C.I.R.P.É.E

Centre Interuniversitaire sur le Risque, les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **10-04**

Comparing Multidimensional Poverty with Qualitative Indicators of Well-Being

Yélé Maweki Batana

Jean-Yves Duclos

Février/February 2010

Batana : Centre de Recherche Léa-Roback, Direction de la santé publique, Montréal (Québec) H2L 1M3 and Université de Montréal

yele.maweki.batana@umontreal.ca

Duclos : Département d'économique and CIRPÉE, Pavillon DeSève, Université Laval, Québec, Canada G1V 0A6

jyves@ecn.ulaval.ca

We are grateful to Canada's SSHRC, to Québec's FQRSC and to the *Programme canadien de bourses de la francophonie* for financial support. This work was also carried out with support from the Poverty and Economic Policy (PEP) Research Network, which is financed by the Government of Canada through the International Development Research Centre (IDRC) and the Canadian International Development Agency (CIDA), and by the Australian Agency for International Development (AusAID).

Abstract:
This paper examines multidimensional stochastic dominance when one of the indicators of well-being, such as household size or place of residence, is qualitative. It also uses a test for strict dominance based on the empirical likelihood ratio. Empirical applications are based on the DHS (Demography and Health Surveys) for several countries in Western Africa. The results show the existence of multidimensional dominance relationships between most of these countries.

Keywords: Stochastic dominance, multidimensional poverty, empirical likelihood tests, bootstrap tests, West Africa

JEL Classification: C10, C11, C12, C30, C39, I32

1 Introduction

Poverty has been increasingly recognized as a multidimensional phenomenon, in large part due to the influential work of Sen (1979, 1985, 1987). A number of variables other than income can indeed provide important information on well-being and poverty, such as the state of health, the level of education, ownership of durable goods, access to basic services, *etc*. One set of variables that affects well-being deals with the size and composition of the households in which individuals live. Other variables are more geographical or describe the nature of one's area of residence (*e.g.*, rural *vs*. urban area).

Many of those indicators of well-being other than income are qualitative and discrete, although many of the multidimensional poverty indices found in the literature have been implicitly or explicitly designed with continuous indicators of well-being in mind. There have, however, been a few attempts at handling discrete indicators of well-being, including in the context of multidimensional dominance analysis. Atkinson (1992), Atkinson and Bourguignon (1987), Bourguignon (1989), Jenkins and Lambert (1993), Chambaz and Maurin (1998) and Duclos and Makdissi (2005) provide for instance conditions explicitly designed for making welfare and poverty dominance comparisons that take into account differences in discrete measures of household size and composition.

Most of the above studies have focused on deriving theoretical conditions for dominance and have usually performed poverty comparisons without taking account sampling variability. One exception is Duclos, Sahn, and Younger (2007). The current paper extends this work by applying statistical inference techniques based on an intersection-union approach to testing dominance, an approach that facilitates inference of dominance (as opposed to non-dominance) relationships. This also extends to the presence of a discrete indicator of multidimensional well-being the work of Davidson and Duclos (2006) on the use of an empirical likelihood ratio statistic.

Section 2 describes the conditions for dominance in welfare, while Section 3 examines dominance in poverty. These conditions are analogous to those found in Atkinson and Bourguignon (1987), Jenkins and Lambert (1993), Chambaz and Maurin (1998), and Duclos, Sahn, and Younger (2007). Section 4 presents an inference methodology based on the empirical likelihood ratio and on resampling techniques. Section 5 features a few empirical illustrations of poverty comparisons based on the proposed method and in the context of West African countries. The last section concludes.

2 Stochastic dominance in welfare

This section draws importantly on the work of Atkinson and Bourguignon (1987), Jenkins and Lambert (1993), and Chambaz and Maurin (1998). We assume a heterogeneous population that can be subdivided into K subgroups along the values of a discrete indicator. We also postulate that these subgroups can be ranked in descending order according to their marginal utility of income x. Let $u_k(x)$ be the concave and increasing utility function of the subgroup k of households, let ϕ_k^F be the proportion of households of subgroup k in a total population denoted F, such that $\sum_{k=1}^{K} \phi_k^F = 1$, and let $F_k(x)$ be the cumulative density function of households in each subgroup k of a population F. The social utility function (W) for this population is given by:

$$W^F = \sum_{k=1}^{K} \phi_k^F \int_0^{\overline{x}} u_k(x) dF_k(x)\, dx, \tag{1}$$

where $\overline{x}$ corresponds to the highest possible income in the population and where $F_k(\overline{x}) = 1 \;\forall\; k = 1, \ldots, K$. We assume that

$$u_1^{(1)}(x) \geq u_2^{(1)}(x) \geq \cdots \geq u_K^{(1)}(x) \geq 0 \;\forall\; x. \tag{2}$$

This says that, at the same income level x, the subgroups with greater needs (subgroups are ordered so that 1 is most needy and K is least needy) value an increase in their income more than less needy subgroups. We also assume that (Jenkins and Lambert 1993, Chambaz and Maurin 1998):

$$u_1(\overline{x}) = u_2(\overline{x}) = \cdots = u_K(\overline{x}). \tag{3}$$

Consider then another population G whose cumulative density function for K subgroups is given by $G_k(x)$. Let ϕ_k^G represent the proportion of households in each subgroup k in population G. The condition for first-order sequential welfare dominance of F by G is given by Proposition 1.

Proposition 1 *$W^G > W^F \;\forall\; u_k(x)$, with $k = 1, \ldots, K$, satisfying assumptions (2) and (3) if and only if $\sum_{k=1}^{L} \left[\phi_k^F F_k(x) - \phi_k^G G_k(x)\right] > 0 \;\forall\; x$ and $\forall\; L = 1, \ldots, K$.*

For a proof, see Chambaz and Maurin (1998), who draw on the demonstration in Atkinson and Bourguignon (1987) but relax the constraint that $\phi_k^F = \phi_k^G$.

