



## INEQUALITY COMPARISONS WITH ORDINAL DATA

BY STEPHEN P. JENKINS\*

*LSE, ISER (University of Essex), and IZA*

Non-intersection of appropriately defined Generalized Lorenz (GL) curves is equivalent to a unanimous ranking of distributions of ordinal data by all Cowell and Flachaire (*Economica*, 2017) indices of inequality and by a new index based on GL curve areas. Comparisons of life satisfaction distributions for six countries reveal a substantial number of unanimous rankings. The GL dominance criteria are compared with other criteria including the dual-*H* dominance criteria of Gravel, Magdalou, and Moyes (*Economic Theory*, 2020).

**JEL Codes:** D31, D63, I31

**Keywords:** inequality, ordinal data, Generalized Lorenz dominance, H dominance, Hammond transfers, life satisfaction, World Values Survey

### 1. INTRODUCTION

Cowell and Flachaire (2017) provide an approach to measuring inequality of ordinal data such as life satisfaction, happiness, and self-assessed health status that differs significantly from the approach taken in most recent research. This paper builds on Cowell and Flachaire’s work by adding dominance results and a new inequality index, illustrates them using cross-national data about life satisfaction distributions, and compares the approach with others.

Since the critique by Allison and Foster (2004), most economists have accepted that it is inappropriate to assess ordinal data inequality using the tools developed to assess the inequality of cardinal data on income and wealth. The latter methods associate greater inequality with greater dispersion about the mean, but the mean is an improper benchmark for an ordinal variable. For ordinal variables, Allison and Foster (2004) propose instead that greater inequality means greater spread about the median and they demonstrate that, for distributions with the same median, a unanimous ordering by all indices incorporating this concept is equivalent to “*S*-dominance”—a particular configuration of cumulative distribution functions.<sup>1</sup> Allison and Foster (2004) and other researchers, including Abul Naga

*Note:* The research was part-supported by core funding of the Research Centre on Micro-Social Change at the Institute for Social and Economic Research by the University of Essex and the UK Economic and Social Research Council (award ES/L009153/1). I gratefully acknowledge the hospitality of the School of Economics, University of Queensland, where I was based when writing the first version of this paper, and thank the anonymous referees for their helpful comments, especially “referee 2,” and Yongsheng Xu for clarificatory discussion.

\*Correspondence to: Stephen P. Jenkins, Department of Social Policy, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK (s.jenkins@lse.ac.uk).

<sup>1</sup>See also Kobus (2015) for characterization results.

© 2020 The Authors. *Review of Income and Wealth* published by John Wiley & Sons Ltd on behalf of International Association for Research in Income and Wealth

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

and Yalcin (2008) and Apouey (2007), have developed inequality indices consistent with  $S$ -dominance. A distinguishing feature of the Allison-Foster approach is that it measures inequality in terms of polarization: “inequality” is maximized when half the population has the lowest value on the ordinal scale and half the population has the largest value. Cowell and Flachaire’s (2017) inequality indices are different because greater inequality reflects greater spread in a sense other than greater polarization. However, no dominance results currently exist for Cowell-Flachaire indices.

I show that non-intersection of appropriately defined Generalized Lorenz (GL) curves is equivalent to a unanimous ranking of distributions by all Cowell and Flachaire (2017) indices of inequality and by a new index based on areas below GL curves. The results are not restricted to distributions with the same median. Comparisons of life satisfaction distributions for six countries derived from World Values Survey data show that the new dominance results reveal a substantial number of unanimous inequality rankings.

I use Cowell and Flachaire’s “peer-inclusive downward-looking” definition of individual status (explained below) because this definition is consistent with the focus in the median-related inequality measurement literature. (Cumulative distribution functions are the building blocks in common.) There are analogous results for Cowell and Flachaire’s “peer-inclusive upward-looking” status definition but, for brevity, I summarize these in the Appendix.<sup>2</sup>

I also demonstrate that rankings according to the GL criteria differ from rankings according to the dual  $H$ -dominance results of Gravel *et al.* (2020) based on the concept of Hammond transfers. I compare the elementary transformations that underlie each approach and provide empirical illustrations of how they order World Values Survey life satisfaction distributions differently.

Various supplementary materials cited in the main text are reported in the Appendix.

## 2. COWELL-FLACHAIRE INEQUALITY INDICES FOR ORDINAL DATA

The well-being of each of  $N$  individuals is measured on an ordinal scale characterized by a set of numerical labels  $(l_1, l_2, \dots, l_K)$ , with  $-\infty < l_1 < l_2 < \dots < l_K < \infty$ , and  $K \geq 3$ . Thus, the distribution of well-being is summarized by an ordered categorical variable. The proportion of individuals in the  $k$ th category is denoted  $f_k$  with  $0 \leq f_k \leq 1$  and  $\sum_{k=1}^K f_k = 1$ . The proportion of individuals in the  $k$ th category or lower is  $F_k$ , with  $F_k = \sum_{j=1}^k f_j$  and  $F_K = 1$ .

Cowell and Flachaire (2017) propose a two-step approach to inequality measurement for ordinal data. First, decide how to summarize “status,”  $s_i$ , for each individual  $i = 1, 2, 3, \dots, N$ . In particular, Cowell and Flachaire’s “peer-inclusive downward-looking” status of an individual with scale level  $k$  is  $F_k$ , and hence does not depend on the specific values attached to  $(l_1, l_2, \dots, l_K)$ . That is, the measure is scale independent.

<sup>2</sup>No researchers have used Cowell and Flachaire’s (2017) “peer-exclusive” definitions in applied work, not even the authors themselves.

