duclos (2000). Notwithstanding the soundness of their technique, the conditions of Duclos *et al.* (2006,2007) do not cover all the potential signs of the cross-derivatives of the welfare functions. By contrast the older conditions laid out by Atkinson and Bourguignon (1982) cover different signs of the cross-derivatives. In this paper we show with a bivariate illustration that the conditions by Atkinson and Bourguignon (1982) include all the conditions of Duclos *et al.* (2006) and are more complete in that they also consider all possibilities in terms of signs of cross-derivatives.[5] Thereby this paper provides justification to finding alternative statistical techniques to test the conditions of Atkinson and Bourguignon which are generalized, in this paper, to n dimensions and up to third-order

---

[2]For instance, Atkinson and Bourguignon (1982), Muller and Trannoy (2003).

[3]For instance, Foster and Shorrocks (1988), Davidson and Duclos (2000), Duclos and Makdissi (2005), Duclos et al. (2006).

[4]For instance, Dardanoni and Forcina (1999)

[5]On the other hand, the robustness work of Davidson and Duclos (2000) as well as Duclos et al. (2006) is not restricted to dominance conditions and tests. They also show how to estimate consistently critical frontiers, i.e. the points at which cumulative distributions or dominance surfaces cross.

stochastic dominance. The importance of the sign of the cross-derivatives is illustrated with an example from the family of association-sensitive welfare measures (Seth, 2009).

In this paper I propose a test of multidimensional stochastic dominance suitable for the conditions of Atkinson and Bourguignon, based on Anderson's (1996), Ibbot's (1998) and Crawford's (2005) nonparametric tests. The test is useful for ordinal, discrete variables since it is based on the standard errors of multivariate, multinomial distributions. Besides the application to ordinal, discrete variables does not require using the trapezoidal approximation to the integral areas which are a source of inconsistency (Barret and Donald, 2003).

In multidimensional welfare comparisons there is no way to escape the problem of choosing a welfare function and therefore robustness analysis, e.g. stochastic dominance, is warranted.[6] As Crawford (2005) notes, real income comparisons, for instance, are not good substitutes. In these comparisons a vector of household demands and demographic characteristics is aggregated using market prices as weights for the marketed goods and adjusted by an equivalence scale to reflect differences in household members' composition like ages and size (Sen, 1979). Real incomes are not useful as subtitutes to welfare functions, even assuming that real incomes could be calculated without controversy or difficulty[7] for at least two reasons: firstly, non-expenditure attributes, e.g. a long and healthy lifetime, being educated, etc. are valued by people on top of consumption items (Sen, 2001). A counter-argument would claim that the service flow of such non-expenditure items could be included in the real income calculation with the appropriate prices, should the exercise be possible. However, and that is the second reason, even in such a scenario the resulting metric would be a linear combination of the welfare-enhancing variables ranging from the consumption items to the non-expenditure ones (e.g. political freedoms) thus implying a perfect degree of substitutability across wellbeing dimensions. Practicioners in multidimensional welfare analysis are not eager to limit the scope of welfare function to such restrictive option.

The paper proceeds as follows. In section 2 a brief introduction to the conditions of Duclos *et al.* (2006) is provided. Then I introduce the conditions of Atkinson and Bourguignon and its extensions by Crawford (2005) to third order of dominance, in order then

[6]Notice that a welfare function is different from a utility function in that with the former there is no intent to represent the preferences of the invidual under evaluation.

[7]That is, for instance, if equivalence scales could be estimated, if relevant prices are observed and they reflect consumers' valuation of the goods, they do not vary across households because of preferences toward risk, regional location, productivity differentials in non-marketed goods, etc.)

to show their completeness in terms of accounting for different signs of the cross-partial derivatives and in terms of the ability of the framework to replicate the conditions of Duclos *et al.* (2006) in the context of welfare functions. I illustrate the importance of having conditions which ensure robustness to the sign of the cross-derivative with an example involving association-sensitive welfare functions. In the third section I propose a test for the stochastic dominance conditions suitable for discrete variables. The estimators and tests for multivariate, multinomial distributions corrects Crawford's (2005) multidimensional extension of Anderson's (1996) univariate test for inconsistency in the trapezoidal approximation by being tailored exclusively for discrete variables and multinomial distributions. Up to section 3 the analysis is performed in a bivariate setting. Section 4 discusses the extension of these tests to higher dimensions with an important contribution in the form of a simple but accurate rule for stacking the probabilities of a multidimensional joint distribution in a matrix, which renders the estimation of the statistics substantially easier. Section 5 applies these tests to a three-dimensional setting using data from Peru. Section 6 draws some conculsions and outlines future research ideas.

2. Multivariate stochastic dominance conditions

So far the literature on multidimensional stochastic dominance has focused on the class of additive welfare functions when deriving conditions under which all functions belonging to specific subclasses of this additive class rank individuals, households, societies or social states unanimously. The additivity property of this general class means that total welfare is measured by the welfare function as the sum of the contributions to welfare of all the economic units involved (e.g. households). These contributions are given by functions evaluated on the vector space of wellbeing-enhancing variables. Following Crawford (2005) the general class κ of additive welfare function is defined as:

$$\kappa = \left\{ W\left(F\right) \mid W\left(F\right) = \int ... \int \psi\left(x_1, ..., x_D\right) dF\left(x_1, ..., x_D\right) \right\} \tag{1}$$

where the function $\psi\left(x_1, ..., x_n\right) : \mathbb{R}^n \rightarrow \mathbb{R}$ measures the contribution of an economic unit (e.g. a household, an individual) to total welfare. Notice that $\psi$is not a utility function. Therefore the welfare functions in class κ are not utilitarian but could be so if ψwere defined as a utility function. Societies ranked by these welfare functions can have any joint distribution function. The set of all possible distributions, $\mathcal{F}$, is defined as:

$$\mathcal{F} = \left\{ \begin{array}{c} F : \mathbb{R}^n \to [0,1] \\ F \text{ nondecreasing and continuous} \\ F(0, ..., 0) = 0; F(a_1, ..., a_n) = 1 \end{array} \right\} \tag{2}$$

where the range of x_i is assumed to be $[0, a_i]$. I also define: $f(x_1, ..., x_n) = \frac{\partial^n F(x_1,...,x_n)}{\partial x_1 ... \partial x_n}$.Since one of the key applications of multidimensional stochastic dominance is on poverty comparisons I follow Duclos et al. (2006) in presenting a class of social poverty functions similar to κ in (1):

$$\Xi = \left\{ P(\lambda) \mid P(\lambda) = \int ... \int \pi(x_1, ..., x_n; \lambda) \, dF(x_1, ..., x_n) \right\} \tag{3}$$

where $\lambda : \mathbb{R}^n \to \mathbb{R} \mid \frac{\partial \lambda}{\partial x_i} \geqslant 0$ is an identificator function which defines whether an economic unit is poor or not depending on the values of x. Moreover $\lambda(x_1, ..., x_n) = 0$ defines the poverty hyperline over the welfare-enhancing variables space. Whenever a unit's values for the variables imply $\lambda(x_1, ..., x_n) \leqslant 0$ the unit is classified as poor and $\pi(x_1, ..., x_n; \lambda) \geqslant 0$. Otherwise $\pi(x_1, ..., x_n; \lambda) = 0$. Thereby the focus axiom is fulfilled (Duclos *et al.* 2006, p. 947).

In the bi-dimensional case Duclos *et al.* (2006) also define a *bi-dimensional stochastic dominance surface* which is at the core of their dominance conditions and resembles a multiplicative FGT poverty index[8] :

$$P^{\alpha_1, \alpha_2}(z_1, z_2) = \int_0^{z_1} \int_0^{z} (z_1 - x_1)^{\alpha_1} (z_2 - x_2)^{\alpha_2} \, dF(x_1, x_2) \tag{4}$$

where $z_1$and z_2 are univariate poverty lines.

