Review of Income and Wealth
Series 52, Number 3, September 2006

A BOUNDED INDEX TEST TO MAKE ROBUST HETEROGENEOUS WELFARE COMPARISONS

BY ANDRÉ DECOSTER*

AND

ERWIN OOGHE

CES, Katholieke Universiteit Leuven

Fleurbaey *et al.* (2003) develop a bounded dominance test to make robust welfare comparisons, which is intermediate between Ebert's (1999) cardinal dominance criterion—generalized Lorenz dominance applied to household incomes, divided and weighted by an equivalence scale—and Bourguignon's (1989) ordinal dominance criterion. In this paper, we develop a more complete, but less robust bounded index test, which is intermediate between Ebert's (1997) cardinal index test—an index applied to household incomes, divided and weighted by the equivalence scale—and a (new) sequential index test—an index applied to household incomes of the most needy only, the most and second most needy only, and so on. We illustrate the power of our test to detect welfare changes in Russia using data of the RLMS-surveys.

1. INTRODUCTION

When income units are homogeneous in non-income characteristics, there exist many tools to evaluate income distributions and the properties of these tools are well-known; see Lambert (2001) for an overview. Basically, these tools can be classified in two groups. Indices map income distributions into a comparable number measuring the welfare of the distribution under consideration, whereas dominance criteria look for unanimity among a "wide" class of such indices. The most well-known dominance criterion is the generalized Lorenz dominance (GLD) criterion due to Shorrocks (1983). Unfortunately, these tools are not well-suited to make reasonable comparisons in practice, because "At the heart of any distributional analysis, there is the problem of allowing for differences in people's non-income characteristics" (Cowell and Mercader-Prats, 1999).

To make robust heterogeneous welfare comparisons, the most well-known result is Atkinson and Bourguignon's (1987) sequential generalized Lorenz dominance (SGLD) criterion: (i) divide all income units into different need types on the basis of non-income characteristics; and (ii) check—on the basis of the GLD criterion—whether the most needy in one distribution dominate the most needy in another distribution, whether the most and second most needy together in the former distribution also dominate the most and second most needy in the other

Note: We are grateful to Bart Capéau, Peter Lambert, Luc Lauwers and two anonymous referees for useful comments and help. Erwin Ooghe is a Postdoctoral Fellow of the Fund for Scientific Research—Flanders.

*Correspondence to: André Decoster, Professor of Economics, Center for Economic Studies, Katholieke Universiteit Leuven, Naamsestraat, 69, B-3000 Leuven, Belgium (andre.decoster@econ.kuleuven.be).

distribution, and so on. The SGLD criterion is very robust—as it is equivalent to unanimity among a wide set of utilitarian welfare orderings—but it has little power to rank distributions. It has been extended by Atkinson (1992), Jenkins and Lambert (1993), Chambaz and Maurin (1998), Lambert and Ramos (2002), and Moyes (1999) to deal with changing demographics, poverty and/or the principle of diminishing transfers. We also refer to Bourguignon (1989) for a related dominance criterion.

The SGLD criterion is often called an "ordinal" dominance criterion because the needs classes have to be defined in an ordinal way only, i.e. a ranking of all non-income types on the basis of needs. In contrast, practitioners often use equivalence scales to cardinalize needs differences between income units, expressing, for example, that (for each income level) a couple needs m times the income of a single to reach the same living standards, with m between 1 and 2. Equivalence scales are defined with respect to a reference type, usually a single. Once defined, practitioners can: (i) transform the heterogeneous distribution of incomes and types into a homogeneous distribution of equivalent incomes (for reference types); and (ii) use a standard tool (an index or dominance criterion) applied to the vector of equivalent incomes. Depending on the chosen tool, we call it either a cardinal index or a cardinal dominance approach.[1]

Fleurbaey *et al.* (2003) consider a dominance criterion which is intermediate between the ordinal and the cardinal approach. They propose to make welfare comparisons using the GLD criterion for a bounded set of equivalence scale vectors. Choosing the bounded set as small as possible, their criterion reduces to Ebert's (1999) cardinal GLD approach—the GLD criterion applied to household incomes, both divided and weighted by the (unique) equivalence scale—and choosing the bounded set as wide as possible, their criterion is equivalent with one of Bourguignon's (1989) dominance criteria.

The different existing ways to deal with heterogeneity, as well as the main contributions, are summarized in Table 1. The rows denote the different ways to measure the well-being of heterogeneous income units: do we use one specific equivalence scale (cardinal), a bounded set of equivalence scales (intermediate) or no scales at all, which is equivalent to a "wide" set of scales (ordinal)? The columns summarize the different ways to aggregate the resulting well-beings: do we use an index or a dominance criterion, for example, the GLD criterion? Moving downwards (resp. rightwards) in Table 1 increases robustness as we consider more equivalence scales (resp. indices), at the cost of completeness, i.e. the power to rank distributions.

In this paper, we explore areas A and B in Table 1. In the next section, we introduce Fleurbaey *et al.*'s (2003) bounded dominance test and propose an alternative *bounded index* test, based on a specific iso-elastic measure (area A in Table 1). Using the same bounds, the bounded index test is less robust, but more powerful compared to Fleurbaey *et al.*'s (2003) bounded dominance test. Choosing bounds as small as possible in the bounded index test, we get a cardinal index test in line with

[1]As noted by Pyatt (1990) and Glewwe (1991), the use of an equivalence scale may give rise to a weighting problem. More precisely, it is not clear whether one should weight each income unit by the number of individuals or by the equivalence scale; see Ebert (1997), Ebert and Moyes (2003), Shorrocks (2004) and Capéau and Ooghe (2004).

TABLE 1
A Classification of the Different Ways to Deal with Heterogeneity

	Index	Dominance
Cardinal	Ebert (1997, 1999) and Shorrocks (2004)	
Intermediate	(A)	Fleurbaey, Hagneré and Trannoy (2003)
Ordinal	(B)	Atkinson and Bourguignon (1987) Bourguignon (1989)

Ebert's (1997) weighting scheme: an index applied to household incomes, both divided and weighted by the (unique) equivalence scale. Choosing bounds as wide as possible, we obtain a (new) *sequential index* test (area B in Table 1), i.e. checking—on the basis of the iso-elastic index—whether welfare is higher for the most needy income units only, for the most and second most needy only, and so on.