Second, define inequality as an aggregate summary of the “distances” between each person’s status and an appropriate common reference value. A shortening of the distances means less inequality. Cowell and Flachaire argue persuasively that the common reference status value for a peer-inclusive status measure should be the maximum value, that is, 1 (the maximum of  $F_k$ ). With some auxiliary axioms including a requirement that the minimum value of the inequality index is 0 (when all individuals are in the same category), Cowell and Flachaire (2017) characterize a one-parameter family of inequality indices,  $I(\alpha)$ , with  $0 \leq \alpha < 1$ :

$$(1) \quad I(\alpha) = \frac{1}{\alpha(\alpha-1)} \left[ \frac{1}{N} \sum_{i=1}^N s_i^\alpha - 1 \right], \quad 0 < \alpha < 1;$$

$$I(0) = -\frac{1}{N} \sum_{i=1}^N \log(s_i).$$

The smaller that  $\alpha$  is, the greater the weight that is put on small status values relative to high status values. Cowell and Flachaire also cite a closely related class of “Atkinson-like” indices,  $A(\alpha)$ :

$$(2) \quad A(\alpha) = 1 - \left( \frac{1}{N} \sum_{i=1}^N s_i^\alpha \right)^{\frac{1}{\alpha}}, \quad 0 < \alpha < 1;$$

$$A(0) = 1 - \left( \prod_{i=1}^N s_i \right)^{\frac{1}{N}}.$$

$A(\alpha)$  equals one minus the generalized mean of order  $\alpha$  of status, and is a monotonically increasing transformation of the corresponding  $I(\alpha)$  index.<sup>3</sup>

Let  $D(s)$  denote an index that is a monotonically decreasing Schur-convex function of the distribution of status  $s$ , and  $\mathcal{D}$  denote the set of all such functions. Clearly, every  $I(\alpha)$  and  $A(\alpha)$  index belongs to  $\mathcal{D}$ . Also consider the class of indices  $W(s) = G(D(s))$  where  $G(\cdot)$  is a monotonically decreasing function.  $W(s)$  is of a similar form to the social evaluation functions commonly used in income distribution analysis: it is a monotonically increasing Schur-concave function of individual status (Marshall *et al.*, 2011, Table 1, case xii). Let  $\mathcal{W}$  denote the set of monotonically increasing Schur-concave social evaluation functions. Indices  $E(\alpha) = 1 - A(\alpha)$ ,  $0 \leq \alpha < 1$ , are examples of members of  $\mathcal{W}$ .

The next section presents a tractable method based on Generalized Lorenz curve comparisons for assessing whether one distribution of status is unambiguously more (un)equal than another according to all indices belonging to classes  $\mathcal{D}$  and  $\mathcal{W}$  regardless of the differences in social judgements encapsulated in them.

<sup>3</sup>Because every individual that provides the same response on the scale has the same status, the index expressions (1) and (2) can also be expressed in terms of sums over scale levels rather than sums (or products) over individuals. See Cowell and Flachaire (2017, eqn. 21) and Section 5 below.

TABLE 1  
CROSS-NATIONAL COMPARISONS OF LIFE SATISFACTION DISTRIBUTIONS: SUMMARY OF *F*- AND  
GL-DOMINANCE CHECKS

		Country <i>y</i>					
		AU	CA	GB	NZ	US	ZA
Country <i>x</i>	AU		<	<	<	–	–
	CA	–		–	–*	>	–
	GB	>	–		–	>	–
	NZ	–	<	<		>	>
	US	–	–	–	–		–
	ZA	<	<	<	<	<	

Notes: Entries above the diagonal summarize checks for *F*-dominance (first-order dominance): “>”, *x F*-dominates *y*; “<”, *y F*-dominates *x*; “–”, no dominance. \*: NZ *S*-dominates CA. Entries below the diagonal summarize GL dominance checks: “>”, *x* is more equal than *y*; “<”, *y* is more equal than *x*; “–”, no dominance.

### 3. GENERALIZED LORENZ CURVES FOR DISTRIBUTIONS OF (PEER-INCLUSIVE) STATUS, AND AN INEQUALITY DOMINANCE RESULT

I define the Generalized Lorenz (GL) curve for the distribution of status,  $GL(s, p)$  given  $0 \leq p \leq 1$ , following Shorrocks (1983) closely. With the elements of the distribution of status placed in ascending order, that is,  $s_1 \leq s_2 \leq s_3 \leq \dots \leq s_N$ , we have:

$$(3) \quad GL\left(s, \frac{m}{N}\right) = \frac{1}{N} \sum_{i=1}^m s_i, m = 1, \dots, N, \text{ and } GL(s, 0) = 0.$$

The GL curve is drawn using straight lines to connect adjacent points of the form  $\{m/N, GL(s, m/N)\}$ . The vertices of the curve are at  $\{p_0 = 0, 0\}$  and  $\{p_k, \sum_{j=1}^k f_j F_j\}$  for each  $k = 1, \dots, K$  with  $p_k = F_k$ . The GL ordinate at  $p = 1$  is the arithmetic mean of the status distribution (with a limiting value of 0.5 as  $K \rightarrow \infty$ ). Figure 1 provides an illustrative example for the case  $K = 4$ .

The 45° ray from (0, 0) to (1, 1) is the GL curve representing complete equality—when all individuals have the same scale value and hence the same status. With inequality, the GL curve lies below the 45° ray and, intuitively, the further below the ray the curve is, the greater is inequality.

One can demonstrate that a unanimous ranking of a pair of distributions in terms of their equality (or inequality) is equivalent to the non-crossing of their GL curves:

*Result 1:* For two status distributions  $s$  and  $s'$ ,  $W(s) \geq W(s')$  for all  $W(\cdot) \in \mathcal{W}$  iff  $GL(s, p) \geq GL(s', p)$  for all  $p$ .

*Result 2:*  $D(s) \leq D(s')$  for all  $D(\cdot) \in \mathcal{D}$  iff  $GL(s, p) \geq GL(s', p)$  for all  $p$ .

Result 1 follows directly from Shorrocks (1983, Theorem 2). Result 2 follows from Result 1 and the relationship between  $D(\cdot)$  and  $W(\cdot)$ .

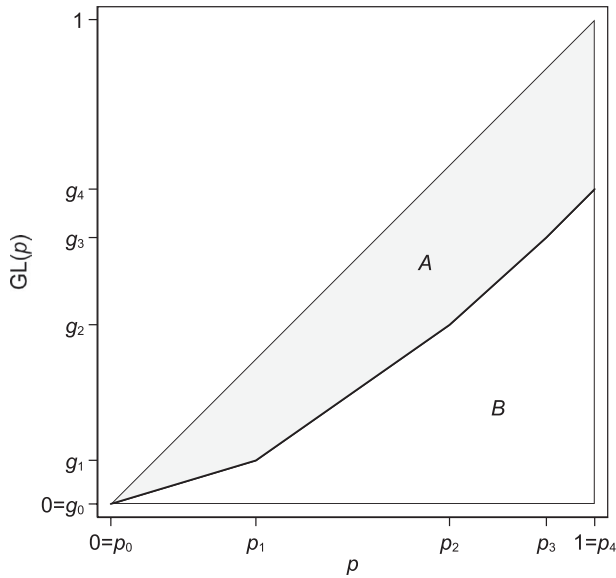


Figure 1. A Generalized Lorenz (GL) Curve for the Distribution of Status ( $K = 4$ )  
 Note: Area-based inequality index  $J = A/(A + B) = 1 - 2B$ .