2.1. The conditions of Duclos *et al.* (2006)

The stochastic dominance conditions of Duclos *et al.* (2006) relate the consistency of rankings of a subclass of additive poverty measures to a relationship between the dominance surfaces of two samples. For different values of $\alpha_1$and α_2 different subclasses of poverty functions rank consistently. As the alpha parameters increase from zero upward more conditions

[8]FGT stands for Foster, Greer and Thorbecke (1984).

are imposed on the derivatives of the poverty functions and thus the subclass is reduced. Duclos *et al.* find the following general result for the bivariate case:

$$\Delta P(\lambda) > 0, \forall P(\lambda) \in \Pi^{\alpha_1+1,\alpha_2+1} \leftrightarrow \Delta P^{\alpha_1,\alpha_2} > 0 \forall (x_1, x_2) \in \Lambda(\lambda^*) \tag{5}$$

where $\Delta P = \Delta P_A - \Delta P_B$, with A and B being two samples, $\Lambda(\lambda^*)$ denotes a set of (x_1, x_2) such that $\lambda^*(x_1, x_2) \leqslant 0$ and λ^* is the maximal poverty hyperline. $\Pi^{\alpha_1+1,\alpha_2+1}$ is an additive subclass. They illustrate with the following two examples of subclasses:

$$\Pi^{1,1} = \left\{ P(\lambda) \mid \begin{array}{c} \Lambda(\lambda) \sqsubset \Lambda(\lambda^*) \\ \pi(x_1, x_2, \lambda) = 0, \text{whenever } \lambda(x_1, x_2) = 0 \\ \pi^{x_i} \leq 0 \quad \forall x_i \\ \pi^{x_1 x_2} \geq 0 \quad \forall x_1, x_2 \end{array} \right\} \tag{6}$$

$$\Pi^{2,1} = \left\{ P(\lambda) \mid \begin{array}{c} P(\lambda) \in \Pi^{1,1} \\ \pi^{x_1}(x_1, x_2, \lambda) = 0, \text{whenever } \lambda(x_1, x_2) = 0 \\ \pi^{x_1 x_1} \geq 0 \quad \forall x_1 \\ \pi^{x_1 x_1 x_2} \leq 0 \quad \forall x_1, x_2 \end{array} \right\} \tag{7}$$

The superscripts on the π denote derivatives (e.g. $\pi^{x_i} = \frac{\partial \pi}{\partial x_i}$). Notice that these conditions allow for increasing the order of dominance dimension-by-dimension as oppossed to just increasing it for all dimensions at the same time. These are the first results in the multidimensional dominance literature to show these extensions.[9] However as I show below the conditions presented by Atkinson and Bourguignon (1982) and Crawford (2005) also allow for increasing the order of dominance in the same way.

2.2. The conditions of Atkinson and Bourguignon (1982), Crawford (2005) and beyond

In this section I summarize the conditions of Atkinson and Burguignon for the bivariate case (both those derived and listed by the authors), and the extension to third-order domi-

[9] The conditions in Duclos et al. (2006a) are very similar. For instance, substitutability between pairs of discrete and continuous variables is also considered in the definition of the class of poverty measures. On the other hand these conditions are more sophisticated in their treatment of differential poverty lines for different values of the discrete variables.

nance by Crawford (2005). Then I show that these conditions allow for increasing the order of dominance dimension-by-dimension. The importance of having conditions for either sign of the cross-derivatives of ψ is illustrated with an example using a class of association-sensitive welfare functions.

Let's start by defining the difference in welfare between two samples as:

$$\Delta W = \int_0^{a_1} \int_0^{a_2} \psi(x_1, x_2) \Delta f(x_1, x_2) \, dx_2 dx_1 \tag{8}$$

Following Atkinson and Bourguignon (1982) the marginal distributions of $F(x_1, x_2)$ are defined as $F_i(x_i)$. The following expressions are also defined: $K(x_1, x_2) \equiv F_1(x_1) + F_2(x_2) - F(x_1, x_2)$, $H(x_1, x_2) \equiv \int_0^{x_1} \int_0^{x_2} F(s,t) \, dsdt$, $H_i(x_i) \equiv \int_0^{x_i} F_i(s) \, ds$, and $L(x_1, x_2) \equiv \int_0^{x_1} \int_0^{x_2} K(s,t) \, dsdt$.

2.3. First-order dominance results

The results are the following:

$$\psi \in \Psi^- \leftrightarrow (\Delta W \geq 0 \leftrightarrow \Delta F_i(x_i) \leq 0 \forall i = 1,2 \wedge \Delta F(x_1, x_2) \leq 0 \quad \forall x_1, x_2) \tag{9}$$

$$\psi \in \Psi^+ \leftrightarrow (\Delta W \geq 0 \leftrightarrow \Delta F_i(x_i) \leq 0 \forall i = 1,2 \wedge \Delta K(x_1, x_2) \leq 0 \quad \forall x_1, x_2) \tag{10}$$

$$\psi \in \Psi^1 \leftrightarrow (\Delta W \geq 0 \leftrightarrow \Delta F_i(x_i) \leq 0 \forall i = 1,2 \wedge \Delta F(x_1, x_2) \leq 0 \wedge \Delta K(x_1, x_2) \leq 0 \quad \forall x_1, x_2) \tag{11}$$

The result in (9) is due to Hadar and Russel (1974).[10] The result in (10) is due to Levy and Paroush (1974). Ψ^- is a subclass of aditive welfare measures: $\Psi^- = \left\{\psi \colon \psi_i \geq 0 \wedge \psi_{ij} \leq 0\right\}$, where $\psi_i = \frac{\partial \psi}{\partial x_i}$. While $\Psi^+ = \left\{\psi \colon \psi_i \geq 0 \wedge \psi_{ij} \geq 0\right\}$ and $\Psi^1 = \Psi^- \cup \Psi^+ = \{\psi \colon \psi_i \geq 0\}$.

[10] They show that the result can be extended to n-dimensions. Notice that $\Delta F(x_1, x_2) \leq 0$ implies $\Delta F_i(x_i) \leq 0 \forall i = 1,2$.

2.4. Second-order dominance results

Second-order dominance corresponds to subclasses of welfare functions which represent preferences for mean-preserving, inequality-reducing transfers. The results are the following:

$$\psi \in \Psi^{--} \leftrightarrow \left(\Delta W \geq 0 \leftrightarrow \Delta H_i\left(x_i\right) \leq 0 \forall i=1,2 \wedge \Delta H\left(x_1,x_2\right) \leq 0 \quad \forall x_1,x_2\right) \tag{12}$$

$$\psi \in \Psi^{++} \leftrightarrow \left(\Delta W \geq 0 \leftrightarrow \Delta H_i\left(x_i\right) \leq 0 \forall i=1,2 \wedge \Delta L\left(x_1,x_2\right) \leq 0 \quad \forall x_1,x_2\right) \tag{13}$$

$$\psi \in \Psi^{2} \leftrightarrow \left(\Delta W \geq 0 \leftrightarrow \Delta H_i\left(x_i\right) \leq 0 \forall i=1,2 \wedge \Delta H\left(x_1,x_2\right) \leq 0 \wedge \Delta L\left(x_1,x_2\right) \leq 0 \quad \forall x_1,x_2\right) \tag{14}$$

The results in (12) and (13) are due to Atkinson and Bourguignon (1982). $\Psi^{--} = \left\{\psi\colon \psi \in \Psi^{-} \wedge \psi_{ii} \leq 0,\ \psi_{iij} \geq 0,\ \psi_{iijj} \leq 0\right\}$, $\Psi^{++} = \left\{\psi\colon \psi \in \Psi^{+} \wedge \psi_{ii} \leq 0,\ \psi_{iij} \leq 0,\ \psi_{iijj} \geq 0\right\}$, and $\Psi^2 = \Psi^{--} \cup \Psi^{++} = \{\psi\colon \psi \in \Psi^1 \wedge \psi_{ii} \leq 0\}$.