To study second-order sequential dominance, we can postulate that:

$$u_1^{(2)}(x) \leq u_2^{(2)}(x) \leq \cdots \leq u_K^{(2)}(x) \leq 0 \;\forall\; x. \tag{4}$$

This says that marginal utility declines more rapidly in needier subgroups. A condition for second-order sequential welfare dominance of F by G is then given by Proposition 2.

Proposition 2 *$W^G > W^F$ $\forall$ $u_k(x)$, with $k = 1, \ldots, K$, satisfying Assumptions (2), (3) and (4) if and only if $\sum_{k=1}^{L} \int_0^x \left[\phi_k^F F_k(y) - \phi_k^G G_k(y)\right] dy > 0$ $\forall$ x and $\forall$ $L = 1, \ldots, K$.*

Jenkins and Lambert (1993) demonstrate the sufficiency of this condition for the general case in which ϕ_k^F and ϕ_k^G are not necessarily equal, and Chambaz and Maurin (1998) show its necessity.

3 Stochastic dominance in poverty

Some further assumptions are needed before we can establish conditions for dominance in poverty. Consider first the following sets of sequential poverty thresholds:

$$Z_1 \geq Z_2 \geq \cdots \geq Z_k. \tag{5}$$

This means that the poverty lines of the various subgroups can be ranked in decreasing order, from the most needy to the least needy. In the case of groups differentiated by household size, the neediest subgroup consists of the largest households (subgroup 1) while the least needy one is made of households of a single individual (subgroup K). In this case, it is reasonable to assume that a household in subgroup k has at least the same level of basic needs as one in subgroup $k+1$ and, that we can therefore set for it a poverty line at least as high as that of subgroup $k+1$.

Define the poverty index in each subgroup k as $P_k(Z_k)$:

$$P_k(Z_k) = \int_0^{Z_k} \pi_k(x)\, dF_k(x)\, dx, \tag{6}$$

where $\pi_k(x)$ is the contribution to group k poverty of a household in subgroup k with income level x. We have $\pi_k(x) = 0$ if $x \geq Z_k$. It seems natural to suppose that, at a common income level x, an identical increase in income will affect the poverty of a subgroup k at least as much as that of a less needy subgroup $k+1$. This is captured by the following assumption:

$$\pi_1^{(1)}(x) \leq \pi_2^{(1)}(x) \leq \cdots \leq \pi_K^{(1)}(x) \leq 0, \ \forall\, x. \tag{7}$$

Since $\pi_k^1(x)$ is non-positive, this also implies the usual monotonicity axiom used in axiomatic definitions of poverty, which says that an increase in anyone's income should not increase total poverty. [1] Now consider the poverty index for the entire population:

$$P(Z) = \sum_{k=1}^{K} \phi_k \int_0^{Z_k} \pi_k(x)\, dF(x)\, dx = \sum_{k=1}^{K} \phi_k P_k(Z_k), \tag{8}$$

with $Z = (Z_1, Z_2, \ldots, Z_K)$. Let $P^F(Z)$ be the poverty measure for the entire population F, and $P^G(Z)$ that for population G, where $\Delta P(Z) = P^F(Z) - P^G(Z)$.

Let $\Pi^1(Z)$ represent the class of poverty measures satisfying the assumptions enumerated above, and define upper limits on the poverty thresholds $Z_1^+ \geq Z_2^+ \geq \cdots \geq Z_k^+$, such that $Z_k \leq Z_k^+, \ \forall\ k$. The condition for first-order poverty dominance of F by G is given by Proposition 3.

Proposition 3 $\Delta P(Z) > 0, \ \forall\ P(Z) \in \Pi^1(Z)$ *and* $\forall\ Z_k \in \left[0, Z_k^+\right], k = 1, \ldots, K$ *if and only if* $\sum_{k=1}^{L} \left[\phi_k^F F_k(Z_k) - \phi_k^G G_k(Z_k)\right] > 0, \quad \forall\ Z_k \in \left[0, Z_L^+\right]$ *and* $\forall\ L = 1, \ldots, K$.

See Jenkins and Lambert (1993) and Chambaz and Maurin (1998) for a proof.

Extending the analysis to second-order dominance, we can assume that

$$\pi_1^{(2)}(x) \geq \pi_2^{(2)}(x) \geq \cdots \geq \pi_K^{(2)}(x) \geq 0, \ \forall\ x. \tag{9}$$

Denote as $\Pi^2(Z)$ the class of poverty measures that also satisfy (9). Proposition 4 establishes the equivalence conditions for second-order dominance of F by G in poverty.

Proposition 4 *The following conditions are equivalent:*

(i) $\Delta P(Z) > 0, \ \forall\ P(Z) \in \Pi^2(Z)$ *and* $\forall\ Z_k \in \left[0, Z_k^+\right], k = 1, \ldots, K$

(ii) $\sum_{k=1}^{L} \left[\phi_k^F P_k^F(Z;1) - \phi_k^G P_k^G(Z;1)\right] > 0, \ \forall\ Z \in \left[0, Z_L^+\right]$ *and* $\forall\ L = 1, \ldots, K$.

with:

$$P_k^F(Z;1) = \int_0^Z (Z - x)\, dF_k(x) \ \text{ and } P_k^G(Z;1) = \int_0^Z (Z - x)\, dG_k(x) \tag{10}$$

Duclos and Makdissi (2005) provide a proof of this.

[1] See Bourguignon and Chakravarty (2002) and Tsui (2002) for instance for a discussion of this and other popular axioms.