The connection between GL curve location and inequality suggests a new index of inequality for ordinal data,  $J$ . With reference to Figure 1,  $J$  is the ratio of area  $A$  to area  $A + B$ ; equivalently,  $J$  equals 1 minus twice area  $B$ . It is Generalized Lorenz-consistent because a ranking of a pair of distributions by  $J$  is the same as the ranking by all  $D(\cdot) \in \mathcal{D}$  when the two GL curves do not cross. Using the expression for the vertices of the GL curve, and applying the Trapezium Formula, one can show that:

$$(4) \quad J = 1 - \sum_{j=0}^{K-1} (p_{j+1} - p_j) (GL_j + GL_{j+1}) = 1 - \sum_{j=0}^{K-1} f_{j+1} (GL_j + GL_{j+1}).$$

The minimum value of  $J$  is 0, achieved when there is perfect equality.

The dominance results relate to GL curves, not to Lorenz curves as some readers might expect. The reason is that the mean of the category labels is an inappropriate reference point (Allison and Foster, 2004), and hence also shares of the (labels) total are not a suitable building-block for inequality measurement in this context. Differences between observed status and a common status reference value are what matter for Cowell-Flachaire indices. (I return to this issue in Section 5.)

The situation considered here has analogies with the measurement of poverty. Non-intersection of two Three Is of Poverty (TIP) curves is equivalent to a unanimous ranking according to all “generalized poverty gap” poverty indices (Jenkins and Lambert, 1997). But a TIP curve shows, at each  $p$ , the vertical distance between two GL curves, one for the distribution of income censored above at the poverty line and the other for the distribution in which every income equals the poverty line (a distribution with perfect equality). Index  $J$ , based on the area

between two GL curves, is analogous to the Shorrocks (1995) modified-Sen poverty index (which equals twice the area beneath a TIP curve, that is, twice the area between two GL curves).

Polarized distributions and uniform distributions provide potential maximum-inequality benchmarks for ordinal data. As mentioned earlier, inequality indices in the Allison-Foster (2004) tradition reach their maximum if the distribution is polarized. But do Cowell-Flachaire indices reach a maximum in this case, and what if the distribution is uniform?

Results 1 and 2 imply that inequality is greater for a uniform distribution than for a polarized distribution according to  $J$  and all  $D(\cdot) \in \mathcal{D}$ .<sup>4</sup> With a polarized distribution,  $N/2$  individuals have the minimum scale value (status  $F_1 = 0.5$ ) and  $N/2$  have the maximum value ( $F_K = 1$ ) for all possible  $K$ . The corresponding GL curve has two segments connecting points  $\{(0, 0), (0.5, 0.25), (1, 0.75)\}$ . In contrast, with a uniform distribution,  $f_k = 1/K$ , all  $k = 1, \dots, K$ , and the GL curve has vertices at  $\{k/K, k(k+1)/(2K^2)\}$  for each  $k$ . Exploiting the expression for a straight line between two points, one can show that the GL curve for the polarized distribution lies above the curve for a uniform distribution at all  $p$  and regardless of the value of  $K$ .<sup>5</sup> For example, to three decimal places (d.p.),  $J = 0.375$  for a polarized distribution and for a uniform distribution,  $J = 0.481$  if  $K = 3$ ,  $0.531$  if  $K = 4$ , and  $0.615$  if  $K = 10$ .  $I(0) = 0.347$  for a polarized distribution and, for a uniform distribution,  $I(0) = 0.501$  if  $K = 3$ ,  $0.592$  if  $K = 4$ , and  $0.807$  if  $K = 10$ .

Results 1 and 2 are also informative about whether  $J$  and all  $D(\cdot) \in \mathcal{D}$  reach their maximum values in the case of a uniform distribution. One can show that, if one starts from a uniform distribution and shifts a small number of individuals from one scale level to the next level up (or down), the pre- and post-shift GL curves intersect. Hence the non-uniform distribution may have greater inequality than the uniform distribution according to some Cowell-Flachaire indices. I provide a numerical illustration of this in Section 5.

#### 4. EMPIRICAL ILLUSTRATION: RANKING COUNTRIES BY LIFE SATISFACTION INEQUALITY

To illustrate the analysis, I use data about life satisfaction from the mid-2000s for six countries (Australia, Canada, Great Britain, New Zealand, Australia, USA, South Africa), drawn from the fifth wave of the World Values Survey (WVS), the latest available when this research was undertaken. The six countries are “white settler” economies plus their colonial mother country (Great Britain).<sup>6</sup> Life satisfaction is measured using a 10-point integer-valued scale ranging from 1

<sup>4</sup>The results cited in this paragraph require that  $N$  is sufficiently large so that any difference in the number of individuals in each category is negligible, with attention restricted to the two populated categories in the case of a polarized distribution.

<sup>5</sup>Similarly, one can also show that the GL curve for a uniform distribution over  $K + 1$  levels lies everywhere on or below the GL curve for a uniform distribution over  $K$  levels. Illustrating these results, Appendix Figure A1 shows Generalized Lorenz curves for a polarized distribution and uniform distributions with  $K = 3, 4, 5$ , and  $10$ .

<sup>6</sup>Ireland could be included in this description but there are no Irish data in the WVS. For additional dominance and inequality index comparisons based on WVS data, see Jenkins (2019).

(completely dissatisfied) to 10 (completely satisfied). The data are the same as those employed by Cowell and Flachaire (2017), though I analyze fewer countries. For more details about the data, see their paper and the WVS documentation (Inglehart *et al.*, 2014). All my estimates use the WVS-supplied sample weights and were derived using my Stata program *ineqord* (Jenkins, 2020).<sup>7</sup> Country-specific relative frequency distributions are shown in Appendix Figure A2.

Figure 2 shows the GL curves for two of the 15 possible pairwise cross-national comparisons. Panel (a) provides an example of inequality dominance: the GL curve for Britain lies every on or above the GL curve for South Africa, and hence life satisfaction inequality is lower in Britain than South Africa according to all indices  $J$  and  $D \in \mathcal{D}$ , and there is more equality according to all indices belonging to  $\mathcal{W}$ . In contrast, panel (b) shows that there is no unambiguous ranking of Australia and New Zealand. Their GL curves intersect four times. To assess whether inequality is higher in one or other of these two countries requires use of indices and the ordering derived may depend on the index used.