2.5. Third-order dominance results

Crawford (2005) extended the results up to the third order of dominance since it has received attention in the univariate literature as corresponding to subclasses of welfare function which favour mean-preserving, inequality-reducing transfers more when these happen at the lower end of the distributions. Let $J\left(x_1,x_2\right) \equiv \int_0^{x_1}\int_0^{x_2} H\left(s,t\right) dsdt,\ J_i\left(x_i\right) \equiv \int_0^{x_i} H_i\left(s\right) ds$, and $M\left(x_1,x_2\right) \equiv \int_0^{x_1}\int_0^{x_2} L\left(s,t\right) dsdt$. The results are the following:

$$\psi \in \Psi^{---} \leftrightarrow \left(\Delta W \geq 0 \leftrightarrow \Delta J_i\left(x_i\right) \leq 0 \forall i=1,2 \wedge \Delta J\left(x_1,x_2\right) \leq 0 \quad \forall x_1,x_2\right) \tag{15}$$

$$\psi \in \Psi^{+++} \leftrightarrow \left(\Delta W \geq 0 \leftrightarrow \Delta J_i\left(x_i\right) \leq 0 \forall i=1,2 \wedge \Delta M\left(x_1,x_2\right) \leq 0 \quad \forall x_1,x_2\right) \tag{16}$$

$$\psi \in \Psi^{3} \leftrightarrow \Delta\left(W \geq 0 \leftrightarrow \Delta J_i\left(x_i\right) \leq 0 \forall i=1,2 \wedge \Delta J\left(x_1,x_2\right) \leq 0 \wedge \Delta M\left(x_1,x_2\right) \leq 0 \quad \forall x_1,x_2\right) \tag{17}$$

$\Psi^{---} = \{\psi \colon \psi \in \Psi^{--} \wedge \psi_{1112}, \psi_{1122} \geq 0,\ \psi_{11122}, \psi_{11222} \geq 0,\ \psi_{111222} \leq 0\},$

$\Psi^{+++} = \{\psi \colon \psi \in \Psi^{++} \wedge \psi_{1112}, \psi_{1122} \geq 0,\ \psi_{11122}, \psi_{11222} \geq 0,\ \psi_{111222} \geq 0\},$

and $\Psi^{3} = \Psi^{---} \cup \Psi^{+++} = \{\psi \colon \psi \in \Psi^{2} \wedge \psi_{1112},\ \psi_{11122}, \psi_{11222} \geq 0\}$.

The results shown so far are for conditions in which the order of dominance is raised simultaneously for all dimensions. In the next subsection I illustrate the potential of this approach to produce more stochastic dominance conditions for additional subclasses.

2.6. More results changing the order of dominance dimension-by-dimension

The following result is analogue to the result of Duclos *et al.* (2006) in (5), for subclass (7) :

$$\psi \in \Psi^{-1} \leftrightarrow \left(\Delta W \geq 0 \leftrightarrow \Delta F_2 (x_2) \leq 0 \wedge \Delta H_1 (x_1) \leq 0 \wedge \int_0^{x_1} \Delta F (t, x_2)\, dt \leq 0 \quad \forall x_1, x_2 \right) \tag{18}$$

where $\Psi^{-1} = \{\psi \colon \psi \in \Psi^{-} \wedge \psi_{11} \leq 0 \wedge \psi_{112} \geq 0\}$. A similar condition can be obtained for the following subclass: $\Psi^{-2} = \{\psi \colon \psi \in \Psi^{-} \wedge \psi_{22} \leq 0 \wedge \psi_{122} \geq 0\}$. The derivation is in the Appendix 1. To relate these results to those of Duclos et al. (2006) W needs to be regarded as an aggregate poverty function and the signs of the derivatives in the subclass have to be reversed accordingly.

2.7. Relevance of the conditions for the literature on multidimensional welfare indicators: an example

Seth (2009) has recently characterized a class of association-sensitive welfare indices based on general means (e.g. similar to CES functions in the traditional microeconomics literature). The index depends on two parameters, α and β. When $\alpha, \beta \neq 0$ the index is:

$$W = \left[\frac{1}{N} \sum_{n=1}^{N} \left[\sum_{d=1}^{D} a_d x_{nd}^{\beta} \right]^{\frac{\alpha}{\beta}} \right]^{\frac{1}{\alpha}} \quad \forall \alpha, \beta \neq 0 \tag{19}$$

Such index strictly increases whenever there is increased association among the dimensions if and only if $\beta < \alpha \leq 1$ (Seth, 2009, p.12). Therefore such parameter range is worth considering if a welfare indicator with that property is desired for a given welfare evaluation. Now consider the case with just two dimensions (variables) and $\alpha = 1$. It yields: $W^{\alpha=1} = \frac{1}{N}\sum_{n=1}^{N}\left[a_1 x_{n1}^{\beta} + a_2 x_{n2}^{\beta}\right]^{\frac{1}{\beta}}$. Define $v_n = \left[a_1 x_{n1}^{\beta} + a_2 x_{n2}^{\beta}\right]^{\frac{1}{\beta}}$. The cross-partial derivative is:

$$v_{12} \equiv \frac{\partial^2 v_n}{\partial x_{n1}\partial x_{n2}} = \beta\left(1-\beta\right) v_n^{-\beta} a_1 a_2 \left(x_{n1}x_{n2}\right)^{\beta-1} \tag{20}$$

If $x_{nd}, a_d \geq 0 \forall n, d$, then $(-\infty < \beta < 0 \vee \beta > 1) \leftrightarrow v_{12} < 0$ and $0 < \beta < 1 \leftrightarrow v_{12} > 0$. Since the choice of β affects the cross-partial derivative the ranking performance of these association-sensitive indicators may or may not be robust to the choice of β. It depends on the pair of compared samples. For example imagine x_1 and x_2 take only two values: a low and a high one. Imagine society A and society B have the following joint distributions of x_1 and x_2 (lowest values are on the top-left corners):

$$\overset{A}{\begin{matrix} 0.1 & 0.2 \\ 0.2 & 0.5 \end{matrix}} \qquad \overset{B}{\begin{matrix} 0.2 & 0.2 \\ 0.2 & 0.4 \end{matrix}} \tag{21}$$

These two societies are ranked unanimously by the aforementioned association-sensitive measures regardless of the choice of β within the interval $]-\infty, 1[$, because condition (11) is fulfilled. However such is not the case for societies C and D whose joint distributions are the following:

$$\overset{C}{\begin{matrix} 0.2 & 0.3 \\ 0.3 & 0.2 \end{matrix}} \qquad \overset{D}{\begin{matrix} 0.3 & 0.2 \\ 0.2 & 0.3 \end{matrix}} \tag{22}$$

In this case condition (9) is fulfilled but not (10). Therefore association-sensitive measures of the type described above only rank C and D consistently as long as the choice of β falls within the interval $]-\infty, 0[$. In other words there is no guarantee that different choices of β within the interval $]-0, 1[$ do not reverse the ranks attributed by the welfare indicator. The analysis illustrated by this example requires an approach offering conditions for different signs of the cross-partial derivatives.

3. Estimation and inference

In this section I propose a test of multidimensional stochatic dominance for the above conditions suitable for discrete variables. The test is a corrected version of the extension that Crawford (2005) made of Anderson's (1996) univariate test. Both Anderson's and Crawford's tests become inconsistent when applied to continuous variables for testing second and higher orders of dominance. (Barret and Donald, 2003) For those tests of second-to-higher dominance, the trapezoidal approximations to the areas under the integrals of the cumulative distribution functions are the source of inconsistency. This problem is corrected when the formulas for the statistics are altered in order to estimate the areas under the cumulative distributions of exclusively discrete, ordinal variables. For each condition the idea is to construct estimates of the distributional differences, e.g. $\Delta F(x_1, x_2)$, depending on the dominance condition under scrutiny.

The tests presented here are different to those of Duclos *et al.* (2007). The latter tests are for a combination of one continuous variable and several discrete ones (although it could be extended to several continuous variables). Those tests are related to a version of the conditions by Duclos *et al.* (2006) that incorporates discrete variables. As with their first set of conditions, these conditions consider dimensions to be substitutes. Besides Duclso *et al.* (2007) rely on the estimation of powers of the poverty gap with respect to the continuous variable as their statistics. In the tests presented below the statistics are based on the probabilities of the multivariate multinomial distributions of well-being.