We illustrate the bounded dominance and the bounded index test by measuring welfare changes in Russia from 1994 to 2002 on the basis of the RLMS (Russian Longitudinal Monitoring Survey) data. The post-communist era (after 1991) was characterized by rising inequality and strongly decreasing GDP per capita, reaching rock bottom with the financial crisis of August 1998. Afterwards, enhanced political stability and increasing oil prices led to strong growth and slowly decreasing inequality. Therefore, we expect welfare to decrease in the first and to rise again in the second period. While the bounded index test is able to detect such a pattern, this is not the case for the bounded dominance test. Robustness with respect to the aggregation of well-being of individuals, rather than with respect to its measurement, turns out to be the main culprit.

2. Robust Welfare Comparisons

2.1. *Notation*

Consider household incomes $y \in \mathbb{R}_+$ and types $k \in \mathbb{K} = \{1, \ldots, K\}$ representing relevant non-income characteristics; types are ordered from least ($k = 1$) to most needy ($k = K$). A heterogeneous distribution is denoted by $F = (p_1, \ldots, p_K, F_1, \ldots, F_K)$, with p_k the proportion of households with type k and F_k the (differentiable) income distribution function of type k households defined over $\mathbb{R}_+$ with a finite support $[0, \bar{s}_k]$. We focus directly on the case where demographics might change, or the proportions p_k may vary over the different distributions. Household utility functions $U_k : \mathbb{R}_+ \to \mathbb{R}$ measure the utility of a household with type k as a function of its income, with $U_k(0)$ finite for all $k \in \mathbb{K}$. Social welfare in a distribution F is measured by the average household utility in society:

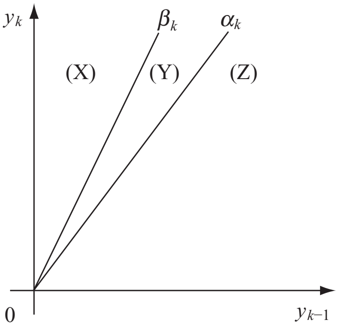

Figure 1. Partial Comparability in Case of Bounded Equivalence Scales

$$W : F \mapsto W(F) = \sum_{k \in \mathbb{K}} p_k \int_0^{\bar{s}_k} U_k dF_k. \tag{1}$$

2.2. *A Bounded Dominance Test*

Fleurbaey, Hagneré and Trannoy (2003) (FHT) consider a lower and upper bound vector $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \mathbb{R}^K$ which satisfy

$$(1, 1, \ldots, 1) \leq (\alpha_1 = 1, \alpha_2, \ldots, \alpha_K) \leq (\beta_1 = 1, \beta_2, \ldots, \beta_K). \tag{2}$$

Type 1 (the least needy type) will be referred to as the reference type. They impose the following conditions on household utility functions, all assumed to be twice continuously differentiable:

A1: $U_k' \geq 0$, for all $k \in \mathbb{K}$,
A2: $U_k'' \leq 0$, for all $k \in \mathbb{K}$,
A3: $U_k'(\alpha_k y) \geq U_{k-1}'(y)$, for all $y \in \mathbb{R}_+$ and for all $k = 2, \ldots, K$,
A4: $U_k'(\beta_k y) \leq U_{k-1}'(y)$, for all $y \in \mathbb{R}_+$ and for all $k = 2, \ldots, K$,
A5: a vector $(a_2, \ldots, a_K)$ exists s.t. $\begin{cases} (a) & U_k(a_k) = U_1(a_1) \quad \text{for all } k = 2, \ldots, K \\ (b) & U_k'(a_k) = U_1'(a_1) \quad \text{for all } k = 2, \ldots, K \end{cases}$

The marginal utility of a type is its social priority, because it tells a utilitarian social planner where to put his money first when maximizing social welfare. Assumptions A1 and A2 are standard: all types have positive, but decreasing, social priority. In terms of money transfers, these conditions require that more income is better (monotonicity) and transfers from rich to poor households of the same type improve social welfare (the (within type) Pigou–Dalton transfer principle).

Assumptions A3 and A4 link the social priority of the different types. As a consequence, they tell us something about the welfare effect of money transfers between types, because a small money transfer from a type with a lower to a type with a higher social priority, must improve social welfare.

Figure 1 illustrates the social priority classification of two households with adjacent types $k - 1$ and k, depending on their household incomes y_{k-1} and y_k.

For all income combinations in zone (X), type k has a lower social priority than type $k - 1$, and vice-versa in zone (Z). In the area (Y), there is disagreement whether type k or $k-1$ has the highest social priority. Notice that the disagreement zone dissapears when choosing $\alpha_k = \beta_k$, while it increases when lowering α_k and/or increasing β_k. The proposals of Ebert (1999) and Bourguignon (1989) correspond with the limiting cases $\alpha_k = \beta_k$ (for all $k = 2, \ldots, K$) and $\alpha_k = 1$, $\beta_k \to \infty$ (for all $k = 2, \ldots, K$), respectively.

Finally, Assumption A5 depends on an exogeneous income level a_1 and is imposed to deal with changing demographics. At a certain income level, social welfare is invariant to transfers of population across need groups (A5a) and transfers of income across need groups (A5b).