Table 1 summarizes the results of all 15 pairwise GL curve comparisons (entries below the main diagonal) as well as for checks for first-order stochastic dominance (“ $F$ -dominance”) based on comparisons of cumulative distribution functions (entries above the main diagonal).<sup>8</sup>

GL curves do not cross in 8 out of 15 pairwise comparisons, demonstrating that the inequality dominance result has empirical usefulness. The most striking finding concerns South Africa: its life satisfaction inequality is unambiguously greater than in each of the other five countries. No country stands out as being unambiguously more equal than every other country, but Great Britain is more equal than its comparator countries in three of its five comparisons.

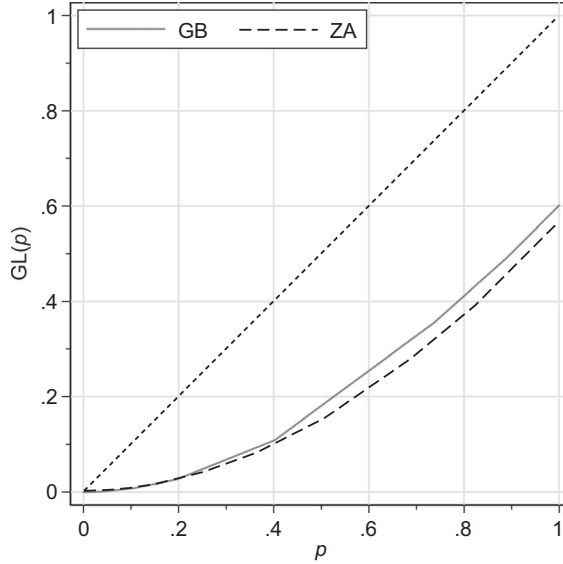
There is  $F$ -dominance in 8 of the 15 pairwise comparisons. For example, average life satisfaction in the USA is lower than in Canada, or Great Britain, or New Zealand, regardless of the life satisfaction scale that is used. However, the  $S$ -dominance criterion has little discriminatory power by comparison with the GL dominance criterion. There is only one case of  $S$ -dominance: there is greater spread away from the median in New Zealand than in Canada.  $S$ -dominance is rare partly because of the prevalence of  $F$ -dominance—if there is  $F$ -dominance, there cannot also be  $S$ -dominance (Allison and Foster, 2004)—and partly because median life satisfaction is lower in South Africa than in the other countries (seven rather than eight).  $S$ -dominance applies only to distributions with a common median—a restriction that does not apply to GL-dominance.

To derive a complete inequality ordering of the six countries an inequality index must be used. However, different indices incorporate different social judgments about how to assess differences in different parts of the life satisfaction distribution. It is, therefore, important to use a portfolio of indices to check the robustness of rankings. I report estimates for six GL-consistent indices in Figure 3:  $I(\alpha)$  for  $\alpha = 0, 0.25, 0.5, 0.75, \text{ and } 0.9$ , plus  $J$ . For comparison, I also include 3  $S$ -dominance-consistent indices from the Abul Naga and Yalcin (2007) class.

<sup>7</sup>The program can also be downloaded from within Stata using the command `ssc install ineqord`.

<sup>8</sup>See Appendix Figures A3 and A4 for charts showing all the pairwise GL curve and cumulative distribution function comparisons.

**(a) Dominance**



**(b) Non-dominance**

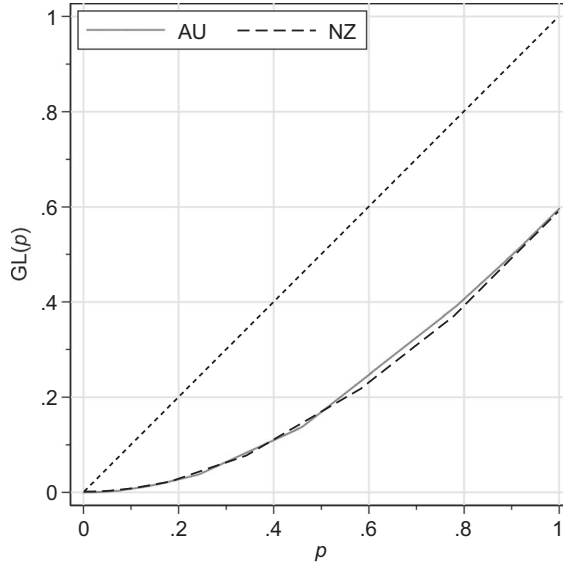


Figure 2. GL Curve Comparisons of Life Satisfaction Distributions: Selected Examples. (A) Dominance. (B) Non-Dominance

$ANY(1, 1)$  weights observations in categories above and below the median equally;<sup>9</sup>  $ANY(4, 1)$  is more sensitive to above-median spread than below-median spread; and  $ANY(1, 4)$  is the opposite (i.e. relatively bottom-sensitive). Figure 3 shows

<sup>9</sup>For a linear integer scale,  $ANY(1,1)$  is also equal to Apouey's (2007) P2(1) index. It is also known as the normalized average jump index.



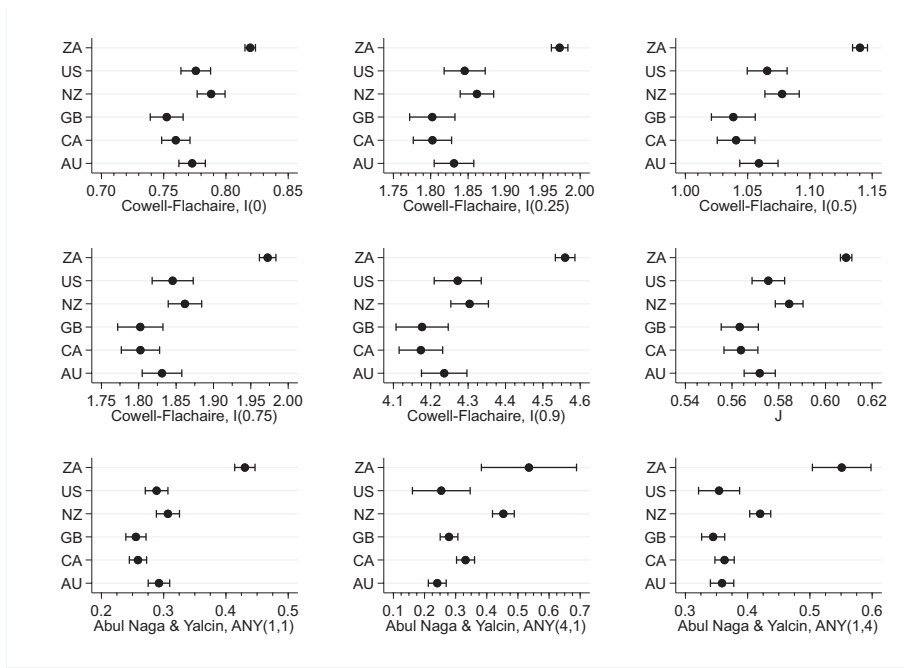


Figure 3. Estimates of Life Satisfaction Inequality Across Countries: 9 Indices

Notes: The figure shows point estimates of inequality indices and their associated 95 percent confidence intervals.

point estimates and 95 percent confidence intervals. I estimate standard errors using a repeated half-sample bootstrap approach in order to appropriately account for the sample weights (Saigo *et al.*, 2001; Van Kerm, 2013) with 500 bootstrap replications.