4 Statistical inference

Consider a population of size N partitioned into K subgroups by a discrete indicator d (level of education, place of residence, household size, *etc.*). We also consider a continuous indicator of well-being x (income, expenses, *etc.*). Let p_i be the probability associated with household i, with $i = 1, \ldots, N$ and $\sum_{i=1}^{N} p_i = 1$, and let p_{ki} be the probability associated with household i relative to subgroup k, such that $p_i = \phi_k p_{ki}$. If we consider y to be some income level from the set $[0, \overline{x}]$, the *non-normalized* cumulative distribution function for a subgroup k can be estimated by:

$$\widehat{F}_k(y) = \sum_{i=1}^{N} p_i I(x_i \leq y, d_i = k), \tag{11}$$

where $I(\cdot)$ is an indicator function assuming the value of one when the condition holds, and of zero otherwise. Assume we have two populations F and G, of size N^F and N^G respectively, for which we want to compare the distributions over (x, d). Proposition 1 states the existence of first-order sequential welfare dominance of F by G in the samples if:

$$\sum_{k=1}^{L} \left[\widehat{F}_k(y) - \widehat{G}_k(y)\right] > 0 \ \forall\ y \text{ and } \forall\ L = 1, \ldots, K. \tag{12}$$

Thus, non-dominance would obtain when we have $\sum_{k=1}^{L} \left[\widehat{F}_k(y) - \widehat{G}_k(y)\right] \leq 0$ for at least one pair (y, L). The tests are thus performed on the null hypothesis of non-dominance versus the alternative hypothesis of dominance:

$H_0 : \sum_{k=1}^{L} \left[\widehat{F}_k(y) - \widehat{G}_k(y)\right] \leq 0$ for at least one pair (y, L)

versus

$H_1 : \sum_{k=1}^{L} \left[\widehat{F}_k(y) - \widehat{G}_k(y)\right] > 0 \ \forall\ y$ and $\forall\ L = 1, \ldots, K$.

Davidson and Duclos (2006) suggest a method based on the empirical likelihood ratio for comparing univariate distributions. This method is extended here to bivariate distributions for which one of the dimensions is discrete. The problem of maximizing the empirical likelihood function is given by:

$$\max_{P_i^F, P_j^G} \sum_{i=1}^{N^F} \log P_i^F + \sum_{j}^{N^G} \log P_j^G, \tag{13}$$

subject to:

$$\sum_i P_i^F = 1 \text{ and } \sum_j P_j^G = 1.$$

Solving this problem yields the maximum of the unconstrained empirical likelihood function $\widehat{EL}$. In a second step, equation (13) is maximized taking into account the additional constraint:

$$\sum_{i=1}^{N^F} P_i^F I(x_i^F \leq y, d_i^F \leq L) - \sum_{j=1}^{N^G} P_j^G I(x_j^G \leq y, d_j^G \leq L) \leq 0 \tag{14}$$

This yields the constrained empirical likelihood function $\widehat{EL_c}$. The likelihood ratio (LR) is obtained by multiplying the difference between $\widehat{EL}$ and $\widehat{EL}_c$ by 2. It is given by:

$$LR(y,L) = 2\left\{\begin{array}{c} N\log N - N_F \log N_F - N_G \log N_G \\ +N_F(y,L)\log N_F(y,L) + N_G(y,L)\log N_G(y,L) \\ +M_F(y,L)\log M_F(y,L) + M_G(y,L)\log M_G(y,L) \\ -\left[N_F(y,L)+N_G(y,L)\right]\log\left[N_F(y,L)+N_G(y,L)\right] \\ -\left[M_F(y,L)+M_G(y,L)\right]\log\left[M_F(y,L)+M_G(y,L)\right] \end{array}\right\} \tag{15}$$

where $N_h(y,L) = \sum_{i=1}^{N_h I_i(y,L)}$, $M_h(y,L) = N_h - N_h(y,L)$ and $I_i(y,L) = I(x_i^h \leq y, d_i^h \leq L)$, with $h = F, G$. When dealing with sequential dominance in poverty, constraint (14) can be rewritten as

$$\sum_{i=1}^{N^F} P_i^F I(x_i^F \leq Z_L, d_i^F \leq L) - \sum_{j=1}^{N^G} P_j^G I(x_j^G \leq Z_L, d_j^G \leq L) = 0, \text{ with } Z_L \in [0, Z_L^+]. \tag{16}$$

As shown in Davidson and Duclos (2006), the square of the t-statistic from Kaur, Prakasa Rao, and Singh (1994) is asymptotically equivalent to the LR statistic and can be used in its place. The result extends easily to the case of bivariate distributions. The expression for the squared t-statistic is:

$$t^2(y,L) = \frac{N_F N_G \left(\widehat{F}(y,L) - \widehat{G}(y,L)\right)^2}{N_G \widehat{F}(y,L)\left(1-\widehat{F}(y,L)\right) + N_F \widehat{G}(y,L)\left(1-\widehat{G}(y,L)\right)}. \tag{17}$$

The equivalence between the two statistics is established in Proposition 5.

Proposition 5 *Under the assumption that* $F(y,L) = G(y,L)$*, with* $F(y,L) = \sum_{k=1}^{L} F_k(y)$ *and* $G(y,L) = \sum_{k=1}^{L} G_k(y)$*, when* $N \to \infty$*, then* $LR(y,L)$ *is equivalent to* $t^2(y,L)$*.*

Proof. See the appendices. ∎

The advantage of using the likelihood ratio approach is that it also yields probabilities P_i^F and P_j^G that can be used to draw bootstrap samples under the null of non-dominance[2]. For first-order dominance, these probabilities can be found analytically. Expressions for P_i^F and P_j^G are indeed given by:

$$P_i^F = \frac{I_i(y,L)}{\theta} + \frac{(1 - I_i(y,L))}{\phi} \tag{18}$$

$$P_j^G = \frac{I_j(y,L)}{N_F + N_G - \theta} + \frac{(1 - I_j(y,L))}{N_F + N_G - \phi} \tag{19}$$

where

$$\theta = \frac{(N_F + N_G) \times N_F(y,L)}{N_F(y,L) + N_G(y,L)} \text{ and } \phi = \frac{(N_F + N_G) \times M_F(y,L)}{M_F(y,L) + M_G(y,L)}.$$

The statistics $LR(y,L)$ and $t(y,L)$ correspond to the respective minima over all possible combinations of (y,L). If this minimum is equal to zero, then the dominance curve given by $\left[\sum_i P_i^F I_i(y,L) - \sum_j P_j^G I_j(y,L)\right]$ assumes at least one negative value for certain pairs (y,L). Otherwise, dominance is possible and bootstrap tests are performed using the approach in Davidson and Duclos (2006). Using P_i^F and P_j^G, we generate 399 sample pairs for the bootstraps in the application below. A p-value can then be computed as the proportion of bootstrap samples for which the LR and t bootstrap statistics respectively exceed the minima $LR(y,L)$ and $t(y,L)$ obtained from the initial samples.