We denote with $\mathcal{U}(\alpha, \beta, a_1)$ the family of utility profiles $(U_1, \ldots, U_K)$ satisfying assumptions A1–A5, given α, β, a_1. We say that a distribution F welfare dominates G according to the family $\mathcal{U}(\alpha, \beta, a_1)$, denoted $F \succsim_{(\alpha,\beta,a_1)} G$, if and only if the welfare difference $\Delta W = W(F) - W(G)$ is positive for all profiles in $\mathcal{U}(\alpha, \beta, a_1)$. The following proposition shows how welfare dominance for $\succsim_{(\alpha,\beta,a_1)}$ can be implemented. Define functions H_k^1 and H_k^2 over $\mathbb{R}_+$ (for all types $k \in \mathbb{K}$) as:

$$(3) \qquad H_k^1(y) = p_k F_k(y) - q_k G_k(y), \quad \text{and} \quad H_k^2(y) = \int_0^y H_k^1(x)\,dx.$$

We get the following bounded dominance criterion:

Fleurbaey, Hagneré and Trannoy (2003). Consider two heterogeneous distributions F and G, an exogeneous income level $a_1 \geq \max\left(\frac{\bar{s}_1}{\alpha_1}, \frac{\bar{s}_2}{\alpha_1\alpha_2}, \ldots, \frac{\bar{s}_K}{\alpha_1\alpha_2\ldots\alpha_K}\right)$ and lower and upper bound vectors α, $\beta \in \mathbb{R}^K$ which satisfy (2). Let $Z_{K+1} : x \mapsto 0$. Define functions Z_k recursively (starting from $k = K$ downwards to $k = 2$) as $Z_k : y \mapsto \max_{\alpha_k y \leq x \leq \beta_k y} \{H_k^2(x) + Z_{k+1}(x)\}$. We have:

$$(4) \qquad F \succsim_{(\alpha,\beta,a_1)} G \Leftrightarrow H_1^2(y) + Z_2(y) \leq 0 \quad \text{for all } y \in [0, a_1].$$

In the next subsection, we present a more powerful, but less robust bounded index test.

2.3. *A Bounded Index Test*

The criterion in equation (4) incorporates robustness in two dimensions: in the aggregation of well-beings of households and in the cardinalization of the needs differences. Eventual lack of power to order distributions by means of (4) cannot easily be attributed to one of these dimensions. Put otherwise, it is possible that the interesting idea of bounded equivalence scales in FHT (2003) is drawn into a too demanding aggregation criterion, reducing its possible relevance for practitioners. In the following we therefore propose a methodology which keeps the robustness of the bounded equivalence scales intact, while sacrifying some of the robustness of the GLD criterion.

We define an iso-elastic household utility function I, which is reminiscent of Clark *et al.*'s (1981) poverty index:

$$(5) \qquad I:\mathbb{R}_+\times\mathbb{R}_{++}\rightarrow\mathbb{R}:(y,m)\rightarrow\begin{cases}\dfrac{m}{1-\rho}\left(\left(\dfrac{y}{m}\right)^{1-\rho}-\left(\dfrac{a_1}{m}\right)^{1-\rho}\right), & \text{for } y\leq a_1\\ 0, \quad \text{for } y>a_1\end{cases},$$

with $a_1 \in \mathbb{R}_+$ an exogeneous income level, ρ the inequality aversion parameter, with $\rho \geq 0$, $\rho \neq 1$, and m an equivalence scale.[2] We briefly explain the different modalities.

Notice first the two terms within brackets, the equivalent income (y/m) and an additional equivalized term (a_1/m). The term (a_1/m) is introduced to ensure that the iso-elastic household utility profiles (see below) become a subset of FHT's (2003) profiles. To put it differently, this term ensures that condition A5 will be satisfied. We stress however that all subsequent results remain valid if one leaves it out to obtain a more standard Kolm–Atkinson–Sen welfare index. Next, both terms are transformed via an iso-elastic function. This functional form is indeed specific, but still allows for sufficient flexibility by varying the parameter ρ, which measures the cost of inequality: the higher this parameter, the more of the average one is willing to give up for an equal society. Finally, notice that the equivalence scale is also used to weight utilities: this is to ensure that households with higher equivalent incomes have a lower marginal utility, or equivalently, a lower social priority; see Ebert (1997, 1999), Ebert and Moyes (2003), Shorrocks (2004) and Capéau and Ooghe (2004) for a discussion of the weighting issue. Notice again that all subsequent results can be adapted for more conventional weighting procedures, such as weighting by household size.

The equivalence scale m will be used to differentiate the household utility functions according to needs. More precisely, to satisfy conditions A3 and A4, we consider equivalence scale vectors $\mathbf{m} = (m_1, \ldots, m_K)$ (consisting of one equivalence scale for each household type) which belong to the following bounded set:

$$\mathcal{M}(\alpha,\beta)=\{\mathbf{m}\in\mathbb{R}^K \mid m_1=1 \quad \text{and} \quad \alpha_k m_{k-1}\leq m_k\leq\beta_k m_{k-1} \quad \text{for all } k=2,\ldots,K\}.$$

Choosing $\alpha_k = 1$ and $\beta_k \rightarrow \infty$, for all $k = 2, \ldots, K$, $\mathcal{M}(\alpha, \beta)$ contains all equivalence scales satisfying $m_1 = 1 \leq m_2 \leq \ldots \leq m_K$; choosing $\alpha_k = \beta_k$, for all $k = 2, \ldots, K$, is choosing one specific equivalence scale vector $\mathbf{m}$, with $m_i = \Pi_{k=1}^{i}\alpha_k$, for all $i = 1, \ldots, K$.

We denote with $\mathcal{I}(\alpha, \beta, a_1, \rho)$ the family of iso-elastic utility profiles ($I(\cdot, m_1), \ldots, I(\cdot, m_K)$), one for each vector $\mathbf{m}$ in $\mathcal{M}(\alpha, \beta)$, and $\succsim_{(\alpha,\beta,a_1,\rho)}$ is the corresponding unanimity quasi-ordering.