Consider the  $I(\alpha)$  estimates. Figure 3 confirms that inequality is distinctly greater in South Africa than in every other country. For example, according to  $I(0)$ , South Africa’s inequality is 4 percent larger than NZ’s (with the null hypothesis of no difference decisively rejected: test statistic = 5.2).<sup>10</sup> According to  $I(0.9)$ , the difference is 6 percent (test statistic = 4.8). Canada and Great Britain appear to have the lowest inequality according to all five  $I(\alpha)$  estimates. (Although the  $I(0)$  point estimate appears slightly smaller for Canada, the CA-GB difference is not statistically different from zero.) The countries ranked second, third, and fourth by  $I(\alpha)$  are NZ, the USA, and Australia, but differences between the three estimates are not statistically significant. (Differences between NZ on the one hand and Great Britain and Canada on the other hand are significantly different, however.) Rankings by  $J$  are very similar to those by  $I(\alpha)$ .

The bottom row of Figure 3 shows that median-based indices can yield different conclusions about inequality orderings from those based on GL-consistent indices. Although the country ranking by  $ANY(1,1)$  mimics those by  $I(0)$  and  $J$ ,

<sup>10</sup>The narrower confidence bands for South Africa’s estimates partly reflect that country’s distinctly larger WVS sample size (Jenkins, 2019).

the top- and bottom-sensitive indices  $ANY(4,1)$  and  $ANY(1,4)$  provide different patterns. For example, according to these two indices, New Zealand is the second-ranked country by life satisfaction inequality after South Africa. For top-sensitive index  $ANY(4,1)$ , Australia moves down the ranking by comparison with the rankings from the other indices. As well, the precision of the estimates of  $ANY(4,1)$  and  $ANY(1,4)$  seems to be lower than for the other indices: look at the width of the confidence intervals for South Africa and the USA in particular.

## 5. GENERALIZED LORENZ DOMINANCE AND DUAL- $H$ DOMINANCE COMPARED

This section draws attention to differences between the GL dominance criterion and the dual- $H$  dominance criteria of Gravel *et al.* (2020). Gravel *et al.* build their approach on the principle that the inequality of an ordinal variable increases if there is a shift in density mass away from a specific scale level (one person moving to a higher level as well as one moving to a lower level, where the number of levels changed by each person need not be the same). This is the concept of a disequalizing “Hammond transfer,” which may be contrasted with the concept of a disequalizing Pigou-Dalton transfer for a cardinal variable such as income.<sup>11</sup>

Gravel *et al.* (2020) define  $H^+$  and  $H^-$  curves (called  $H$  and  $\bar{H}$  curves in their paper), the ordinates of which are specially defined recursive cumulations of CDFs over the levels of the ordinal variable from  $k = 1, \dots, K-1$ . The authors prove a dual dominance result: distribution  $A$  being more equal than distribution  $B$  according to the Hammond transfer concept is equivalent to (1) the  $H^+$  curve for  $A$  lying nowhere above the  $H^+$  curve for  $B$ , and (2) the  $H^-$  curve for  $A$  lying nowhere above the  $H^-$  curve for  $B$ . They also show that  $F$ -dominance implies  $H^+$  dominance. The dual dominance check can be applied if the distributions have different medians.

Table 2 summarizes the results of  $H^+$ -dominance and  $H^-$ -dominance checks for each of the countries analyzed in the previous section, using the same format as Table 1.<sup>12</sup> Entries above the diagonal refer to  $H^+$ -dominance and entries below the diagonal to  $H^-$ -dominance.

Table 2 shows that there are only four comparisons out of 15 for which there is dual- $H$  dominance: according to this criterion, AU is more equal than ZA; US is more equal than ZA; CA is more equal than NZ; and GB is more equal than NZ. In all four cases, the pairwise ranking by the dual- $H$  criteria coincides with the ranking by the GL dominance criterion (see Table 1). However, there are four additional rankings by the GL dominance criterion for which there is not dual  $H$ -dominance. (All of the eight country-pairs for which there is GL dominance are country-pairs for which there is  $H^+$ -dominance.) For the remaining seven pairwise comparisons, there is neither GL nor dual  $H$ -dominance.

These comparisons demonstrate that Cowell-Flachaire (2017) and related indices such as  $J$  are not Hammond-transfer consistent in general. Another way of seeing this is to compare the social evaluation functions that correspond to the dual- $H$

<sup>11</sup>Allison and Foster’s (2004) median-preserving spread is a specific example of a disequalizing Hammond transfer.

<sup>12</sup>See Appendix Figures 5 and 6 for charts showing all the pairwise comparisons.

TABLE 2  
CROSS-NATIONAL COMPARISONS OF LIFE SATISFACTION DISTRIBUTIONS: SUMMARY OF DUAL  $H$ -DOMINANCE CHECKS

		Country $y$					
		AU	CA	GB	NZ	US	ZA
Country $x$	AU		>	>	>	>	<
	CA	>		<	<	<	<
	GB	>	<		<	<	<
	NZ	>	>	>		<	<
	US	>	<	<	<		<
	ZA	>	<	<	<	>	

Notes: Entries above the diagonal summarize checks for  $H^+$ -dominance: “<”,  $x$ ’s  $H^+$  curve lies nowhere above  $y$ ’s  $H^+$  curve; “>”,  $x$ ’s  $H^+$  curve lies nowhere below  $y$ ’s  $H^+$  curve; “-”, no dominance. Entries below the diagonal summarize checks for  $H^-$ -dominance: “<”,  $x$ ’s  $H^-$  curve lies nowhere above  $y$ ’s  $H^-$  curve; “>”,  $x$ ’s  $H^-$  curve lies nowhere below  $y$ ’s  $H^-$  curve; “-”, no dominance.  $x$  dual- $H$  dominates  $y$  if  $x$   $H^+$ -dominates  $x$  and  $x$   $H^-$ -dominates  $y$ .