3.1. Extending Anderson's test for applications to discrete variables

Let's start with the bivariate case.[11] Without loss of generality, let x_1 and x_2 be two discrete variables taking values on the natural line in the range $[1, C_k]\, \forall k = 1, 2$. The number of observations in cell ij is n_{ij} so that: $\sum_{i=1}^{C_1} \sum_{j=1}^{C_2} n_{ij} = n$.The probability of being in cell ij is: $p_{ij} \equiv P\{x_1 = i \wedge x_2 = j\}$. Its empirical counterpart is:

$$\widehat{p_{ij}} = \frac{1}{n} \sum_{k=1}^{n} I\left(x_1^k = i \wedge x_2^k = j\right) \tag{23}$$

As Formby *et al.* (2004) show, the empirical probabilities of multinomial distributions

[11] Ibbot(1998) also described a test for first-order dominance in bivariate distributions. Here I describe tests for first and higher orders of dominance.

are asymptotically distributed as normal with mean μ and variance Ω, where:

$$\mu = \begin{bmatrix} p_{11} & \cdots & p_{1C_2} \\ \vdots & \ddots & \vdots \\ p_{C_1 1} & \cdots & p_{C_1 C_2} \end{bmatrix} \tag{24}$$

$$\Omega = \begin{bmatrix} p_{11}\left(1-p_{11}\right) & -p_{11}p_{12} & \cdots & -p_{11}p_{C_1C_2} \\ -p_{11}p_{12} & p_{12}\left(1-p_{12}\right) & \cdots & -p_{12}p_{C_1C_2} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{11}p_{C_1C_2} & -p_{12}p_{C_1C_2} & \cdots & p_{C_1C_2}\left(1-p_{C_1C_2}\right) \end{bmatrix} \tag{25}$$

The multidimensional, multinomial distributions of two samples can be compared by stacking the probabilities in respective column vectors, $\widehat{P^A}$ and $\widehat{P^B}$ and denoting the difference as:

$$\widehat{V} = \left(\widehat{P^A} - \widehat{P^B}\right) \tag{26}$$

Under the null hypothesis of homogenenous distributions:

$$\widehat{V} \xrightarrow{d} N\left(0, \frac{n^A + n^B}{n^A n^B}\Omega\right) \tag{27}$$

The extensions of Anderson (1996)'s test are derived straightforwardly to test for the distributional conditions by noticing that the latter are all linear combinations of the probabilities in (26). The following examples illustrate the approach:

Estimation of ΔF, ΔH, ΔJ:

In the bivariate case a straightforward estimation follows by rewriting $\widehat{V}$ as a matrix, $\widehat{V}_{C_1,C_2}$, and by considering L_r to be a r-dimension, lower triangular matrix of ones:

$$\widehat{\Delta}^Z = \left[L_{C_1}\right]^Z \widehat{V}_{C_1,C_2} \left[L'_{C_2}\right]^Z \tag{28}$$

Such that: $\widehat{\Delta}^1 = \widehat{\Delta F}$, $\widehat{\Delta}^2 = \widehat{\Delta H}$, $\widehat{\Delta}^3 = \widehat{\Delta J}$. Notice that for (discretized) continuous variables these estimations become imprecise when $z \geq 2$. A better approximation, for instance, is the trapezoidal one (Anderson, 1996; Crawford, 2005). However even such approximation has been deemed to render tests potentially inconsistent (Barrett and Donald, 2003). That is why I restrict the application of this extended test to ordinal, discrete variables for whom no special assumption is made regarding any underlying continuous latent variable with an unknown distribution function and no meaningful distance is attributed to the differences in values.

Estimation of ΔF_i, ΔH_i, ΔJ_i:

In the bivariate case the estimation again is straightforward:

$$\widehat{\Delta}_1^Z = [L_{C_1}]^Z \, \widehat{V}_{C_1,C_2} 1_{C_2} \tag{29}$$

where 1_r is a column vector of ones of dimension r. Now: $\widehat{\Delta}_1^1 = \widehat{\Delta F}_1$, $\widehat{\Delta}_1^2 = \widehat{\Delta H}_1$, $\widehat{\Delta}_1^3 = \widehat{\Delta J}_1$. And:

$$\widehat{\Delta}_2^Z = 1'_{C_1} \widehat{V}_{C_1,C_2} \left[L'_{C_2}\right]^Z \tag{30}$$

Hence: $\widehat{\Delta}_2^1 = \widehat{\Delta F}_2$, $\widehat{\Delta}_2^2 = \widehat{\Delta H}_2$, $\widehat{\Delta}_2^3 = \widehat{\Delta J}_2$.

Estimation of ΔK, ΔL, ΔM:

$$\widehat{\Delta}_{1+2-12}^Z = [L_{C_1}]^{Z-1} \begin{bmatrix} -I_{C_1} & \vdots & I_{C_1} & \vdots & I_{C_1} \end{bmatrix} \begin{bmatrix} L_{C_1}\widehat{V}_{C_1,C_2}L'_{C_2} \\ \cdots \\ L_{C_1}\widehat{V}_{C_1,C_2}1_{C_2}1'_{C_2} \\ \cdots \\ 1_{C_1}1'_{C_1}\widehat{V}_{C_1,C_2}L'_{C_2} \end{bmatrix} \left[L'_{C_2}\right]^{Z-1} \tag{31}$$

Where I_r is an identitiy matrix of dimension r. $\widehat{\Delta}_{1+2-12}^1 = \widehat{\Delta K}$, $\widehat{\Delta}_{1+2-12}^2 = \widehat{\Delta L}$, $\widehat{\Delta}_{1+2-12}^3 = \widehat{\Delta M}$

Estimation of $\int_0^{x_1} \Delta F\left(t, x_2\right) dt \; (\forall x_1, x_2)$:

It follows from $\widehat{\Delta}^1$ (28): $\widehat{\Delta}_1^{1\frac{1}{2}} = L_{C_1}\widehat{\Delta}^1$

The respective standard errors of these functions, which are linear combinations of the probabilities, are based on the result in (27). For instance for the family $\widehat{\Delta}^Z$ they are the square root of the elements of $var\left(\widehat{\Delta}^Z\right)$:

$$var\left(\widehat{\Delta}^Z\right) = \left[L_{C_2}\right]^Z \otimes \left[L_{C_1}\right]^Z \left(\frac{n^A + n^B}{n^A n^B}\widehat{\Omega}\right) \left[L'_{C_2}\right]^Z \otimes \left[L'_{C_1}\right]^Z \tag{32}$$

where the probability elements of $\widehat{\Omega}$, $\widehat{p_{ij}}^H$, are estimated under the null hypothesis of homogeneous distributions for A and B, following Anderson (1996):

$$\widehat{p_{ij}}^H = \frac{n^A}{n^A + n^B}\widehat{p_{ij}}^A + \frac{n^B}{n^A + n^B}\widehat{p_{ij}}^B \tag{33}$$

At this point, the extension of Anderson's test involves constructing z-statistics[12] by dividing each element of the $\widehat{\Delta}$ by their respective standard errors in the diagonal of $\sqrt{var\left(\widehat{\Delta}\right)}$. These statistics are used for a test of the dominance conditions of the previous section, in which the null hypothesis is of homogeneity or common distribution and the alternative is of dominance. For instance, using the convention adopted in Anderson (1996) and Bishop et al. (1989), for condition (9) we test $H_o\colon \Delta F(.) = 0$ against $H_{1F}\colon \Delta F(.) \leq 0$. Rejection of H_o requires:

$$\Delta F(x_1, x_2) \leq 0 \; \forall x_1, x_2 \tag{34}$$

$$\exists x_1, x_2 \mid \Delta F(x_1, x_2) < 0 \tag{35}$$

These criteria, (34) and (35) mean that no element of the matrix $\widehat{\Delta F}$ is significantly greater than zero and at least one element is significantly less than zero[13]. Rejection of the null in favour of this alternative hypothesis requires that there exit both significantly positive and significantly negative elements in $\widehat{\Delta F}$ (Anderson, 1996). Note also that the statistics based on the marginal distributions (ΔF_i, ΔH_i, ΔJ_i) are those of Anderson's tests

[12] These statistics have an asymptotic t-distribution with infinite degrees of freedom.

[13] There is also a further alterantive hypothesis of indeterminancy. For instance in the case of testing condtion (9): it is $H_{1F}\colon \Delta F(.) \nleq \wedge \ngeq 0$.

for univariate distributions should they be applied to discrete variables. Several null and alternative hypotheses for the dominance conditions in the bivariate and the univariate cases are in tables 6 and 7. Since multiple comparisons are involved for each $\widehat{\Delta}$, corresponding to each of its elements, and each involves a comparison in means, I follow Anderson (1996) and Crawford (2005) in using the critical values from the studentised maximum modulus (SMM) distribution (Stoline and Ury, 1979) for $C_1C_2 - 1$ contrasts when $Z = 1$, and C_1C_2 contrasts when $Z \geq 2$.

These tests are symmetric in that whenever one can not reject H_o in favour of the alternative hypotheses for a given criterion of dominance, one can not establish such dominance in the way the differences in $\widehat{\Delta}$ have been constructed (e.g. $H^A - H^B$), but one can establish dominance in the reverse order for the same subclass of welfare functions.