There is no analytical solution for stochastic dominance of higher orders. We can then consider a constraint of the form:

$$\sum_{i=1}^{N^F} P_i^F (y - x_i^F)^s I_i(y,L) - \sum_{j=1}^{N^G} P_j^G (y - x_j^G)^s I_j(y,L) \leq 0, \tag{20}$$

[2]The statistics $LR^{1/2}(y,L)$ and $t^{1/2}(y,L)$ are asymptotically pivotal. Indeed, under the null $F(y,L) = G(y,L)$, $LR^{1/2}(y,L)$ is equivalent to $t^{1/2}(y,L)$ which, under the same hypothesis, asymptotically follows $N(0,1)$. This is also valid for their minima.

where $s+1$ is the order of dominance. We let the Lagrangian ($\mathcal{L}$) associated with the maximization problem be:

$$\mathcal{L} = \sum_i \log P_i^F + \sum_j \log P_j^G + \lambda_F \left(1 - \sum_i P_i^F\right) + \lambda_G \left(1 - \sum_j P_j^G\right)$$
$$- \mu \left[\sum_{i=1}^{N^F} P_i^F \left(y - x_i^F\right)^s I_i(y,L) - \sum_{j=1}^{N^G} P_j^G (y - x_j^G)^s I_j(y,L)\right] \quad (21)$$

where λ_F, λ_G and $\mu \in R$ are Lagrange multipliers. There is no analytical solution for $s \geq 1$, but there are some transformations that simplify the computation of numerical solutions. For this, we first need the first-order condition from the preceding maximization problem:

$$\lambda_F + \lambda_G = N_F + N_G$$

$$P_i^F = \frac{1}{\lambda_F + \mu(y - x_i^F)^s I_i(y,L)} \text{ and } P_j^G = \frac{1}{N_F + N_G - \lambda_F - \mu(y - x_j^G)^s I_j(y,L)}. \quad (22)$$

We can now solve the problem by finding $\widehat{\lambda}_F$ and $\widehat{\mu}$:

$$(\widehat{\lambda}_F, \widehat{\mu}) = \underset{\lambda_F, \mu \in R}{\arg\min} - \sum_i \log \left[\lambda_F + \mu(y - x_i^F)^s I_i(y,L)\right]$$
$$- \sum_j \log \left[N_F + N_G - \lambda_F - \mu(y - x_j^G)^s I_j(y,L)\right] \quad (23)$$

Next we replace λ_F and μ by their estimates $\widehat{\lambda}_F$ and $\widehat{\mu}$ in equation (21) to obtain the probabilities $\widehat{P}_i^F$ and $\widehat{P}_j^G$. These probabilities are then used to compute LR. The statistic in which we are interested is given by the minimum over the set of statistics given by all possible pairs (y, L). Extending Davidson and Duclos (2006)'s methods, Davidson (2007) shows that, in the univariate case and under the null hypothesis of second-order non-dominance for at least one (y), this statistic is equivalent to the corresponding squared t-statistic. The extension to the multivariate case is immediate.

5 Empirical Results

We use data from the Demographic and Health Surveys (DHS) for six WAEMU countries (Benin, Burkina Faso, Côte d'Ivoire, Mali, Niger and Togo) to illustrate the existence of sequential stochastic dominance in the presence of a discrete indicator of well-being. The first variable we consider is a multidimensional index of wealth calculated using factor analysis. This is estimated using multiple correspondence analysis (MCA) based on 11 qualitative indicators about the ownership of durable goods (radio, television, refrigerator, bicycle, motorcycle, car) and the access to other goods and services (electricity, type of toilette, quality of flooring, potable water, education)[3]. As a discrete indicator, we use household size. We next partition the samples into 10 groups, so that the first consists of households with 10 or more individuals and the tenth includes only single-individual households. Another discrete indicator considered in the study is the area of residence, with division of the samples into two groups: those living in rural areas and those living in urban areas. The education level of the head of household is the third discrete indicator considered.

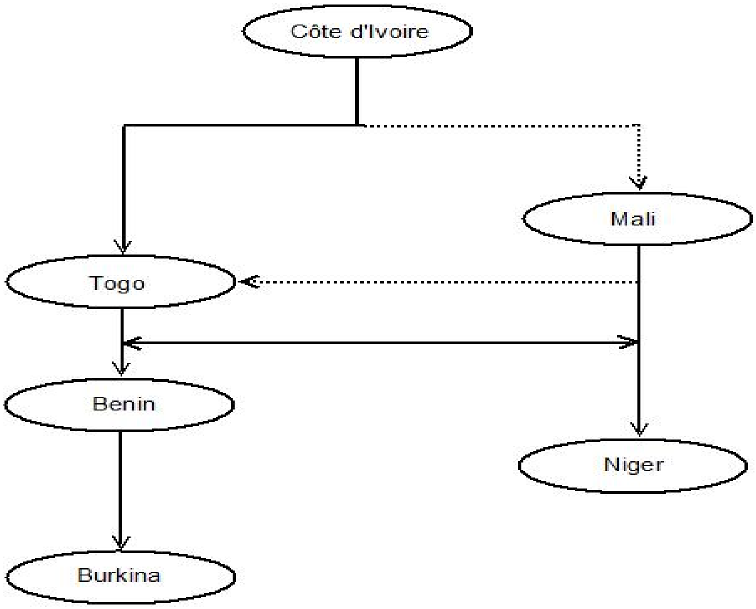
To perform poverty comparisons using household size as a discrete indicator, we construct a grid of points (z_y, L), 20 points for z_y and 10 for L. The points z_y for the wealth dimension are given by the 20 wealth quantiles obtained after merging the two distributions to be compared. We define $L = [1, 2, \ldots, 10]$, such that $L = 1$ corresponds to households of 10 individuals or more, $L = 2$ to those with 9 individuals, and so on until $L = 10$ for households comprising a single individual.