dominance criterion with ones associated with the GL dominance criterion. Gravel *et al.*’s (2020) Theorem 5 states that there is an equivalence between (1) distribution  $A$  dual- $H$  dominating distribution  $B$  (i.e.  $A$  being more equal than  $B$ ) and (2) the ranking of  $A$  over  $B$  by an additive social evaluation function defined as follows:<sup>13</sup>

$$(5) \quad \sum_{k=1}^K f_k^A \pi_k \geq \sum_{k=1}^K f_k^B \pi_k$$

for all lists of numbers  $(\pi_1, \dots, \pi_K)$  incorporating social evaluations of the levels that are increasing and “‘strongly concave’ with respect to [the] categories in the sense that the utility gain from moving from a category to a better one is always larger when moving from categories in the bottom of the scale than when moving in the upper part of it” (Gravel *et al.*, 2020, p. 11). Specifically, the lists of numbers  $(\pi_1, \dots, \pi_K)$  must belong to the set  $\mathcal{H}$ , where:<sup>14</sup>

$$(6) \quad \mathcal{H} = \{(\pi_1, \dots, \pi_K) \in \mathbb{R}^K \mid (\pi_i - \pi_g) \geq (\pi_l - \pi_j), \text{ for } 1 \leq g < i \leq j < l \leq K\}.$$

Contrast the social evaluation function used in (5) with, for example, the class  $\mathcal{W}$  defined in Section 2. Recall that indices  $E(\alpha) = 1 - A(\alpha)$ ,  $0 \leq \alpha < 1$ , are members of  $\mathcal{W}$ . So too are:

$$(7) \quad (E(\alpha))^\alpha = \sum_{k=1}^K f_k (F_k)^\alpha, \quad 0 < \alpha < 1, \text{ and}$$

$$\log [E(0)] = \sum_{k=1}^K f_k \log (F_k).$$

<sup>13</sup>Equation (5) is a modification of Gravel *et al.*’s (2020) equation (6) to allow for different population sizes in  $A$  and  $B$ . I also use different notation from them in order to be consistent with usage elsewhere in the current paper.

<sup>14</sup>See Gravel *et al.* (2020, eq. 22).

Equation (7) highlights that the counterparts of the  $\pi_k$  here are functions of  $F_k$ , the measure of status in the peer-inclusive downward-looking status case. These terms are non-decreasing in the scale levels (increasing if there are responses for every level of the ordinal variable), but there is no guarantee that they are “strongly concave” with respect to the categories in the sense specified by (6) and required for consistency with dual- $H$  dominance.

The differences between the GL dominance and dual- $H$  dominance criteria relate to differences in the elementary transformations that characterize them. As explained earlier, the Cowell and Flachaire (2017) approach separates the measurement of inequality into two steps, the first being the definition of status (mapping the ordinal data about levels into a cardinal indicator), and the second being the summary of the dispersion of status across individuals. The elementary transformations associated with the GL dominance criterion refer to changes in individual status, and there is no reference to how these might arise in terms of shifts in the distribution of responses across scale levels. In particular, “distribution  $A$  GL-dominates distribution  $B$ ” is equivalent to “distribution  $A$  can be obtained from  $B$  by a finite sequence of either increases in individual status or status-mean-preserving progressive transfers of status (or both).” In contrast, the elementary transfers in the dual- $H$  dominance approach refer to Hammond transfers and these are defined with reference to changes in the spread of density mass across the scale levels. The constituent  $H^+$ -dominance and  $H$ -dominance criteria also refer to an “increment” and a “decrement,” that is, a shift of one person to a higher level or to a lower level, respectively. What matters for GL dominance is not these changes expressed in relation to *levels* but how they translate into changes in the distribution of *status*.

These points are illustrated using Table 3 which shows seven distributions of a four-level ordinal variable. (The table is modelled on Cowell and Flachaire’s (2017) Tables 2 and 3.) For each distribution, the top half of the table shows the numbers of individuals ( $n_k$ ) at each level  $k = 1, \dots, 4$ , and their status ( $F_k$ ), and the bottom half shows the mean of status and inequality according to three GL-consistent indices.

Consider first a shift from Case 0 to Case 1, achieved by a promotion of 25 individuals from level 3 to level 4, that is, a sequence of increments according to Gravel *et al.* (2020). Case 1  $F$ -dominates Case 0: the shift represents a social improvement according to this criterion characterized with references to scale levels. However, the shift from Case 0 to Case 1 also means that there is a decrease in status of  $\frac{1}{4}$  for each of those who remain at level 3 (and average status falls), and Case 1 is GL-dominated by Case 0. That is, Case 1 represents a socially worse situation than Case 0 according to the GL criterion, and all GL-consistent inequality indices are greater for Case 1. A conflicting assessment by the  $F$ - and GL criteria can also arise when mean status does not change. This is illustrated by a shift from Case 5 to Case 6. Case 5  $F$ -dominates Case 6, but the shift downwards of 20 individuals from level 4 to level 3 is associated with an increase in status for these movers. Case 6 GL-dominates Case 5, and the  $I(0.75)$ ,  $I(0)$ , and  $J$  indices decrease.

Now consider the shift from Case 1 (uniform distribution) to Case 4. This is achieved by five individuals moving up from level 1 to level 2 and five moving

TABLE 3  
DISTRIBUTIONS OF AN ORDINAL VARIABLE

Levels	Case 0		Case 1		Case 2		Case 3		Case 4		Case 5		Case 6	
	$n_k$	$F_k$	$n_k$	$F_k$	$n_k$	$F_k$	$n_k$	$F_k$	$n_k$	$F_k$	$n_k$	$F_k$	$n_k$	$F_k$
1 (N)	25	0.25	25	0.25	0	0	0	0	20	0.20	20	0.20	20	0.20
2 (G)	25	0.50	25	0.50	50	0.5	50	0.50	30	0.50	20	0.40	20	0.40
3 (E)	50	1	25	0.75	50	1	25	0.75	30	0.80	20	0.60	40	0.80
4 (B)	0		25	1	0		25	1	20	1	40	1	20	1
Mean ( $F_k$ )		0.6875		0.6250		0.75		0.6875		0.63		0.64		0.64
$I(0.75)$		1.402		1.661		1.081		1.340		1.643		1.617		1.607
$I(0)$		0.520		0.592		0.347		0.418		0.597		0.607		0.594
$J$		0.484		0.531		0.375		0.422		0.525		0.536		0.520