4. Extensions to higher dimensions

In the preceding sections I have discussed the formation of estimators of the functions of the probability distributions and for their respective covariance matrices in the cases of discrete variables. These ideas also apply to higher dimensions. Dominance conditions for three dimensions have been provided by Muller and Trannoy (2003). For D dimensions Crawford (2005) shows that the first-order dominance conditions stem from:

$$\begin{aligned}\Delta W &= -\sum_{i=1}^{D}\int_0^{a_i}\psi_i \Delta F_i dx_i + \sum_{i=1}^{D-1}\sum_{j=i+1}^{D}\int_0^{a_j}\int_0^{a_i}\psi_{ij}\Delta F_{ij} dx_i dx_j - \\ &\sum_{i=1}^{D-2}\sum_{j=i+1}^{D-1}\sum_{k=j+1}^{D}\int_0^{a_k}\int_0^{a_j}\int_0^{a_i}\psi_{ijk}\Delta F_{ijk} dx_i dx_j dx_k \cdots \\ &+ (-1)^D \int_0^{a_D}\cdots\int_0^{a_i}\psi_{i\ldots D}\Delta F dx_i \ldots dx_D\end{aligned} \tag{36}$$

For first-order dominance there are $2^D - 1$ partial derivatives of ψ to consider: D partial derivatives of the function with respect to individual dimensions, which are assumed positive by monotonicity, and $2^D - D - 1$ cross-partial derivatives to which signs have to be assigned. In order to implement the test for discrete variables to D dimensions I propose following these steps:

1. Stack the probabilities using as a guide a matrix of positions which is in turn the Kronecker product of the vectors of positions of individual variables. The vector of positions, vp_i simply shows the location of every probability in an ordinal way (i.e. the probability of having the lowest value comes first at the left corner, then followed by the second lowest and so on). Therefore if the number of possible values for variable x_i $(i = 1, ..., D)$ is C_i then the length of the vector of positions is C_i and its elements are the natural numbers from 1 to C_i, each with an indication that these numbers belong to variable i (this is helpful to pinpoint the position of the joint probabilities $\widehat{p_{jk...z}} = \frac{1}{n}\sum_{A=1}^{n} I\left(x_1^A = j \wedge x_2^A = k \ldots \wedge x_D^A = z\right)$, in the matrix of positions. For instance if we have three dimensions, $1, 2, 3$, with $C_1 = 2, C_2 = 2, C_3 = 3$, then the matrix of positions, MP, is calculated as follows:

$$MP = vp_1' \otimes vp_2 \otimes vp_3' \tag{37}$$

The actual order does not matter much as long as the rest of the procedure is kept consistently, but it is recommended to transpose non-adjacent vectors. The matrix of postions in (37) looks like:

$$MP = \begin{bmatrix} 1^1 1^2 1^3 & \cdots & 1^1 1^2 3^3 & \cdots & \cdots & 2^1 1^2 3^3 \\ 1^1 2^2 1^3 & \cdots & \cdots & \cdots & \cdots & 2^1 2^2 3^3 \end{bmatrix} \tag{38}$$

For instance the probability corresponding to the top right position in (38) is $\widehat{p_{123}} = \frac{1}{n}\sum_{A=1}^{n} I\left(x_1^A = 1 \wedge x_2^A = 2 \wedge x_3^A = 3\right)$. The matrix of probabilities, $\widehat{P^A}$, is therefore constructed following MP and then define $\widehat{V}$ as in (26).

1. In order to accumulate the probabilities to generate $\widehat{\Delta}^z, \widehat{\Delta}_1^z, \widehat{\Delta}_{1+2-12}^z$, as well as other functions relevant to cases with more than two variables, we pre-multiply and post-multiply $\widehat{V}$ by powers of the Kronecker products of sets of L_{Ci} and 1_{C_i} chosen especially to yield the linear combinations of the probabilities. The standard errors are then estimated following the same approach as in the previous section. For instance in the previous example:

$$\widehat{V} = \begin{bmatrix} \widehat{p_{111}^A} - \widehat{p_{111}^B} & \widehat{p_{112}^A} - \widehat{p_{112}^B} & \widehat{p_{113}^A} - \widehat{p_{113}^B} & \widehat{p_{211}^A} - \widehat{p_{211}^B} & \widehat{p_{212}^A} - \widehat{p_{212}^B} & \widehat{p_{213}^A} - \widehat{p_{213}^B} \\ \widehat{p_{121}^A} - \widehat{p_{121}^B} & \widehat{p_{122}^A} - \widehat{p_{122}^B} & \widehat{p_{123}^A} - \widehat{p_{123}^B} & \widehat{p_{221}^A} - \widehat{p_{221}^B} & \widehat{p_{222}^A} - \widehat{p_{222}^B} & \widehat{p_{223}^A} - \widehat{p_{223}^B} \end{bmatrix}$$

Since the number of rows in the example is determined solely by x_2 and it is equal to $C_2 = 2$, while the number of columns is determined by x_2 and x_3, being equal to

$C_1C_3 = 6$, $\widehat{\Delta}^z$ is computed as:

$$\widehat{\Delta}^z = [L_{C_2}]^Z \widehat{V} [L_{C_1} \otimes L_{C_3}]'^Z \tag{39}$$

The standard errors are found by taking the square root of the diagonal of the covariance matrix $var\left(\widehat{\Delta}^Z\right)$:

$$var\left(\widehat{\Delta}^Z\right) = [L_{C_1} \otimes L_{C_3}]^Z \otimes [L_{C_2}]^Z \left(\frac{n^A + n^B}{n^A n^B}\widehat{\Omega}\right) [L_{C_1} \otimes L_{C_3}]'^Z \otimes \left[L'_{C_2}\right]^Z \tag{40}$$

Crawford (2005) warns that extension to D dimensions demands considering two problems: one of interpretability and one of data requirements. Regarding interpretability, as the number of dimensions increase, so does the number of cross-partial derivatives, which calls for difficult interpretations of their signs. However this problem is ameliorated when dominance analysis focuses on welfare functions to which more structure has been imposed. Usually this means, conciously or not, attributing signs to cross-partial derivatives or just making them equal to zero. Such attributions often are reflected in separability properties of the functions. As for data requirements, the problem is that sample size requirements to maintain precision in the nonparametric estimation of multivariate distributions increase exponentially as the number of dimensions increase. In the non-parametric statistics literature this a manifestation of Bellman's "curse of dimensionality" (e.g. see Silverman, 1986). For instance if there are $n = 1000$ observations distributed uniformly over a $5D$ hypercube, $[0,1]^5$ the expected number of observations in the neighborhood of one of its 0.2^5 is $n\ 0.2^5 = 0.32$, i.e. less than one observation. To get 50 data points upon which to base an estimate of a cell frequency one needs to average over a 0.55^5 cube. Hence as the number of dimensions rises either increasing sample size or taking larger neighbourhoods become necessary.

5. Empirical application

I test for stochastic dominance with three dimensions using data from Peru. As with the bivariate case (and any other), several dominance conditions can be derived. We test one of the first-order dominance conditions for three dimensions:

$$\psi \in \Psi^{FO3D-} \leftrightarrow \left(\begin{array}{c} \Delta W \geq 0 \leftrightarrow \Delta F_i(x_i) \leq 0 \forall i = 1,2,3 \wedge \Delta F_{ij}(x_i, x_j) \leq 0 \forall i,j = 1,2,3 \\ \wedge \Delta F_{123}(x_1, x_2, x_3) \leq 0 \quad \forall x_1, x_2, x_3 \end{array}\right) \tag{41}$$

$$\Psi^{FO3D-} == \left\{\psi\colon \psi_i, \geq 0 \forall i = 1,2,3,\ \psi_{ij} \leq 0 \forall i,j = 1,2,3,\ \psi_{123} \leq 0\right\}$$

Condition (41) stems from setting $D = 3$ in equation (36):

$$\Delta W = -\sum_{i=1}^{3} \int_0^{a_i} \psi_i \Delta F_i dx_i + \sum_{i=1}^{2} \sum_{j=i+1}^{3} \int_0^{a_j} \int_0^{a_i} \psi_{ij} \Delta F_{ij} dx_i dx_j - \int_0^{a_1} \int_0^{a_2} \int_0^{a_3} \psi_{123} \Delta F dx_3 dx_2 dx_1 \tag{42}$$

Notice that $\Delta F_{123}(x_1, x_2, x_3) \leq 0 \quad \forall x_1, x_2, x_3$ implies $\Delta F_{ij}(x_i, x_j) \leq 0 \forall i, j = 1, 2, 3$ and $\Delta F_i(x_i) \leq 0 \forall i = 1, 2, 3$. Therefore a test on $\Delta F_{123}(x_1, x_2, x_3)$ suffices to ascertain stochastic dominance over the class of welfare functions considered.