[2]In case $\rho = 1$, the usual logarithmic case applies, i.e.

$$I:\mathbb{R}_+\times\mathbb{R}_{++}\rightarrow\mathbb{R}:(y,m)\rightarrow\begin{cases}m\left(\ln\left(\dfrac{y}{m}\right)-\ln\left(\dfrac{a_1}{m}\right)\right), & \text{for } y\leq a_1\\ 0, \quad \text{for } y>a_1\end{cases}.$$

We obtain a bounded index criterion (proofs can be found in the appendix):

Proposition 1. Consider two heterogeneous distributions F and G, an exogeneous income level $a_1 \geq \max(\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_K)$, lower and upper bound vectors $\alpha, \beta \in \mathbb{R}^K$ which satisfy (2) and an inequality aversion parameter $\rho \geq 0$. Let $Z^{\circ}_{K+1} : x \mapsto 0$ and $b_k = \int_0^{\bar{s}_k} \frac{1}{1-\rho}\left((y)^{1-\rho} - (a_1)^{1-\rho}\right) dH^1_k(y)$, for all $k \in \mathbb{K}$. Define functions Z°_k recursively (starting from $k = K$ downwards to $k = 3$) as $Z^{\circ}_k : m \mapsto \min_{\alpha_k m \leq x \leq \beta_k m} \{b_k x^{\rho} + Z^{\circ}_{k+1}(x)\}$. We have:

(6) $$F \succsim_{(\alpha,\beta,a_1,\rho)} G \text{ if and only if } b_1 + b_2 m^{\rho} + Z^{\circ}_3(m) \geq 0 \text{ for all } m \in [\alpha_2, \beta_2].$$

Notice that the functions Z°_k for $k = 3, \ldots, K$ can be easily calculated, because monotonicity guarantees that the minimum can be found at one of the extremes. Furthermore, the bounded dominance and bounded index criteria are nested, i.e. $F \succsim_{(\alpha,\beta,a_1)} G$ implies $F \succsim_{(\alpha,\beta,a_1,\rho)} G$, for all $\rho \in \mathbb{R}_+$. Finally, choosing $\alpha = \beta$, we obtain Ebert's cardinal approach for indices, i.e. apply an index to household incomes, divided and weighted by the equivalence scale. Choosing $\alpha_k = 1$ and $\beta_k \to \infty$, for all $k \in \mathbb{K}$, our next proposition tells us that $\succsim_{(\alpha,\beta,a_1,\rho)}$ reduces to a (new) sequential index test in the spirit of Atkinson and Bourguignon (1987):

Proposition 2. Consider two heterogeneous distributions F and G, an exogeneous income level $a_1 \geq \max(\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_K)$, lower and upper bound vectors $\alpha = (1, \ldots, 1)$ and $\beta \to (1, \infty, \ldots, \infty)$ and an inequality aversion parameter $\rho \geq 0$. Define all b_k's as in Proposition 1. We have:

(7) $$F \succsim_{(\alpha,\beta,a_1,\rho)} G \text{ if and only if } \sum_{k=i}^{K} b_k \geq 0 \text{ for all } i = 1, \ldots, K.$$

3. Welfare Changes in Russia 1994–2002

We illustrate and compare the bounded dominance and the bounded index test by measuring welfare changes in Russia from 1994 to 2002 on the basis of the RLMS (Russian Longitudinal Monitoring Survey) data. First, we briefly describe the data and the Russian socio-economic background.

3.1. *Data and Socio-economic Background*

The RLMS surveys start in 1992 and describe in detail the living conditions, expenditures, incomes, and socio-economic characteristics of a representative panel of Russian households.[3] The first phase (the first four rounds) being merely a pilot survey, we use data of the second phase only, starting from Round 5 in 1994 up to Round 11 in 2002. In each round, we use the appropriate sample weights, delivered by the RLMS team, to gross up the sample to a nationally representative

[3]See the website http://www.cpc.unc.edu/projects/rlms and Mroz *et al.* (2004) for detailed information on this survey. The data can be freely downloaded.

population of Russian households. The RLMS datasets contain expenditures both in current and in constant prices, where the RLMS researchers have converted the nominal ones into constant prices of 1992 by means of region specific (but not commodity specific) price indices. To measure living standards of Russian households we use non-durable expenditures in constant prices.

In the Appendix (Table A1), we sketch the evolution of average per capita real expenditures, of inequality measured by the Gini, and of the proportion and the average real expenditures of different needs groups in the Russian population during the period covered by the different rounds of the RLMS surveys. The picture emerging from the use of microdata roughly confirms the by now well-documented U-shaped pattern found in aggregate data, such as the evolution of GDP per capita: a steady and sharp decline of (average) well-being up to the financial crisis of 1998 and a recovery from 2000 onwards.[4] For the evolution of inequality we find the inverse pattern: the Gini increases from 1994 up to 1996 (rounds 5, 6 and 7), and it decreases from 1998 to 2002.[5] Contrary to the existing empirical literature, we focus on welfare rather than inequality rankings. On the one hand, we are prepared to accept at least some partiality of the ranking of the different years, due to the required robustness. On the other hand, the RLMS data suggest a clear (but non-robust) picture of welfare changes in Russia. The steeply decreasing average income and the slightly increasing inequality in the first half of the period, and the reverse in the second half of the period, leads one to expect social welfare to go down in the first and catch up again in the second period. At least, we expect a reasonably robust welfare measure to detect parts of this U-pattern.

It is striking that the extensive literature on the inequality evolution in Russia during the transition did not pay any attention to the issue of equivalence scales. Most authors seem to take for granted that the most sensible choice is to work with per capita concepts.[6] Yet, preliminary results on the RLMS data do show a sensitivity to the scale. If we calculate the Gini coefficient for a continuum of equivalence scales, defined by the number of persons to the power θ, where θ varies from 0 to 1, and we then rank the years from lowest to highest Gini of equivalent income, the ranking is *not* robust. The year 1995, for example, has the lowest Gini when calculated on household expenditures ($\theta = 0$), but only the fourth lowest Gini

[4]See, for example, the World Development Indicators published by the World Bank (2004). The U-shape, with a recovery from 2000 onwards, is nevertheless more pronounced in the microdata of the RLMS-survey. This is in line with recent findings in the debate on the evolution of world income inequality, where one observes large discrepancies between the growth of consumption in the surveys and the growth of either GDP or the consumption aggregate of GDP for many countries (see Deaton, 2001). No satisfactory explanation has been given up to now for these large differences.