Notes: The table is modelled on Tables 2 and 3 of Cowell and Flachaire (2017) which refer to Cases 0–3. I have added Cases 4–6. (See also Cases 7 and 8 discussed in the main text.) Categories  $k = 1–4$  are ordered from the worst to the best; N, G, E, and B are Cowell and Flachaire’s labels for the levels. Each individual’s status is measured using the peer-inclusive downward-looking concept; hence the status of an individual reporting level  $k$  is  $F_k$ .  $n_k$  is the number of individuals reporting level  $k$ .

down from level 4 to level 3, and hence corresponding to a sequence of progressive Hammond transfers. There are multiple changes in status, for those who remain in the categories with movers as well as for the movers themselves. This is a situation where there is no GL dominance (nor  $F$ -dominance), and so whether there is a social improvement according to GL-consistent indices depends on the index chosen. Inequality is lower for the uniform distribution (Case 1) than the non-uniform distribution (Case 4) according to  $I(0.75)$ , but the reverse is the case according to  $I(0)$ .

Conflicting conclusions need not arise. A shift from Case 0 to Case 2, achieved by the promotion of 25 individuals from level 1 to level 2, corresponds to an increment in both level space and status space. Case 2  $F$ -dominates Case 0 and also GL-dominates it. The shift from Case 1 (a uniform distribution) to Case 2 represents a case without  $F$ -dominance but is achieved by a sequence of progressive Hammond transfers and hence a social improvement by the dual- $H$  criterion. In addition, Case 1 is GL-dominated by Case 2. The shift from Case 0 to Case 3 is achieved by a promotion of 25 individuals from level 1 to level 2 (increasing their status by  $\frac{1}{4}$ ) and a promotion of 25 individuals from level 3 to level 4 (decreasing the status of the level 3 stayers by  $\frac{1}{4}$ ). This is a status-mean-preserving spread of status and so Case 3 is preferred according to the GL criterion. There is also  $F$ -dominance.

Finally, observe that the same distribution of status may correspond to very different distributions of individuals across the levels. Consider a seventh case (not shown in Table 3) of a four-level distribution in which 50 individuals respond with scale level 1 and 50 respond with level 4; and an eighth case (also not shown), in which 50 individuals respond with scale level 3 and 50 respond with level 4. The distributions of status are identical in both cases and yet, when moving from Case 7 to Case 8, “inequality is unambiguously decreased on the space of the numerical labels” as a referee put it.

Part of the issue here concerns how inequality measures should account for situations in which some categories of the ordinal variable are empty. Cowell and Flachaire argue for what they call the mergers principle: “[i]f two adjacent categories are merged, then this has no effect on the status of any person outside these two categories (2017, p. 293). Application of the principle leads to their peer-inclusive and peer-exclusive status definitions, later combined with their downward- and upward-looking status concepts, and thence underpinning their inequality measures. The mergers principle means that, in Case 7, merging the two empty categories 2 and 3 and then, also merging this combined category with level 4 does not affect our assessment of the status of people in level 1. A counter-argument is that the social welfare gap between level 1 and level 4 is greater than the gap between level 3 and level 4—even if we do not know how much larger because of the ordinal nature of the data—and that this should be taken into account by an inequality measure.

With reference to examples similar to those discussed above, a referee commented that it is inappropriate to refer Cowell-Flachaire indices as “inequality” indices, also suggesting that the measures comprising  $E(\alpha)$  are better described as social evaluation functions rather than as equality indices. The argument is that Cowell-Flachaire inequality indices always react positively not only to

status-mean-preserving progressive transfers of status (which is as one would expect) but also to increments in an individual's status (which is arguably surprising). I believe that Cowell and Flachaire would demur. They argue that they provide “an alternative approach to inequality analysis that is rigorous, has a natural interpretation, and embeds both the ordinal data problem and the well-known cardinal data problem” (Cowell and Flachaire, 2017, p. 290). In doing so, they develop a particular approach to “inequality” which encapsulates the idea that “[w]hat is important in the ordinal data case is the pattern of individual moves towards or away from the reference point” (Cowell and Flachaire, 2017, p. 299). Put differently, Cowell and Flachaire are less concerned about embedding their approach to inequality within a broader framework of distributional comparisons based on social evaluation functions.

In sum, on the one hand and as the old saying goes, “you pays your money and you takes your choice” between the two approaches. On the other hand, and more positively, this paper has shown that it is straightforward to undertake analysis using both approaches. Where they yield different conclusions, researchers can investigate the sources of the differences by more detailed analysis of the distributions.

Currently missing from the analysis toolbox are indices that are guaranteed to be Hammond-transfer-consistent. Filling the gap requires specification of sets of  $\pi_k$  values that satisfy Gravel *et al.*'s (2020) requirement that these numbers are increasing and strongly concave in the manner given by (6).

I now show that the index recently proposed by Apouey *et al.* (2020) only partially fits the bill. Their index is:<sup>15</sup>

$$(8) \quad \text{ASX}(\gamma) = \sum_{k=1}^K f_k \left( \frac{1 - \gamma^{k-1}}{1 - \gamma^{K-1}} \right), \quad 0 < \gamma < 1.$$

ASX( $\gamma$ ) is increasing and concave in the levels of the ordered categorical variable. The  $\gamma$  parameter encapsulates aversion to inequality, with smaller values corresponding to greater aversion. The index reaches its minimum, 0, if all individuals report the lowest level ( $k = 1$ ) and it reaches its maximum, 1, if all individuals report the highest level ( $k = K$ ). ASX( $\gamma$ ) = 0.5 when the distribution is totally polarized, regardless of  $\gamma$ . In the case of a uniform distribution, ASX( $\gamma$ ) =  $[(K - 1) - K\gamma + \gamma^K]/[K(1 - \gamma - \gamma^{K-1} + \gamma^K)]$  which ranges between 0.5, which is the limiting value as  $\gamma \rightarrow 1$ , and  $(K - 1)/K$  which is the limiting value as  $\gamma \rightarrow 0$ . That is, a uniform distribution is always socially preferred to the totally polarized distribution, regardless of the degree of inequality aversion—a result contrasting with the GL dominance ordering (see earlier). However, this social preference represents more than concern about inequality alone; there is an interplay with increasingness. By comparison with the polarized distribution, the higher levels achieved by some in the uniform distribution more than offset the lower levels achieved by others. For

<sup>15</sup>My statement of the ASX() index differs from that in Proposition 3 of Apouey *et al.* (2020, p. 274) because I use the conventional labelling scheme in which  $k = 1$  refers to the worst category and  $k = K$  refers to the best category. Apouey *et al.*'s coding scheme is the reverse.

the same reason, if every individual reports level  $K$ ,  $ASX(\gamma)$  takes a larger value than in a situation in which every individual reports level  $k < K$ . In sum,  $ASX(\gamma)$  is an index of the “social welfare” associated with an ordinal distribution, not an index of inequality alone.