5.1. Data

The wellbeing of Peruvian adults at least 25 years old living in Lima is compared against that of adults living elsewhere in the country. I use the Peruvian National Household Survey, ENAHO 2001, covering 16,515 households.For illustrative purposes, the comparison is over the following three discretely measured dimensions:

- School attendance by type of school: Calonico and Nopo (2007) document a positive earnings return to having attendend private school in Peru. This variable takes three values: 0 if the person did not attend school, 1 if the person went to public school and 2 if the person went to private school.
- Crime victimization: It is a binary variable. The question in the survey is whether the respondent or any other household member was the victime of the crime during the past year. The assumption behind using this variable is that being a victime of crime, and/or being a relative of a crime victim, has a negative impact on an individual's welfare. The value of 1 is attributed to those whose households and/or selves did not undergo crime victimization.
- Serious accident or illness: Also a binary variable. The question is whether the respondent or any other household member experienced a serious accident or illness during the past year. The assumption is that having suffered a serious accident or illness and/or

having a relative suffering from such events affects individual welfare negatively. The value of 1 is attributed to those whose households and/or selves did not suffer a serious accident or illness.

The respective sample sizes are in tables A-1a through A-2c. Denoting school attendance by T, crime victimization by C and serious accident or illness by S, we constructed the following matrix of positions:

$0^S0^C0^T$	$0^S0^C1^T$	$0^S0^C2^T$	$1^S0^C0^T$	$1^S0^C1^T$	$1^S0^C2^T$
$0^S1^C0^T$	$0^S1^C1^T$	$0^S1^C2^T$	$1^S1^C0^T$	$1^S1^C1^T$	$1^S1^C2^T$

The respective joint distributions are:

Peruvian adults in Lima					
0.001	.0203	0.005	0.004	0.135	0.045
0.002	.0590	0.015	0.025	0.554	0.133
Peruvian adults outside Lima					
0.002	0.012	0.001	0.013	0.093	0.008
0.010	0.050	0.004	0.113	0.658	0.036

The joint distribution tables should be read using the matrix of positions. For instance it is interesting to note that both in Lima and elsewhere in Peru the percentage of adults who reported crime victimization and serious illness is highest among those who attended public school vis-a-vis those who attended private school and those who did not go to school. In the case of Lima it is 2% versus 0.5% (private school) and 0.1% (those who did not go to school) and in the case of the rest of the country it is 1.2% versus 0.1% (private school) and 0.2% (no school). (See three top-left cells of the respective tables).

5.2. Results

Since I am testing for first-order stochastic dominance using $\Delta F_{123}(x_1, x_2, x_3)$, the respective matrix yields seven contrasts for testing because $\Delta F_{123}(a_1, a_2, a_3) = 0$. The matrix of differences is constructed so that the joint distribution of Peruvians living outside Lima is subtracted from that of Peruvians living in Lima. The seven z-statistics located according to the matrix of positions are in Table 1:

Table 1: Z-statistics of the 3-D first-order stochastic dominance test for class Ψ^{FO3D-}. Ho: $\Delta F_{123}(x_1, x_2, x_3) = 0 \forall x_1, x_2, x_3$.Peruvian adults living in Lima minus those living elsewhere in the country[14]

-1.271073	3.9410663	13.986603	-3.8936571	9.0505295	15.979511
2.4095175	2.2402767	12.671916	-18.167451	-23.950389	N/A

For a two-tailed test with 10 contrasts and infinite degrees of freedom (as is the case with a z-statistic when using t-statistics tables) the value of the 0.01 upper point of the studentized maximum modulus (SMM) distribution is 3.289 (Stoline and Ury, 1979). Since there are 11 contrasts one should use a higher value, e.g. 3.3. For the alternative hypothesis of indeterminacy, i.e. crossing of the joint distribution hypersurfaces at least one of the statistics has to be higher than 3.3 and another one lower than -3.3 (with about 98% of confidence). In this empirical application that is the case since one statistic is equal to less than -23 and another one is almost 16. Therefore I conclude that the hypersurfaces cross and there is no dominance relationship between adults living in Lima and adults living elsewhere in Peru with respect to the three variables in question, in terms of class Ψ^{FO3D-}. I also show other dominance tests with pairs of these variables as well as with the variables taken individually one at a time in order to explore (not exhaustively) whether other dominace orderings are possible when the range of variables is restricted. Such exercise can also shed light on the source of curve-crossing demonstrated by the 3-D test.

Tables 2, 3 and 4 show the z-statistics for the tests on $\Delta F_{12}(x_1, x_2)$ to ascertain first-order stochastic dominance over the class Ψ^{-}, which assumes substitutability across dimensions. The tests are performed for the three possible combination pairs of the three variables used. The respective critical values for 98% of confidence from the SMM distribution for three contrasts (since again $\Delta F_{12}(a_1, a_2) = 0$) are -2.934 and 2.934. For five contrasts (which is the case of combinations of the schooling variable with any of the others), the values are slightly below -3.143 and 3.143 (these are the values for 6 contrasts available in Stolien and Ury, 1979). The evidence is in favour of rejecting homogeneity for the alternative hypothesis of curve-crossing for school attendance and crime victimization (Table 2), i.e. no dominance ordering for the two samples is possible for welfare measures belonging to class Ψ^{-}. A similar

[14] The z-statistics are ordered and displayed according to the matrix of positions. For instance the statistic equal to 2.4095175 stems from the difference between the two samples in the cumulative probabilities of not having experienced a serious accident ($S = 0$) *and* not having attended school ($T = 0$) *and* not having experienced crime victimization *or* having experienced crime victimization ($C \leq 1$).

result in favour of the alternative of indeterminacy, i.e. curve-crossing, holds for the combination of school attendance and serious illness/accident (Table 3). By contrast, the null hypothesis of homogeneity is rejected in favour of the alternative hypothesis of first-order stochastic dominance of the sample of adults living outside Lima over those living in Lima for the combination of serious illness/accident and crime victimization (Table 4). Such ordering holds for all welfare functions belonging to class Ψ^{-}.

Table 2: Z-statistics of the 2-D first-order stochatic dominance tests for class Ψ^{-}. Ho: $\Delta F_{12}(x_1, x_2) = 0 \forall x_1, x_2$. School attendance and crime victimization[15]

-5.708766982	8.27440639	15.97951227
-22.06532709	-39.15119123	N/A

Table 3: Z-statistics of the 2-D first-order stochatic dominance tests for class Ψ^{-}. Ho: $\Delta F_{12}(x_1, x_2) = 0 \forall x_1, x_2$. Private school attendance and serious illness/accident

-5.539818323	2.240277135	6.001220773
-22.06532709	-39.15119123	N/A

Table 4: Z-statistics of the2-D first-order stochatic dominance tests for class Ψ^{-}. Ho: $\Delta F_{12}(x_1, x_2) = 0 \forall x_1, x_2$. Serious illness/accident and crime victimization

6.021721854	15.97951227
6.001220773	N/A

Finally, univariate first-order dominance tests are performed for each variable. With critical values for 98% of confidence for two contrasts (around -2.9 and 2.9; see Stoline and Ury, 1979) and the critical values for one contrast and 95% of confidence (-1.96 and 1.96) Peruvians living in Lima dominate those living outside Lima regarding school attendance but are dominated by the latter regarding both crime victimization and serious illness/accident (Table 5). These results illustrate the point made by other authors (e.g. Duclos *et al.*, 2006) whereby studying dominance over individual dimensions may lead to different ordering possibilities from those attainable by studying multivariate dominance. But also these results help to explain the three-dimensional curve-crossing of this empirical application: there is at least one dimension in which each sample dominates the other one.

[15] This table is read using the same reasoning as with Table 1, but know the category of serious illness/accident has been integrated out. For instance, the value of -39.15119123 is that of the statistic corresponding to the difference between samples of the cumulative probability of not having been victimized *or* having been victimized ($C \leq 1$) and not having attended school or having attended public school ($T \leq 1$). The remaining tables are read accordingly.

Table 5: Z-statistics of the 1-D first-order stochatic dominance tests for class Φ^1(see Table 2). Ho: $\Delta F_i(x_i) = 0 \forall i = 1, 2, 3$.