Figure 1 depicts the dominance relations between the six countries. Solid arrows indicate first-order dominance and dashed arrows show second-order dominance. Côte d'Ivoire dominates all other countries, followed by Mali, which dominates the next four, and then Togo dominates the remaining three. We observe a lack of dominance between Benin and Niger, as well as between Niger and Burkina. Second-order dominance of Côte d'Ivoire over Mali is only obtained if we exclude households of eight or more, nine or more, and ten or more individuals from the comparisons. In fact, a comparison limited to those households yields no dominance relationship because the curves intersect. However, when we expand the sample to include households with fewer than eight individuals, dominance appears. In the case of Mali and Togo, it is by excluding households of four or fewer individuals that we can find dominance.

More detailed results from the tests are reported in Tables 1, 2, and 3 in the appendices. Table 1 presents the results for first-order dominance. The p-values are the probabilities associated with H_0, *i.e.* that the second country does not sequentially dominate the first. The LR and t-statistic results are identical. Table 2 presents the second-order dominance surface in the case of Mali-Cote d'Ivoire, including the LR values. Values of LR equal to 0 mean that the curves intersect, precluding the possibility of dominance around that section of the surface. Testing for restricted dominance yields minima of 1.919 for the LR and 2.409 for the t-statistic. The p-value obtained by the bootstrap method equals 0.000 for both

[3]See Greenacre (1993) and Greenacre and Blasius (2006) for more details on MCA methods.

Figure 1: Diagram of sequential dominance between countries in the WAEMU, with household size as a discrete indicator



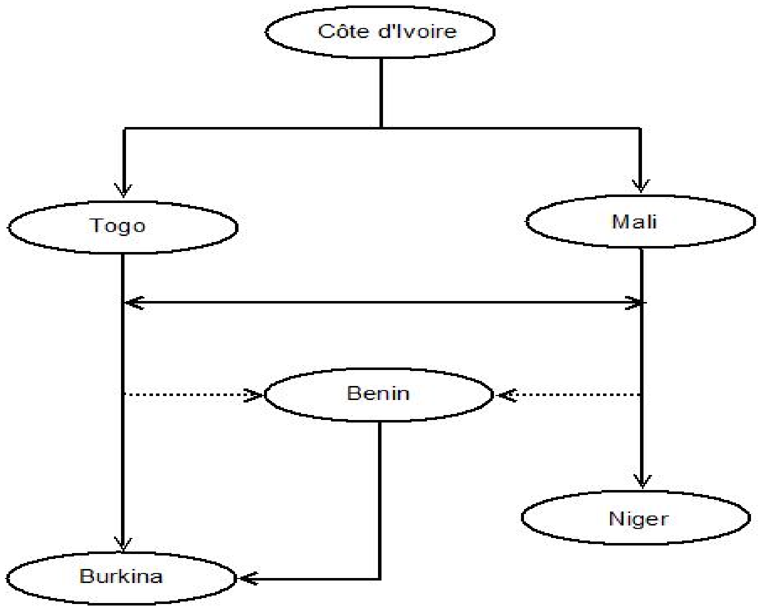
statistics, so the assumption of non-dominance is rejected with an error probability of less than 1%. Table 3 reports the results for Togo-Mali. Here, the LR and t-statistics are 2.193 and 2.195 respectively, with p-values also equal to 0.000.

The alternate specification treats the place of residence as a discrete indicator. Rural areas are assumed to be more needy that urban ones. Figure 2 shows the dominance diagram for this case. As previously, Côte d'Ivoire sequentially dominates all other countries in poverty, followed by Mali and Togo, between which there is no dominance. We find no dominance between Benin and Niger or between Niger and Burkina. Mali and Togo dominate Benin in the second order. Apart from the dominance relations involving Côte d'Ivoire, other dominance relations are restricted in that extreme values of z_y must generally be excluded to obtain dominance.

Detailed results for first-order dominance are presented in Table 4 in the Appendices. Table 5 reports the results for second-order dominance.

The third specification uses education level as the discrete variable. In doing so, this variable is excluded from the estimation of the asset index. The education level of the head of household consists in three categories, which are the lack of education, primary education, and secondary education or more. Figure 3 shows the diagram of dominance obtained in this third case. The results are close to those from the second specification, excepted that here Togo dominates Benin at first-

Figure 2: Diagram of sequential dominance between countries in WAEMU, with place of residence as a discrete indicator

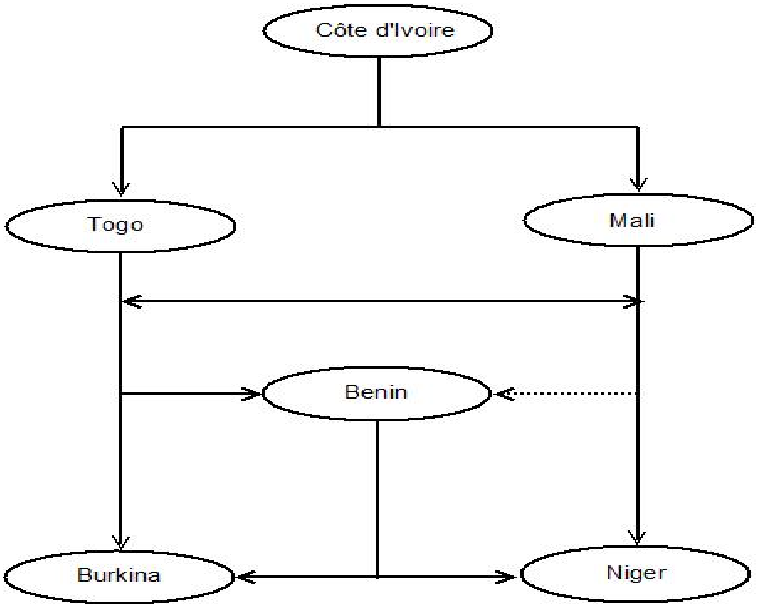


order while the latter dominates Niger at the same order. The lack of dominance relation is still observed between Togo and Mali on the one hand, and between Burkina and Niger, on the other hand. The detailed first-order dominance results are reported in Table 6.