[5]Also these findings fit quite well with the extensive literature on the evolution of Russian inequality. During the first years of the transition (from 1990 to 1995) there was an unprecendented rise in inequality, well documented, for example, in Kislitsyna (2003), Yemtsov (2003), Commander *et al.* (1999) and Lokshin and Popkin (1999). For the second half of the 1990s the picture differs, depending on whether or not one uses Goskomstat data. Kislitsyna (2003) finds moderately increasing inequality from 1996 to 2001 with Goskomstat data, but clearly declining inequality in the RLMS data. Our declining Gini from 1998 onwards corresponds well with her results. Galbraith *et al.* (2004) sketch a very deviating picture of sharply increasing inequality since 1997. They use Goskomstat aggregate data.

[6]Exceptions are Commander *et al.* (1999) and Förster *et al.* (2002). The former show graphs of the evolution of the Gini for different equivalence scales. But, although they find some rank reversals, they do not discuss this sensitivity. The latter use the square root of household size as the equivalence scale.

when calculated on per capita values ($\theta = 1$). There are corresponding rank reversals for other years. Hence some analysis of the robustness of the results for different equivalence scales seems appropriate here.

Equally surprising is the lack of a robust analysis with respect to the choice of the inequality measure and its underlying normative assumptions. As usual, the majority of the papers use the Gini coefficient to investigate inequality changes. Yet, the reported findings do not seem to be robust to this choice either. In Commander *et al.* (1999), for example, inequality increases between 1992 and 1996 when judged by means of the Gini or the bottom sensitive Theils. But when inequality is measured by means of the top sensitive Theil, ordinally equivalent to the coefficient of variation, inequality unambiguously decreases over the same period. More robust methods, like the ones presented, are definitely appropriate.

3.2. *Empirical Illustration*

We use household size to divide households in seven different needs groups, ranging from 1 to 7+ (7 or more individuals). Next, we choose the lower bounds equal to unity: larger households need at least the same *household* income compared to smaller ones to reach the same living standards, or $\alpha = (1, 1, \ldots, 1)$. For the upper bounds, we ensure that the scale itself is bounded by the number of persons in the household: in terms of per capita income, larger households need not more *per capita* income compared to smaller ones to reach the same living standards, or $\beta = (1, 2/1, 3/2, \ldots, 7/6)$. Finally, we set a_1 equal to the maximal household income over the different rounds. We discuss the sensitivity of our results with respect to these choices later on.

Table 2 summarizes our results for the bounded index test, for different values of the inequality aversion parameter ρ. In the last column, we mention the dominances found by the FHT-criterion (for the same bounds α, β and the same a_1).

The total number of rankings (in the last row) obviously depends on the choice of the parameter ρ. But for a wide range of ρ-values from 0.20 to 10, the number of dominances ranges from a minimum of 15 to a maximum of 20 (out of 21 possible comparisons). It is clear that the serious decline in social welfare in the first half of the period, followed by a recovery afterwards, is detected properly. In contrast, the performance of the FHT-criterion is disappointing: only 3 of 21 comparisons can be ranked unambiguously: 1998 is dominated by 2000, 2001 and 2002. It is quite striking that it cannot identify the steep fall in average per capita expenditures up to 1998 in combination with a slightly increasing inequality as a social welfare loss.

In addition, the lack of power of the FHT-criterion is robust with respect to the choices made. First, we assess the FHT criterion for all incomes $y \in [0, a_1]$. Reducing the impact of insignificant crossings, we could choose instead to check the FHT-criterion on a finite number of points n, e.g. $\{a_1/n, 2a_1/n, \ldots, a_1\}$ for some n. This typically adds two dominances (even for large values of n): 2000 is dominated by 2001 and 2002. Second, choosing a different needs classification, for example, by dividing households in n groups, with $3 \leq n \leq 7$ (household size equal to $1, \ldots, n+$), adds one dominance (2000 dominated by 2002). Third, choosing

TABLE 2

Dominances According to the Bounded Index Test for Different Values of ρ

year	ρ	0.20	0.50	1.00	1.50	2.00	3.00	5.00	10.0	all ρ	FHT
94 was better (+) or worse (–) compared to year (in rows) using ρ or FHT (in columns)											
95		+	+	+	+	+	+	+	+	+	
96		+	+	+	+	+	+				
98		+	+	+	+	+	+	+	+	+	
00		+	+	+	+	+	+	+	+	+	
01		+	+	+	+	+	+	+	+	+	
02		+	+	+	+	+					
95 was better (+) or worse (–) compared to year (in rows) using ρ or FHT (in columns)											
96		+	+	+	+	+	–	–	–		
98		+	+	+	+	+	+				
00		+	+	+	+	+	–	–	–		
01		+	+	+	+	+	–	–	–		
02		+	+	+	+	+	–	–	–		
96 was better (+) or worse (–) compared to year (in rows) using ρ or FHT (in columns)											
98		+	+	+	+	+	+	+	+	+	
00			+	+	+	+					
01								+	+		
02				–	–	–	–	–	–		
98 was better (+) or worse (–) compared to year (in rows) using ρ or FHT (in columns)											
00		–	–	–	–	–	–	–	–	–	–
01		–	–	–	–	–	–	–	–	–	–
02		–	–	–	–	–	–	–	–	–	–
00 was better (+) or worse (–) compared to year (in rows) using ρ or FHT (in columns)											
01			–	–	–	–					
02			–	–	–	–	–	–			
01 was better (+) or worse (–) compared to year (in rows) using ρ or FHT (in columns)											
02		+		–	–	–	–	–	–		
total		18/21	18/21	20/21	20/21	19/21	17/21	15/21	15/21	8/21	3/21

higher values for a_1 (smaller values are not allowed) decreases the number of successful rankings for the FHT-criterion.

Finally, choosing a smaller set of bounded equivalence scales (e.g. by setting α and β such that the equivalence scale for each household type k (which is also the household size) belongs to $[k^{0.5-n}, k^{0.5+n}]$ for $n \in [0, 0.5]$) might increase the number of dominances, but only for small n. For example, choosing $n = 0$ (equivalence scale is square-root of household size) or $n = 0.1$ adds four dominances, whereas choosing $n = 0.2, 0.3, 0.4$ does not add dominances.