Private school	Crime victimization	Serious illness/accident
-22.06532709	15.97951227	6.001220773
-39.15119123	N/A	N/A

6. Conclusions and further research

This paper's first objective has been to motivate the need to find tests of multidimensional stochastic dominance to a broad range of conditions based on extensions of the conditions compiled and expanded by Atkinson and Bourguignon (1982) and Crawford (2005). In this paper I compiled these conditions up to third order of dominance and to D dimensions (for first-order dominance) and then proceeded to show that many more conditions, beyond those originally derived by Atkinson and Bourguignon (1982), can be obtained, for instance by increasing the order of dominance over one dimension at a time. This approach was originally pioneered by Duclos et al. (2006), but in this paper I showed that the conditions of Duclos *et al.* (2006) can be replicated and then conditions representing complementarity between attributes can be incorporated, which were missing in Duclos et al. (2006). These conditions are meant for welfare functions but with some adjustments (e.g. reversing the sign of the derivatives, incorporating poverty lines) they are applicable to poverty functions as well. With this argument of completeness this paper motivates the need to find tests for multidimensional stochastic dominance, at least for the "missing conditions", i.e. those not covered by Duclos et al. (2006).

The second part of the paper proposes an encompassing test based on a multivariate extension of Anderson's (1996) non-parametric test. This test is only applied to discrete, ordinal variables since it would not be consistent for continuous variables. I show how the test works and how to implement it for D dimensions using an original approach of a *matrix of positions*, which indicates how to stack in an orderly manner probabilities from several dimensions in a matrix. The test was applied to first-order dominance over three dimensions using data from Peru. The empirical example illustrates the well-established point that the choice of number of dimensions may generate alternative dominance orderings. It also illustrates that by looking at dominance over fewer dimensions one can explain failure to reach partial orderings when more dimensions are considered. In the Peruvian case the fact that one sample dominated in terms of school attendance but the other sample dominated over crime victimization and serious illness/accident helped to explain why dominance over the three variables jointly was rejected.

Pending research should focus on finding new tests for continuous variables and for combinations of continuous and discrete variables. Duclos *et al.* (2007) are pioneering this approach with an extension of their test (Duclos *et al.*, 2006) to combinations of discrete and continuous variables. They have so far the only available multidimensional tests for continuous variables and combinations of continuous and discrete variables, but they have not been shown yet to apply to the "missing conditions". It could be applied if a gap representation can be shown for objects like $\Delta K(x_1, x_2)$ and its linear combinations. Should that be the case both the test by Duclos *et al.* (2006) and that of Barret and Donald (2003) or even Bennett (2008) would be applicable. Alternatively a kernel estimation approach may be necessary. The extension of tests like that of Linton et al. (2008) to multiple dimensions looks like a promising route to complete this part of the multidimensional stochastic dominance testing literature.

REFERENCES

Anderson, Gordon (1996), "Nonparametric tests of stochastic dominance in income distributions", *Econometrica*, 64: 1183-93.

Atkinson, Anthony B. and Bourguignon, Francois (1982), "The comparison of multi-dimensioned distributions of economic status", *Review of Economic Studies*, XLIX: 183-201.

Bennett, Christopher (2007), "Consistent tests for completely monotone stochastic dominance", manuscript.

Bennett, Christopher (2008), "Consistent integral-type tests for stochastic dominance", manuscrpit.

Barret, Gary and Stephen Donald (2003), "Consistent tests for stochastic dominance", *Econometrica*, 71: 71-104.

Bishop, J.A., Chakarborti, S and Thistle, P.D. (1989), "Asymptotically distribution-free statistical inference for generalized Lorenz curves", *Review of Economics and Statistics*, 71: 725-7.

Bourguignon, Francois and Satya Chakravarty (2002), "Multi-dimensional poverty orderings", DELTA, Paris.

Chen, Wen-Hao and Jean-Yves Duclos (2008), "Testing for poverty dominance: an application to Canada", CIRPEE Working Paper, 08-36.

Crawford, Ian (2005), "A nonparametric test of stochastic dominance in multivariate distributions", manuscript.

Calonico, Sebastian and Hugo Nopo (2007), "Where did you go to school? Private-public differences in schooling trajectories and their role on earnings", *Well-being and social policy*, 3(1): 25-46.

Chambaz, Christine and Eric Maurin (1998), "Atkinson and Bourguignon's dominance criteria: extended and applied to the measurement of poverty in France", *Review of Income and Wealth*, 44(4): 497-513.

Cowell, Frank and Maria-Pia Victoria-Feser (2007), "Robust stochastic dominance: a semiparametric approach", *Journal of Economic Inequality*, 5: 21-37.

Dardanoni, Valentino and Antonio Forcina (1999), "Inference for Lorenz curve orderings", *Econometrics Journal*, 2: 49-75.

Davidson, Russell and Jean-Yves Duclos (2000), "Statistical inference for stochastic dominance and for the measurement of poverty and inequality", *Econometrica*, 68: 1435-64.

Deaton, Angus S. and John Muellbauer (1980), *Economics and Consumer Behaviour*, Cambridge: Cambridge University Press.

Duclos, Jean-Yves and Paul Makdissi (2005), "Sequential stochastic dominance and the robustness of poverty orderings", *Review of Income and Wealth*, 51(1): 63-87.

Duclos, Jean-Yves, David Sahn and Stephen Younger (2007), "Robust multidimensional poverty comparisons with discrete indicators of well-being" in Jenkins, Stephen and John Micklewright (editors) , *Inequality and Poverty Re-examined*, Oxford: Oxford University Press.

Duclos, Jean-Yves, David Sahn and Stephen Younger (2006), "Robust multidimensional poverty comparisons", *The Economic Journal*, 116(514): 943-68.

Goodman, A., P. Johnson and S. Webb (1997), *Inequality in the UK*, Oxford: Oxford University Press.

Gregg, P. and J. Wadsworth (1996), "More work in fewer households" in Hills, J. (editor), *New Inequalities: the changing distribution of income and wealth in the United Kingdom*, Cambridge: Cambridge University Press.

Hadar, J. and W. R. Russell (1974), "Stochastic dominance in choice under uncertainty", in Balch, M.S., D.L. McFadden and S.Y. Wu (editors), *Essays on Economic Behaviour under Uncertainty*, Amsterdam: North-Holland.

Harrison, A. (1982), "A tale of two distributions", *Review of Economic Studies*, 48:621-31.

Harsanyi, J.C. (1955), "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility", *Journal of Political Economy*, 63:

Hicks, John (1940), "The valuation of the social income", *Economica*, 7: 105-24.

Howes, Stephen (1996), "A new test for inferring dominance from sample data", Discussion Paper, STICERD, London School of Economics.

Ibbott, P. (1998), "Intergenerational changes in the disttribution of consumption", unpublished manuscript.

Kendall, Maurice G. and Stewart, Allan (1979), *The Advanced Theory of Statistics*, London: Griffin.

Levy, H. and Paroush J. (1974), "Toward multivariate efficiency criteria", *Journal of Economic Theory*, 7: 129-42.

Linton, Oliver, Esfandiar Maasoumi and Yoon-Jae Whang (2005), "Consistent testing for stochastic dominance under general sampling schemes", *The Review of Economic Studies*, 72(3): 735-65.

Maasoumi, Esfandiar and Almas Heshmati (?), "Stochastic dominance amongst Swedish income distributions", manuscript.

Magnus, J.R. and Neudecker H. (1988), *Matrix Differential Calculus*, New York: Wiley.

Muller, Christophe and Alain Trannoy (2004), "A dominance approach to well-being inequality across countries", IVIE Working Paper, 2004-27.

Neary, J.P. and Roberts K.W.S. (1980), "The theory of household behaviour under rationing", *European Economic Review*, 13: 25-42.

Rotbharth, E. (1941), "The measurement of changes in real income under conditions of rationing", *Review of Economic Studies*, 8: 100-7.

Sen, Amartya K. (1979), "The welfare basis of real income comparisons", *Journal of Economic Literature*, 17: 1-45.

Sen, Amartya K. (2001), *Development as Freedom*, Oxford: Oxford University Press.

Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis,* Monographs on Statistics and Applied Probability, London: Chapman & Hall.

Stoline, M.R. and H.A. Ury (1979), "Tables of the Studentised Maximum Modulus Distribution and an application to multiple comparsions among means", *Technometrica*, 21: 87-93.