For the first specification (with household size as a discrete indicator), we obtain 13 dominance relations out of a total of 15 — 11 instances of first-order and 2 of second-order dominance. For the second specification (with place of residence as a discrete indicator), there are 12 dominance relations, including 10 of the first-order. Finally, for the last specification (with education level as a discrete indicator), there are 13 dominance relations, 12 of them being first-order relations.

When there is dominance in the last column of the dominance surface (for example, the last column of Table 2), dominance in the univariate case is achieved since this abstracts from the role of the discrete indicator. Reducing the number of classes of households, *e.g.*, by considering the first class of households as those composed of 5 and more individuals, there is full dominance of Côte d'Ivoire on Mali. Idem for the dominance of Mali on Togo (Table 3). Moreover, dominance in welfare is more difficult to observe than dominance in poverty since the former involves the latter. For instance, in Table 3, a possible poverty dominance excluding welfare dominance would be a situation where null values tend to appear in the bottom and the right of the table. It would be possible, in this case, to obtain

Figure 3: Diagram of sequential dominance between countries in WAEMU, with education level as a discrete indicator



a sequence of thresholds $Z_1^+ \geq Z_2^+ \geq \cdots \geq Z_k^+$ for which poverty dominance holds statistically, but welfare dominance does not.

6 Conclusion

Aside from income, which is typically used as an indicator of well-being, several other variables would also seem to be able to capture important aspects of well-being and poverty, such as household size, place of residence, or literacy. We can often usefully think of these variables as capturing differences in household needs, and it would therefore seem preferable to take into account their distribution in conducting poverty and welfare comparisons. The discrete nature of these variables makes them amenable to the application of sequential dominance techniques to compare welfare and poverty, techniques derived among others by Atkinson (1992), Atkinson and Bourguignon (1987), Jenkins and Lambert (1993), Chambaz and Maurin (1998) and, more recently, by Duclos and Makdissi (2005) and Duclos, Sahn, and Younger (2007).

This paper implements such sequential dominance techniques using statistical inference methods based on the empirical likelihood ratio. This approach allows us to test a null hypothesis of non-dominance versus an alternative hypothesis of dominance. It therefore facilitates inference of multidimensional dominance relationships. Illustrations of comparisons across six West African countries are provided using DHS data and wealth indices. The first set of illustrations uses household size as a discrete indicator of needs, the second set considers place of residence as the discrete indicator of needs, and the third uses the educational level of the head of household.

Results from bootstrap tests lead to the inference of several dominance relationships in all of these specifications. Côte d'Ivoire generally ends up dominating all the other countries, followed by Mali and Togo. These results also suggest that tests based on the likelihood ratio can be useful for analyzing multidimensional poverty and welfare dominance when one of the dimensions of welfare is discrete.

7 Appendices

7.1 Proof of Proposition 5

This proof is based on that provided by Davidson and Duclos (2006). First, an asymptotical expression of the squared t-statistic in equation (17) is derived. Assume that $F(y, L) = G(y, L)$ and $\Delta(y, L) \equiv \widehat{F}(y, L) - \widehat{G}(y, L)$. Moreover, when $N \to \infty$, assume that $N_F/N \to r$, where r is a constant between 0 and 1. If $N \to \infty$, the following equalities hold:

$$\widehat{F}(y, L) = \widehat{G}(y, L) = F(y, L) + O_P\left(N^{-\frac{1}{2}}\right) \tag{24}$$

and

$$\Delta(y, L) = O_P\left(N^{-\frac{1}{2}}\right). \tag{25}$$

The statistic (17) could be asymptotically reexpressed as follows:

$$t^2(y, L) = \frac{r(1-r)}{F(y, L)(1 - F(y, L))} P\lim_{N\to\infty} N\Delta^2(y, L) + O_P\left(N^{-\frac{1}{2}}\right) \tag{26}$$

Davidson and Duclos (2006) show that the LR-statistic is asymptotically equivalent to the above statistic. Indeed, consider the statistic $LR(y, L)$ given in equation (15). Knowing that $N_F(y, L) = N_F F(y, L)$, $N_G(y, L) = N_G G(y, L)$, $M_F(y, L) = N_F - N_F(y, L)$ and $M_G(y, L) = N_G - N_G(y, L)$, (15) could be transformed as the sum of the two following expressions multiplied by 2:

$$\left\{ -N_F\widehat{F}(y, L)\log\left(\frac{N_F\widehat{F}(y,L)+N_G\widehat{G}(y,L)}{N\widehat{F}(y,L)}\right) - N_G\widehat{G}(y, L)\log\left(\frac{N_F\widehat{F}(y,L)+N_G\widehat{G}(y,L)}{N\widehat{G}(y,L)}\right) \right\} \tag{27}$$

and

$$\left\{ \begin{array}{l} -N_F(1-\widehat{F}(y,L))\log\left(\frac{N-(N_F\widehat{F}(y,L)+N_G\widehat{G}(y,L))}{N-(1-\widehat{F}(y,L))}\right) \\ -N_G(1-\widehat{G}(y,L))\log\left(\frac{N-(N_F\widehat{F}(y,L)+N_G\widehat{G}(y,L))}{N-(1-\widehat{G}(y,L))}\right) \end{array} \right\}. \tag{28}$$

(27) can be rewritten as

$$\begin{array}{c} -(N_F\widehat{F}(y,L)+N_G\widehat{G}(y,L))\log(N_F\widehat{F}(y,L)+N_G\widehat{G}(y,L)) \\ +N_F\widehat{F}(y,L)\log(N\widehat{F}(y,L))+N_G\widehat{G}(y,L)\log(N\widehat{G}(y,L)) \end{array} \tag{29}$$

Introducing $\Delta(y,L)$ in (29), and knowing that $\Delta(y,L) \equiv \widehat{F}(y,L) - \widehat{G}(y,L)$ and therefore that $N_F\widehat{F}(y,L)+N_G\widehat{G}(y,L) = N\widehat{G}(y,L)+N_F\Delta(y,L)$, a reduced expression is obtained by a Taylor expansion:

$$\frac{N_F N_G \Delta^2(y,L)}{2N\widehat{F}(y,L)} + O_P\left(N^{-\frac{1}{2}}\right). \tag{30}$$