In addition, there is no guarantee that  $ASX(\gamma)$  is always Hammond-transfer-consistent in the sense of Gravel *et al.* (2020). From (6), if  $K = 3$ , the only relevant condition is  $(\pi_2 - \pi_1) \geq (\pi_3 - \pi_2)$ , which is satisfied for all  $\gamma$ . If  $K = 4$ , there are four conditions that need to be satisfied. Three are satisfied but the condition  $(\pi_2 - \pi_1) \geq (\pi_4 - \pi_2)$  holds only if  $\gamma < \gamma^*$  where threshold  $\gamma^* \approx 0.62$ . When  $K = 5$ , there are more conditions and  $(\pi_2 - \pi_1) \geq (\pi_5 - \pi_2)$  holds only if  $\gamma < \gamma^*$  with  $\gamma^* \approx 0.55$ . As  $K$  increases further, Hammond-transfer-consistency requires smaller and smaller  $\gamma^*$ . In other words, consistency is guaranteed only if the degree of inequality aversion is sufficiently great and this ensures the “strong concavity” that Gravel *et al.* (2020) refer to. This result arises because Apouey *et al.*’s Equity Principle is a weaker requirement than Gravel *et al.*’s: it “requires the existence of ... a ‘locally inequality-reduced’ achievement vector that is ranked higher than the initial achievement vector, while Hammond’s equity principle insists on *any* locally inequality-reduced achievement vector being ranked higher than the initial achievement vector” (Apouey *et al.*, 2020, p. 272. Emphasis in original).

## 6. SUMMARY AND CONCLUSIONS

Cowell and Flachaire’s (2017) approach to inequality measurement with ordinal data complements the predominant approach to date that conceptualizes greater inequality as greater spread around the median. This paper builds on Cowell and Flachaire’s work by adding dominance results and a new inequality index. I show that non-intersection of appropriately defined GL curves is equivalent to a unanimous ordering of distributions according to all Cowell-Flachaire (2017) inequality indices and the new index based on GL curve areas. In contrast with  $S$ -dominance, the results presented here can be applied when distributions do not have a common median. Cross-national comparisons based on WVS data show that the GL curve-based results have useful empirical content in the sense of revealing a substantial number of unanimous rankings.

I have also discussed how the GL-dominance criterion discussed here differs from the dual  $H$ -dominance criteria developed by Gravel *et al.* (2020) which relate to the Hammond transfer principle. It is clear that comparisons of distributions of ordinal data expressed in terms of numbers of individuals with different levels of “status” differ fundamentally from comparisons based on distributions expressed in terms of numbers of individuals at different scale levels.

## REFERENCES

- Abdul Naga, R. and T. Yalcin, “Inequality Measurement for Ordered Response Health Data,” *Journal of Health Economics*, 27, 1614–25, 2008.
- Allison, R. A. and J. E. Foster, “Measuring Health Inequality Using Qualitative Data,” *Journal of Health Economics*, 23, 505–24, 2004.



- Apouey, B., “Measuring Health Polarization with Self-Assessed Health Data,” *Health Economics*, 16, 875–94, 2007.
- Apouey, B., J. Silber, and Y. Xu, “On Inequality-Sensitive and Additive Social Achievement Measures Based on Ordinal Data,” *Review of Income and Wealth*, 66, 267–86, 2020.
- Cowell, F. A. and E. Flachaire, “Inequality with Ordinal Data,” *Economica*, 84, 290–321, 2017.
- Gravel, N., B. Magdalou, and P. Moyes, “Ranking Distributions of an Ordinal Attribute,” *Economic Theory*, online ahead of print, 2020. <https://doi.org/10.1007/s00199-019-01241-4>.
- Inglehart, R., C. Haerpfer, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin, and B. Puranen (eds), *World Values Survey: All Rounds - Country-Pooled Datafile Version*, JD Systems Institute, Madrid, 2014. <http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp>.
- Jenkins, S. P., “Better off? Distributional Comparisons for Ordinal Data About Personal Well-Being,” *New Zealand Economic Papers*, online ahead of print, 2019. <https://doi.org/10.1080/00779954.2019.1697729>.
- \_\_\_\_\_, “Comparing Distributions of Ordinal Data,” Discussion Paper 12811, IZA, Bonn, 2020. *The Stata Journal*, forthcoming.
- Jenkins, S. P. and P. J. Lambert, “Three ‘I’s of Poverty Curves, with an Analysis of U.K. Poverty Trends,” *Oxford Economic Papers*, 49, 317–27, 1997.
- Kobus, M., “Polarization Measurement for Ordinal Data,” *Journal of Economic Inequality*, 13, 275–97, 2015.
- Marshall, A. W., I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*. 2nd ed. Springer, New York, 2011.
- Saigo, H., J. Shao, and R. R. Sitter, “A Repeated Half-Sample Bootstrap and Balanced Repeated Replications for Randomly Imputed Data,” *Survey Methodology*, 27, 189–96, 2001.
- Shorrocks, A. F., “Ranking Income Distributions,” *Economica*, 50, 3–17, 1983.
- \_\_\_\_\_, “Revisiting the Sen Poverty Index,” *Econometrica*, 63, 1225–30, 1995.
- Van Kerm, P., “RHSBSAMPLE: Stata Module for Repeated Half-Sample Bootstrap Sampling,” Statistical Software Components S457697, Boston College Department of Economics, revised 17 Nov 2013.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher’s web site:

**Figure A1:** Generalized Lorenz Curves for a Polarized Distribution and Uniform Distributions ( $N = 500$ ;  $K = 3, 4, 5, 10$ )

**Figure A2:** Life Satisfaction Relative Frequency Distributions, by Country

**Figure A3:** Life Satisfaction Distributions: GL Dominance Checks

**Figure A4:** Life Satisfaction Distributions: First-Order Dominance ( $F$ -Dominance) Checks

**Figure A5:** Life Satisfaction Distributions:  $H^+$  Dominance Checks

**Figure A6:** Life Satisfaction Distributions:  $H^-$  Dominance Checks