Thuysbaert, Bram and Ricardas Zitikis (2005), "Consistent testing for poverty dominance", WIDER Research Paper 2005/64.

Ulph, D.T. (1978), "On labour supply and the measurement of inequality", *Journal of Economic Theory*, 19: 492-512.

Xu, Kuan and Lars Osberg (1998), "A distribution-free test for deprivation dominance", *Econometric Reviews*, 17(4): 415-29.

This preprint was prepared with the AAS LATEX macros v5.2.

Table 6: Bivariate dominance criteria, null and alternative hypotheses

Null	**Alternative**	**Class Ψ**
First order dominance		
$Ho: \Delta F_i(.), \Delta F(.) = 0$	$H_{1F}: \Delta F_i(.), \Delta F(.) \leq 0$	Ψ^-
$Ho: \Delta F_i(.), \Delta K(.) = 0$	$H_{1K}: \Delta F_i(.), \Delta K(.) \leq 0$	Ψ^+
$Ho: \Delta F_i(.), \Delta F(.), \Delta K(.) = 0$	$H_1: H_{1F} \wedge H_{1K}$	Ψ^1
2nd order dominance		
$Ho: \Delta H_i(.), \Delta H(.) = 0$	$H_{2H}: \Delta H_i(.), \Delta H(.) \leq 0$	Ψ^{--}
$Ho: \Delta H_i(.), \Delta L(.) = 0$	$H_{2L}: \Delta H_i(.), \Delta L(.) \leq 0$	Ψ^{++}
$Ho: \Delta H_i(.), \Delta H(.), \Delta L(.) = 0$	$H_2: H_{2H} \wedge H_{2L}$	Ψ^2
3nd order dominance		
$Ho: \Delta J_i(.), \Delta J(.) = 0$	$H_{3J}: \Delta J_i(.), \Delta J(.) \leq 0$	Ψ^{---}
$Ho: \Delta J_i(.), \Delta M(.) = 0$	$H_{3J}: \Delta J_i(.), \Delta M(.) \leq 0$	Ψ^{+++}
$Ho: \Delta J_i(.), \Delta J(.), \Delta M(.) = 0$	$H_3: H_{3J} \wedge H_{3M}$	Ψ^3
Example of condition (18)		
$Ho: \Delta F_2(.), \Delta H_1(.), \int_0^{x_1} \Delta F(t, x_2)\, dt = 0$	$Ha: \Delta F_2(.), \Delta H_1(.), \int_0^{x_1} \Delta F(t, x_2)\, dt \leq 0$	Ψ^{-1}

Table 7: Univariate dominance criteria, null and alternative hypotheses

Null	**Alternative**	**Class Φ**
First order dominance		
$Ho: \Delta F_i(.) = 0$	$H_1: \Delta F_i(.) \leq 0$	Φ^1
2nd order dominance		
$Ho: \Delta H_i(.) = 0$	$H_2: \Delta H_i(.) \leq 0$	Φ^2
3nd order dominance		
$Ho: \Delta J_i(.) = 0$	$H_3: \Delta J_i(.) \leq 0$	Φ^3

$$\Phi^1 = \{\psi \colon \psi_i \geq 0\}; \Phi^2 = \{\psi \colon \psi_i \geq 0, \psi_{ii} \leq 0\}; \Phi^3 = \{\psi \colon \psi_i \geq 0, \psi_{ii} \leq 0, \psi_{iii} \leq 0\}$$

7. Appendix 1:

Derivation of condition (18):

$$\psi \in \Psi^{-1} \leftrightarrow \left(\Delta W \geq 0 \leftrightarrow \Delta F_2 (x_2) \leq 0 \wedge \Delta H_1 (x_1) \leq 0 \wedge \int_0^{x_1} \Delta F (t, x_2)\, dt \leq 0 \quad \forall x_1, x_2 \right)$$

where $\Psi^{-1} = \{\psi \colon \psi \in \Psi^- \wedge \psi_{11} \leq 0 \wedge \psi_{112} \geq 0\}$

We start by integrating by parts (8), first with respect to x_2 and then the two elements each with respect to x_1 which yields the condition from which the first-order dominance result (9) is deduced:

$$\begin{aligned} \Delta W &= -\int_0^{a_1} \psi_1 (x_1, a_2)\, \Delta F_1 (x_1)\, dx_1 - \int_0^{a_1} \psi_2 (a_1, x_2)\, \Delta F_2 (x_2)\, dx_2 \\ &\quad + \int_0^{a_2} \int_0^{a_1} \psi_{12} (x_1, x_2)\, \Delta F (x_1, x_2)\, dx_1 dx_2 \end{aligned} \tag{43}$$

Define: $I \equiv -\int_0^{a_1} \psi_1 (x_1, a_2)\, \Delta F_1 (x_1)\, dx_1$, $II \equiv -\int_0^{a_1} \psi_2 (a_1, x_2)\, \Delta F_2 (x_2)\, dx_2$, $III \equiv \int_0^{a_2} \int_0^{a_1} \psi_{12} (x_1, x_2)\, \Delta F (x_1, x_2)\, dx_1 dx_2$. Integrating I by parts with respect to x_1 yields:

$$I = -\psi_1 (a_1, a_2)\, \Delta H_1 (a_1) + \int_0^{a_1} \psi_{11} (x_1, a_2)\, \Delta H_1 (x_1)\, dx_1 \tag{44}$$

Integrating also III by parts with respect to x_1 yields:

$$III = \int_0^{a_2} \psi_{12} (a_1, x_2) \left[\int_0^{a_1} \Delta F (t, x_2)\, dt \right] dx_2 - \int_0^{a_2} \int_0^{a_1} \psi_{112} (a_1, x_2) \left[\int_0^{x_1} \Delta F (t, x_2)\, dt \right] dx_1 dx_2 \tag{45}$$

Plugging results (44) and (45) into (43) yields the condition:

$$
\begin{aligned}
\Delta W &= -\psi_1(a_1,a_2)\,\Delta H_1(a_1) + \int_0^{a_1} \psi_{11}(x_1,a_2)\,\Delta H_1(x_1)\,dx_1 \\
&\quad - \int_0^{a_1} \psi_2(a_1,x_2)\,\Delta F_2(x_2)\,dx_2 \\
&\quad + \int_0^{a_2} \psi_{12}(a_1,x_2)\left[\int_0^{a_1} \Delta F(t,x_2)\,dt\right] dx_2 \\
&\quad - \int_0^{a_2}\int_0^{a_1} \psi_{112}(a_1,x_2)\left[\int_0^{x_1} \Delta F(t,x_2)\,dt\right] dx_1 dx_2
\end{aligned} \tag{46}
$$

Condition (18) ensues from the result in (46).

8. Appendix 2:

Table A-1a: Incidence of serious accident/illness and crime victimization

Peruvian adults who did not attend school and live in Lima

	Crime victim	No crime victim	Total
Serious accident/illness	5	14	19
No serious accidente/illness	24	140	164
Total	29	154	183

Table A-1b: Incidence of serious accident/illness and crime victimization

Peruvian adults who attended public school and live in Lima

	Crime victim	No crime victim	Total
Serious accident/illness	114	332	446
No serious accidente/illness	762	3,120	3,882
Total	876	3,452	4,328

Table A-1c: Incidence of serious accident/illness and crime victimization

Peruvian adults who attended private school and live in Lima

	Crime victim	No crime victim	Total
Serious accident/illness	29	84	113
No serious accidente/illness	255	749	1,004
Total	284	833	1,117

Table A-2a: Incidence of serious accident/illness and crime victimization

Peruvian adults who did not attend school and do not live in Lima

	Crime victim	No crime victim	Total
Serious accident/illness	44	271	315
No serious accidente/illness	357	3,092	3,449
Total	401	3,363	3,764

Table A-2b: Incidence of serious accident/illness and crime victimization

Peruvian adults who attended public school and do not live in Lima

	Crime victim	No crime victim	Total
Serious accident/illness	342	1,372	1,714
No serious accidente/illness	2,561	18,063	20,624
Total	2,903	19,435	22,338

Table A-2c: Incidence of serious accident/illness and crime victimization

Peruvian adults who attended private school and do not live in Lima

	Crime victim	No crime victim	Total
Serious accident/illness	24	101	125
No serious accidente/illness	220	985	1,205
Total	244	1,086	1,330