Given (24), the above equals

$$\frac{1}{2}\frac{N_F N_G \Delta^2(y,L)}{NF(y,L)}. \tag{31}$$

By reducing the second expression 28 similarly and multiplying it by 2, the equation (15) becomes:

$$\frac{N_F N_G \Delta^2(y,L)}{NF(y,L)(1-F(y,L))} + O_P\left(N^{-\frac{1}{2}}\right). \tag{32}$$

Knowing that $N_F/N \to r$ when $N \to \infty$, this ultimately becomes:

$$\frac{r(1-r)}{F(y,L)(1-F(y,L))} + O_P\left(N^{-\frac{1}{2}}\right) P\lim N\Delta^2(y,L), \tag{33}$$

which is equivalent to the right-hand-side expression in equation (26).

7.2 Main tables

Table 1: First-order dominance tests, with household size as a discrete indicator

F not dominated by G	LR		t-statistic	
	LR	$p - value$	t	$p - value$
Benin-Côte d'Ivoire	2.295	0.000***	2.317	0.000***
Burkina-Côte d'Ivoire	2.360	0.000***	2.377	0.000***
Niger-Côte d'Ivoire	2.337	0.000***	2.379	0.000***
Togo-Côte d'Ivoire	3.180	0.000***	3.259	0.000***
Benin-Mali	2.478	0.000***	2.491	0.000***
Burkina-Mali	7.197	0.000***	6.227	0.000***
Niger-Mali	2.656	0.025**	2.668	0.025**
Benin-Togo	2.635	0.005***	2.512	0.005***
Burkina-Togo	4.173	0.003***	4.071	0.003***
Niger-Togo	2.922	0.090*	2.921	0.090*
Burkina-Benin	2.584	0.035**	2.631	0.035**

Table 2: Second-order dominance surface, Mali *vs* Côte d'Ivoire

z_y	**Household size**									
	10 & +	*9 & +*	*8 & +*	*7 & +*	*6 & +*	*5 & +*	*4 & +*	*3 & +*	*2 & +*	*1 & +*
-0.46	3.88	1.13	1.21	1.92	2.51	2.54	2.51	2.99	2.99	2.99
-0.44	2.01	1.10	1.44	2.33	3.06	3.91	4.38	5.43	5.43	5.43
-0.35	1.69	0.86	1.78	2.84	4.37	6.19	6.97	8.71	8.73	8.73
-0.30	1.94	1.33	2.39	3.64	5.38	7.26	8.03	9.86	9.89	9.89
-0.17	1.05	0.67	1.56	3.12	5.02	6.95	7.68	9.62	9.60	9.60
-0.13	0.24	0.00	0.83	2.57	4.53	6.51	7.30	9.29	9.27	9.27
-0.12	0.02	0.00	0.67	2.49	4.48	6.50	7.33	9.36	9.33	9.33
-0.04	0.00	0.00	0.31	2.53	4.69	6.91	8.01	10.20	10.14	10.14
-0.03	0.00	0.00	0.30	2.58	4.77	7.02	8.16	10.38	10.32	10.32
-0.02	0.00	0.00	0.48	2.81	5.03	7.32	8.52	10.77	10.71	10.71
-0.00	0.00	0.00	0.69	3.08	5.33	7.66	8.92	11.20	11.14	11.14
0.05	0.00	0.27	1.67	4.27	6.69	9.30	10.80	13.30	13.23	13.23
0.11	0.25	0.93	2.47	5.21	7.77	10.60	12.28	14.98	14.91	14.91
0.18	0.75	1.60	3.26	6.18	8.92	12.06	14.03	16.99	16.92	16.92
0.33	1.03	2.14	4.04	7.17	10.21	13.82	16.33	19.78	19.72	19.72
0.71	1.02	2.41	4.72	8.12	11.60	15.99	19.66	24.22	24.14	24.14
1.02	0.58	2.04	4.51	7.99	11.58	16.30	20.62	25.88	25.80	25.80
1.24	0.00	1.19	3.70	7.26	10.88	15.82	20.60	26.43	26.35	26.35
1.52	0.00	0.00	2.42	6.04	9.65	14.81	20.12	26.70	26.62	26.62
1.82	0.00	0.00	0.92	4.54	8.10	13.46	19.29	26.78	26.70	26.70

Table 3: Second-order dominance surface, Togo *vs* Mali

	Household size									
$\mathbf{z}_y$	*10 & +*	*9 & +*	*8 & +*	*7 & +*	*6 & +*	*5 & +*	*4 & +*	*3 & +*	*2 & +*	*1 & +*
-0.47	2.40	3.04	3.29	3.56	3.45	2.75	2.95	2.64	2.64	2.64
-0.45	4.68	5.68	6.37	5.16	5.76	4.61	4.01	2.91	2.91	2.91
-0.39	4.61	5.48	6.11	5.71	4.98	2.81	1.43	0.00	0.00	0.00
-0.34	4.62	5.50	5.99	5.67	4.70	2.32	0.77	0.00	0.00	0.00
-0.31	5.03	6.00	6.29	5.96	4.83	2.34	0.66	0.00	0.00	0.00
-0.24	7.15	8.46	8.79	8.67	7.72	5.50	3.75	1.11	1.16	1.16
-0.19	8.66	10.12	10.50	10.49	9.64	7.65	5.93	3.30	3.36	3.36
-0.14	10.29	11.83	12.28	12.35	11.62	9.82	8.11	5.53	5.60	5.60
-0.12	10.70	12.25	12.72	12.76	12.05	10.30	8.56	5.98	6.06	6.06
-0.05	11.40	12.89	13.35	13.23	12.48	10.75	8.80	6.27	6.37	6.37
-0.03	11.52	12.97	13.42	13.24	12.49	10.77	8.76	6.23	6.34	6.34
-0.01	11.40	12.78	13.16	12.93	12.14	10.39	8.31	5.76	5.88	5.88
0.01	11.04	12.30	12.60	12.27	11.43	9.60	7.40	4.76	4.89	4.89
0.09	9.95	10.96	11.17	10.63	9.75	7.72	5.29	2.41	2.56	2.56
0.14	9.18	10.04	10.22	9.58	8.67	6.54	3.96	0.91	1.09	1.09
0.23	8.14	8.78	8.94	8.23	7.31	5.13	2.32	0.00	0.00	0.00
0.38	7.10	7.49	7.69	6.88	5.96	3.79	0.64	0.00	0.00	0.00
0.76	5.84	5.90	6.14	5.29	4.47	2.43	0.00	0.00	0.00	0.00
1.14	5.44	5.35	5.56	4.79	4.08	2.19	0.00	0.00	0.00	0.00
1.60	5.27	5.07	5.30	4.67	4.11	2.55	0.00	0.00	0.00	0.00