To conclude, recall Table 1, which classifies the different ways to deal with heterogeneous welfare comparisons. In Table 3, we list the number of dominances (on a total of 21 bilateral comparisons) using six different methods.

While the bounded index test finds between 15 and 20 dominances—depending on the inequality aversion parameter—the FHT-criterion only detects three dominances. If we move upwards—choosing $\alpha = \beta$—we (obviously) get a complete ranking (21 dominances) for the bounded index test and between 3 and

TABLE 3

The Number of Dominances for the Different Criteria

	Index	Dominance
Cardinal	21	[3,7]
Intermediate	[15,20]	3
Ordinal	[1,11]	0

7 dominances for the FHT-criterion (depending on the equivalence scale specification). This points to the fact that the lack of ranking power of the FHT-criterion is not caused by the robustness with respect to the needs specification, but to the robustness with respect to the well-being aggregation. If we move downwards—keeping $\alpha = (1, \ldots, 1)$ and letting $\beta \to (1, \infty, \ldots, \infty)$—we find between 1 and 11 dominances, using the sequential index test (Proposition 2). For example, considering moderate values of ρ equal to 1.5 and 2, we can make 11 bilateral comparisons each. This is in sharp contrast with the the zero score of the ordinal dominance criteria (Bourguignon's dominance criterion and the SGLD criterion) in the lower-right corner.

4. Conclusion

Fleurbaey *et al.* (2003) introduce a criterion to measure welfare in a robust way, i.e. robust with respect to both the needs specification (via a bounded set of equivalence scales) and the aggregation procedure (via the generalized Lorenz dominance (GLD) criterion). Choosing the bounded set of equivalence scales as small as possible, their criterion reduces to Ebert's (1999) cardinal GLD approach, i.e. the GLD criterion applied to household incomes, both divided and weighted by the (unique) equivalence scale. Choosing the bounded set as wide as possible, their criterion is equivalent with one of Bourguignon's (1989) dominance criteria.

We propose a bounded (iso-elastic) index test to make welfare comparisons which are robust with respect to the needs specification, but depend on the chosen inequality aversion parameter. Choosing the bounded set as small as possible, we get a cardinal index test in line with Ebert's (1997) weighting scheme: an index applied to household incomes, both divided and weighted by the (unique) equivalence scale. Choosing bounds as wide as possible, we obtain a (new) sequential index test, i.e. checking—on the basis of the iso-elastic index—whether welfare is higher for the neediest income units only, for the most and second most needy only, and so on.

In comparison with Fleurbaey *et al.*'s (2003) bounded dominance criterion, our criterion is more complete, but less robust. To illustrate the trade-off between completeness and robustness, we compare the ranking power of the bounded dominance and the bounded index test using the Russian RLMS (Russian Longitudinal Monitoring Survey) data between 1994 and 2002.

Our empirical illustration suggests why the bounded index test might be an interesting tool for practitioners. Indeed, the cost of robustness with respect to the well-being aggregation turns out to be high. Contrary to the bounded index test, the bounded dominance criterion can hardly detect welfare changes in Russia, in spite of the increasing inequality and strongly declining average well-being in the period before the financial crisis (1994–98), and the opposite afterwards (1998–2002). Our empirical illustration revealed that the bounded dominance criterion also performs badly when using a unique equivalence scale, which indicates that using generalized Lorenz dominance (and hence the robustness with respect to aggregation over individuals) is the main culprit. Therefore the bounded index might be an attractive alternative. It is capable to keep full robustness with respect to the needs specification, some robustness with respect to the aggregation issue (by selecting some reasonable values of the inequality aversion parameter), while still detecting major parts of the welfare evolution. The bounded index allows practitioners to go from the most complete (a bounded index test choosing bounds as small as possible) to the least complete welfare analysis (a bounded dominance test choosing reasonable bounds) and to see to what extent the drop in completeness is due to the robust aggregation and to the robust needs specification.

Appendix

Proof of Proposition 1

We focus on the case $\rho \neq 1$; the other case $\rho = 1$ is analogous. By definition of the unanimity quasi-ordering $\succsim_{(\alpha,\beta,a_1,\rho)}$, we have $F \succsim_{(\alpha,\beta,a_1,\rho)} G$ if and only if

(8) $$\Delta W = \sum_{k\in\mathbb{K}} \int_0^{\bar{s}_k} I(y, m_k)\, dH_k^1(y) \geq 0 \quad \text{for all } \mathbf{m} \in \mathcal{M}(\alpha, \beta).$$

Because (for all $k \in \mathbb{K}$) (i) $a_1 \geq \bar{s}_k$ and (ii) the function dH_k^1 is zero outside its support, we can rewrite the welfare difference ΔW using the definition of I as follows:

$$\begin{aligned}\Delta W &= \sum_{k\in\mathbb{K}} \int_0^{\bar{s}_k} \frac{m_k}{1-\rho}\left(\left(\frac{y}{m_k}\right)^{1-\rho} - \left(\frac{a_1}{m_k}\right)^{1-\rho}\right) dH_k^1(y) \\ &= \sum_{k\in\mathbb{K}} (m_k)^{\rho} \int_0^{\bar{s}_k} \frac{1}{1-\rho}\left((y)^{1-\rho} - (a_1)^{1-\rho}\right) dH_k^1(y).\end{aligned}$$

Define

$$b_k = \int_0^{\bar{s}_k} \frac{1}{1-\rho}\left((y)^{1-\rho} - (a_1)^{1-\rho}\right) dH_k^1(y) \quad \text{for all } k = 1, \ldots, K.$$

We have $F \succsim_{(\alpha,\beta,a_1,\rho)} G$ if and only if

(9) $$\sum_{k\in\mathbb{K}} b_k (m_k)^{\rho} \geq 0 \quad \text{for all } \mathbf{m} \in \mathcal{M}(\alpha, \beta).$$