Table 4: First-order dominance tests, with place of residence as a discrete indicator

F not dominated by G	LR-statistic		t-statistic		Type of
(F-G)	LR	p-value	t	p-value	dominance
Benin-Côte d'Ivoire	1.545	0.000***	1.539	0.000***	no restriction
Burkina-Côte d'Ivoire	5.466	0.000***	5.021	0.000***	no restriction
Niger-Côte d'Ivoire	7.150	0.000***	5.573	0.000***	no restriction
Togo-Côte d'Ivoire	4.358	0.000***	5.303	0.000***	no restriction
Mali-Côte d'Ivoire	4.463	0.000***	4.665	0.000***	no restriction
Burkina-Mali	6.244	0.098*	6.359	0.100*	restricted
Niger-Mali	7.391	0.000***	7.232	0.000***	restricted
Burkina-Togo	6.396	0.000***	6.230	0.000***	restricted
Niger-Togo	7.360	0.040**	7.301	0.040**	restricted
Burkina-Benin	4.312	0.013**	4.351	0.013**	restricted

Table 5: Second-order dominance tests, with place of residence as a discrete indicator

F not dominated by G	LR-statistic		t-statistic	
(F-G)	LR	p-value	t	p-value
Benin-Mali	1.998	0.000***	2.000	0.000***
Benin-Togo	2.811	0.000***	2.813	0.000***

Table 6: First-order dominance tests, with education level as a discrete indicator

F not dominated by G	LR-statistic		t-statistic		Type of
(F-G)	LR	p-value	t	p-value	dominance
Benin-Côte d'Ivoire	6.073	0.000***	5.649	0.000***	no restriction
Burkina-Côte d'Ivoire	4.987	0.000***	4.0784	0.000***	no restriction
Mali-Côte d'Ivoire	2.279	0.000***	2.314	0.000***	no restriction
Niger-Côte d'Ivoire	8.250	0.000***	6.339	0.000***	no restriction
Togo-Côte d'Ivoire	2.752	0.000***	2.742	0.000***	restricted
Benin-Togo	2.920	0.003***	2.942	0.003***	restricted
Burkina-Togo	5.516	0.000***	5.409	0.000***	restricted
Niger-Togo	6.423	0.063*	6.373	0.063*	restricted
Burkina-Mali	9.200	0.000***	8.366	0.000***	restricted
Niger-Mali	7.177	0.000***	7.026	0.000***	restricted
Burkina-Benin	3.786	0.005***	3.846	0.005***	restricted
Niger-Benin	3.951	0.058*	4.001	0.058*	restricted

References

ATKINSON, A. (1992): "Measuring Poverty and Differences in Family Composition," *Economica*, 59, 1–16.

ATKINSON, A. B. AND F. BOURGUIGNON (1987): *Income Distribution and Differences in Needs*, New York: Macmillan.

BOURGUIGNON, F. (1989): "Family Size and Social Utility: Income Distribution Dominance Criteria," *Journal of Econometrics*, 42, 67–80.

BOURGUIGNON, F. AND S. CHAKRAVARTY (2002): "Multi-dimensional poverty orderings," Tech. Rep. 2002-22, DELTA.

CHAMBAZ, C. AND E. MAURIN (1998): "Atkinson and Bourguignon's Dominance Criteria: Extended and Applied to the Measurement of Poverty in France," *Review of Income and Wealth*, 44, 497–513.

DAVIDSON, R. (2007): "Testing for Restricted Stochastic Dominance: Some Further Results," Tech. rep., McGill University.

DAVIDSON, R. AND J.-Y. DUCLOS (2006): "Testing for Restricted Stochastic Dominance," Working Paper 06-09, CIRPEE.

DUCLOS, J.-Y. AND P. MAKDISSI (2005): "Sequential Stochastic Dominance and the Robustness of Poverty Orderings," *Review of Income and Wealth*, 51, 63–88.

DUCLOS, J.-Y., D. E. SAHN, AND S. D. YOUNGER (2007): "Robust Multidimensional Poverty Comparison with Discrete Indicators of Well-being," in *Inequality and Poverty Re-examined*, ed. by S. P. Jenkins and J. Micklewright, Oxford: Oxford University Press, 185–206.

GREENACRE, M. (1993): *Correspondence Analysis in Practice*, Academic Press.

GREENACRE, M. AND J. BLASIUS (2006): *Multiple Correspondence Analysis and Related Methods*, Statistics in the Social and Behavioral Sciences Series, CRC Press.

JENKINS, S. AND P. LAMBERT (1993): "Ranking Income Distributions When Needs Differ," *Review of Income and Wealth*, 39, 337–56.

KAUR, A., B. L. S. PRAKASA RAO, AND H. SINGH (1994): "Testing for Second-Order Stochastic Dominance of Two Distributions," *Econometric Theory*, 10, 849–66.

SEN, A. (1979): "Personal Utilities and Public Judgements: Or what's Wrong with Welfare Economics," *The Economic Journal*, 89, 537–558.

——— (1985): *Commodities and Capabilities*, Amsterdam: North-Holland.

——— (1987): *The Standard of Living*, Cambridge: Cambridge University Press.

TSUI, K. Y. (2002): "Multidimensional Poverty Indices," *Social Choice and Welfare*, 19, 69–93.