Let $Z_{K+1}^{\circ} : x \mapsto 0$. Define functions Z_k° recursively (starting from $k = K$ downwards to $k = 3$) as:

$$Z_k^{\circ}:\left[\prod_{i=1}^{k-1}\alpha_i,\prod_{i=1}^{k-1}\beta_i\right]\to\mathbb{R}: m\mapsto \min_{\alpha_k m\le x\le\beta_k m}\left\{b_k(x)^\rho+Z_{k+1}^{\circ}(x)\right\}.$$

We get

$$\begin{aligned}
(9) &\Leftrightarrow b_1+\sum_{k=2}^{K}b_k(m_k)^\rho \quad \text{for all } \mathbf{m}\in\mathcal{M}(\alpha,\beta)\\
&\Leftrightarrow b_1+\sum_{k=2}^{K-1}b_k(m_k)^\rho+Z_K^{\circ}(m_{K-1})\ge 0 \quad \text{for all } \mathbf{m}\in\mathcal{M}(\alpha,\beta)\\
&\Leftrightarrow b_1+\sum_{k=2}^{K-2}b_k(m_k)^\rho+Z_{K-1}^{\circ}(m_{K-2})\ge 0 \quad \text{for all } \mathbf{m}\in\mathcal{M}(\alpha,\beta)\\
&\Leftrightarrow \dots\\
&\Leftrightarrow b_1+b_2(m_2)^\rho+Z_3^{\circ}(m_2)\ge 0 \quad \text{for all } \mathbf{m}\in\mathcal{M}(\alpha,\beta)\\
&\Leftrightarrow b_1+b_2(m_2)^\rho+Z_3^{\circ}(m_2)\ge 0 \quad \text{for all } \alpha_2\le m_2\le\beta_2, \text{ as required.}
\end{aligned}$$

Proof of Proposition 2

Again, we focus on the case $\rho \neq 1$; the other case is analogous. Recall equation (9) and the definition of $\mathcal{M}(\alpha,\ \beta)$. Choosing $\alpha = (1, \dots, 1)$ and $\beta \to (\infty, \dots, \infty)$, we have $F \succsim_{(\alpha,\beta,a_1,\rho)} G$ if and only if

$$b_1+\sum_{k=2}^{K}b_k(m_k)^\rho\ge 0 \quad \text{for all } m_K\ge m_{K-1}\ge\dots\ge m_2\ge 1, \tag{10}$$

with

$$b_k=\int_0^{\bar{s}_k}\frac{1}{1-\rho}\left((y)^{1-\rho}-(a_1)^{1-\rho}\right)dH_k^1(y) \quad \text{for all } k=1,\dots,K.$$

We show that equation (10) is equivalent with

$$\sum_{k=i}^{K}b_k\ge 0 \quad \text{for all } i=1,\dots,K. \tag{11}$$

Sufficiency

Suppose (11) holds; thus, choosing $i = 1$, we must have $b_1+\Sigma_{k=2}^{K}b_k\ge 0$. Since $(m_2)^\rho \geq 1$, for all $m_2 \geq 1$, and $\Sigma_{k=2}^{K}b_k\ge 0$ (from (11) for $i = 2$) we must have

$$\begin{aligned}
b_1+(m_2)^\rho\sum_{k=2}^{K}b_k\ge 0, &\quad \text{for all } m_2\ge 1,\\
b_1+(m_2)^\rho b_2+(m_2)^\rho\sum_{k=3}^{K}b_k\ge 0, &\quad \text{for all } m_2\ge 1.
\end{aligned}$$

Since $(m_3)^\rho \geq (m_2)^\rho$, for all $m_3 \geq m_2$ and $\Sigma_{k=3}^{K} b_k \geq 0$ (from (11) for $i = 3$) we must have

$$b_1 + (m_2)^\rho\, b_2 + (m_3)^\rho \sum_{k=3}^{K} b_k \geq 0, \quad \text{for all } m_3 \geq m_2 \geq 1,$$

$$b_1 + \sum_{k=2}^{3} b_k (m_k)^\rho + (m_3)^\rho \sum_{k=4}^{K} b_k \geq 0, \quad \text{for all } m_3 \geq m_2 \geq 1.$$

We might proceed in this way, until we finally get

$$b_1 + \sum_{k=2}^{K} b_k (m_k)^\rho \geq 0, \quad \text{for all } m_K \geq m_{K-1} \geq \ldots \geq m_2 \geq 1, \text{as required.}$$

Necessity

Suppose (10) holds, but not (11). More precisely, there exists a $j \in \mathbb{K}$ such that

$$\sum_{k=j}^{K} b_k < 0. \tag{12}$$

1. First, suppose $j = 1$. As (10) holds, we might choose an equivalence scale vector $\mathbf{m} = (1, \ldots, 1)$, and we obtain

$$\sum_{k=1}^{K} b_k \geq 0, \tag{13}$$

   which contradicts equation (12) for $j = 1$.
2. Suppose $1 < j \leq K$. Equation (12) and (13) together, we must have

$$\sum_{k=1}^{j-1} b_k > 0. \tag{14}$$

   Choose an equivalence scale vector $\mathbf{m}$ with $1 = m_1 = \ldots = m_{j-1} \leq m_j = \ldots = m_K = \eta$ in (10); we must have

$$\sum_{k=1}^{j-1} b_k + (\eta)^\rho \sum_{k=j}^{K} b_k \geq 0,$$

   which cannot be true for all values of $\eta \geq 1$, given equations (12) and (14).

Some Summary Statistics for the RLMS

In Table A1 we present for each round of the RLMS-survey: (i) the number of observations in this round (denoted by n); (ii) the average per capita real expenditures (denoted by μ) expressed in Roubles of 1992; (iii) the Gini coefficient of per capita real expenditures (weighted by the number of individuals); (iv) the proportions (denoted by p); and (v) the average real expenditures (denoted by y) in the different needs groups (based on household size).

TABLE A1

RLMS Round	n	μ	Gini		Household Size 1	2	3	4	5	6	7+
5 (1994)	3,761	2,982	41.29	p	17.6	28.6	23.1	21.0	6.3	1.9	1.5
				y	3,641	7,585	9,174	11,029	11,649	10,431	16,007
6 (1995)	3,594	2,595	42.06	p	18.8	27.9	22.8	20.5	6.7	1.9	1.4
				y	3,546	6,750	7,842	9,121	10,367	11,700	11,295
7 (1996)	3,464	2,295	44.38	p	19.3	27.8	22.6	20.3	6.4	2.4	1.2
				y	2,990	5,611	7,174	8,139	9,811	9,274	13,225
8 (1998)	3,268	1,717	43.08	p	19.6	28.1	22.6	20.0	6.1	2.1	1.5
				y	2,225	3,980	5,233	6,417	7,369	7,447	9,959
9 (2000)	3,126	1,985	41.05	p	20.3	27.9	22.1	19.9	6.1	2.3	1.4
				y	2,367	4,351	6,378	7,644	8,105	9,072	12,767
10 (2001)	3,225	2,214	39.74	p	21.5	27.8	21.6	19.5	6.2	1.8	1.6
				y	2,698	4,982	6,703	8,509	9,253	11,448	13,693
11 (2002)	3,224	2,267	39.87	p	21.1	27.7	21.9	19.6	6.2	2.0	1.5
				y	2,772	4,900	7,483	8,784	9,147	9,459	12,471

References

Atkinson A. B., "Measuring Poverty and Differences in Family Composition," *Economica*, 59, 1–16, 1992.

Atkinson, A. B. and F. Bourguignon, "Income Distribution and Differences in Needs," in G. R. Feiwel (ed.), *Arrow and the Foundations of the Theory of Economic Policy*, Macmillan, London, 1987.

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

Capéau, B. and E. Ooghe, "On Comparing Heterogeneous Populations: Is There Really a Conflict Between the Pareto Criterion and Inequality Aversion?" CES discussion paper DPS 04.07, Leuven, 2004.

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

Clark, S., R. Hemming, and D. Ulph, "On Indices for the Measurement of Poverty," *Economic Journal*, 91, 515–26, 1981.

Commander, S., A. Tolstopiatenko, and R. Yemtsov, "Channels of Redistribution: Inequality and Poverty in the Russian Transition," *The Economics of Transition*, 7(2), 411–47, 1999.

Cowell, F. A. and M. Mercader-Prats, "Equivalence Scales and Inequality," in J. Silber (ed.), *Income Inequality Measurement: From Theory to Practice*, Kluwer, Dewenter, 1999.

Deaton, A., "Counting the World's Poor: Problems and Possible Solutions," *World Bank Research Observer*, 16(2), 125–47, 2001.

Ebert, U., "Social Welfare When Needs Differ: An Axiomatic Approach," *Economica*, 64, 233–44, 1997.

———, "Using Equivalent Income of Equivalent Adults to Rank Income Distributions," *Social Choice and Welfare*, 16, 233–58, 1999.

Ebert, U. and P. Moyes, "Equivalence Scales Reconsidered," *Econometrica*, 71(1), 319–43, 2003.

Fleurbaey, M., C. Hagneré, and A. Trannoy, "Welfare Comparisons with Bounded Equivalence Scales," *Journal of Economic Theory*, 110, 309–36, 2003.

Förster, M., D. Jesuit, and T. Smeeding, "Regional Poverty and Income Inequality in Central and Eastern Europe: Evidence from the Luxembourg Income Study," LIS Working Paper 324, Maxwell School of Citizenship and Public Affairs, Syracuse University, New York, 2002.

Galbraith, J. K., L. Krytynskaia, and Q. Wang, "The Experience of Rising Inequality in Russia and China During the Transition," *European Journal of Comparative Economics*, 1(1), 87–106, 2004.

Glewwe, P., "Household Equivalence Scales and the Measurement of Inequality," *Journal of Public Economics*, 44, 211–16, 1991.

Jenkins, S. P. and P. J. Lambert, "Ranking Income Distributions When Needs Differ," *Review of Income and Wealth*, 39, 337–56, 1993.

Kislitsyna, O., "Income Inequality in Russia During Transition. How Can It Be Explained?" Working Paper Series No 03/08, Economics Education and Research Consortium, Moscow, 2003.

Lambert, P. J., *The Distribution and Redistribution of Income. A Mathematical Analysis*, 3rd edition, Manchester University Press, Manchester, 2001.
Lambert, P. J. and X. Ramos, "Welfare Comparisons: Sequential Procedures for Heterogeneous Populations," *Economica*, 69, 549–62, 2002.
Lokshin, M. and B. M. Popkin, "The Emerging Underclass in the Russian Federation: Income Dynamics 1992–96," *Economic Development and Cultural Change*, 47, 803–29, 1999.
Moyes, P., "Comparaisons de distributions hétérogènes et critères de dominance," *Economie et Prévision*, 138–9, 125–46, 1999.
Mroz, T. A., L. B. O. M. A. Henderson, and B. M. Popkin, "Monitoring Economic Conditions in the Russian Federation: The Russian Longitudinal Monitoring Survey 1992–2003," Carolina Population Center, University of North Carolina at Chapel Hill, North Carolina, 2004.
Pyatt G., "Social Evaluation Criteria," in C. Dagum and M. Zenga (eds), *Income and Wealth Distribution, Inequality and Poverty*, Springer-Verlag, Heidelberg, 243–53, 1990.
Shorrocks, A. F., "Ranking Income Distributions," *Economica*, 50, 3–17, 1983.
———, "Inequality and Welfare Evaluation of Heterogeneous Income Distributions," *Journal of Economic Inequality*, 2, 193–218, 2004.
World Bank, *2004 World Development Indicators*, Washington, DC, 2004.
Yemtsov, R., "Quo Vadis: Inequality and Poverty Dynamics Across Russian Regions in 1992–2000," WIDER Discussion Paper 2003/67, Helsinki, 2003.