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

OPHI WORKING PAPER NO. 26b

Rank Robustness of Composite Indices: Dominance and Ambiguity*

James E. Foster†, Mark McGillivray‡ and Suman Seth§

November 2012

Abstract

Many structures in economics – from development indices to expected utility – take the form of a linear composite index, which aggregates linearly across multiple dimensions using a vector of weights. Judgments rendered by composite indices are often given great importance, yet by definition are contingent on an initial vector of weights. A comparison made with one weighting vector could be robust to variations in the weights or, alternatively, it may be reversed at some other plausible vector. This paper presents criteria to discern between these two cases. A general robustness quasi-ordering is defined that requires dominance or unanimous comparisons for a set of weighting vectors, and methods from the

* This paper has benefited from discussions with Sabina Alkire, Bart Lipman, Nora Markova, Farhad Noorbakhsh, Tony Shorrocks, the participants of the 2008 Oxford Poverty and Human Development Initiative (OPHI) "Workshop on Weighting" at the University of Oxford, the 2009 American Economic Association Annual Meeting, the 2010 Social Choice and Welfare Meeting, the 2010 International Association for Research in Income and Wealth Conference, and seminars at Johns Hopkins and George Washington Universities. We thank Alison Kennedy of the UNDP for making the HDI data available for our use and the Institute for International Economic Policy at the Elliott School of George Washington University for research support for Foster. All errors in this paper remain our own. Correspondence to Suman Seth at: suman.seth@qeh.ox.ac.uk.

† Professor of Economics and International Affairs, George Washington University, 1957 E Street, NW Suite 502, Washington, D.C. 20052, +1 202 994 8195, fosterje@gwu.edu, and Research Associate, Oxford Poverty and Human Development Initiative (OPHI), Department of International Development, University of Oxford, 3 Mansfield Road, Oxford, OX1 3TB, UK +44 1865 271915.

‡ Research Professor of International Development, Alfred Deakin Research Institute, Deakin University, Geelong 3217, Australia, +613 5527 8011, mark.mcgillivray@deakin.edu.au, and Research Associate, Oxford Poverty and Human Development Initiative, Department of International Development, University of Oxford, 3 Mansfield Road, Oxford, OX1 3TB, UK +44 1865 271915.

§ Research Officer, Oxford Poverty and Human Development Initiative, Department of International Development, University of Oxford, 3 Mansfield Road, Oxford, OX1 3TB, UK +44 1865 271915, suman.seth@qeh.ox.ac.uk.

This study has been prepared within the OPHI theme on multidimensional methodologies.

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

ISSN 2040-8188 ISBN 978-1-907194-43-6

Bewley model of Knightian uncertainty are invoked to characterize it. We focus on a particular set of weighting vectors suggested by the epsilon-contamination model of ambiguity, which allows the degree of confidence in the initial weighting vector to play a role. We provide a practical vector-valued representation of the resulting epsilon-robustness quasi-ordering and propose a numerical measure to gauge the robustness of a given comparison. An empirical illustration reports on the robustness of Human Development Index country rankings. We extend our methods to certain nonlinear composite indices and explore the links with decision theory, partial comparability in social choice, and the measurement of the freedom of choice.

Keywords: Composite index, robust comparisons, dominance, ambiguity, epsilon-contamination, Knightian uncertainty, Human Development Index

JEL Classifications: I31, O12, O15, C02.

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

Contents

1. Introduction

Linear composite indices, or weighted-sum aggregation methods, are commonplace in economic theory and are frequently invoked in social and economic assessments involving multiple dimensions. Expected utility and price indices have this structure. So too does the Human Development Index (HDI), which aggregates across three equally weighted achievements in health, education and standard of living. Another example is provided by the ubiquitous annual college rankings that convert data on tests scores, educational facilities, and other dimensions into an index to rank universities or departments. The published rankings of certain composite indices generate substantial interest and can impact resource allocation decisions and other economically relevant outcomes.[1]

A ranking generated by a composite index is, however, contingent on the choice of weights: a slight variation in the weights may alter the judgment across a pair of alternatives. And while there are usually plausible methods for selecting initial weights (Decancq and Lugo, 2013), none may be so compelling or precise as to exclude all alternative weights. Given that any comparison made using a composite index is a candidate for reversal, it would seem advisable to explore methods that could evaluate its level of robustness. This is the motivation for the present paper, which provides tractable methods for characterizing and evaluating the robustness of rankings generated by composite indices.

A number of recent papers have considered related issues, particularly in the context of empirical composite indices. Cherchye et al. (2007) and Saisana et al. (2005) consider various “sources of uncertainty” in the construction of a composite index, such as weights, normalization, aggregation techniques, and choice of indicators. Variation in the specification of these factors leads to a distribution of values around the initial composite index value, which they estimate using Monte Carlo methods. McGillivray and Noorbakhsh (2006) evaluate the effect of changing weights by computing the correlations between the original HDI country ranks and those arising from alternative weights. Each of these papers considers a few representative sets of alternative weights in their empirical analysis. Cherchye et al. (2008) consider simultaneous changes in the weights, the normalization, and the aggregation method used by the HDI and derive conditions under which an original comparison is preserved; however, their use of a specialized formula for weighting and normalization limits the applicability of their method.

In what follows, we provide an approach to evaluating robustness with respect to weights that is based on dominance quasi-orderings. This approach is similar to stochastic dominance and related techniques used in income distribution analysis, such as poverty comparisons over a range of poverty lines or inequality comparisons given by the

[1] Country-based composite indices have proliferated of late and include indices of sustainability, corruption, rule of law, economic policy efficacy, institutional performance, happiness, human well-being, transparency, globalization, human freedom, peace and vulnerability. See, for example, Bandura (2008).

Lorenz quasi-ordering.[2] A "robustness quasi-ordering" is defined based on an a priori specification of a set of weighting vectors. A comparison made by the composite index is said to be robust if it is not reversed for any weighting vector in the set. The analysis draws on structures found in the literatures on Knightian uncertainty (Bewley 1986, 2002) and on multiple prior models of ambiguity (Gilboa and Schmeidler 1989). Motivated by a result from Bewley, we characterize this specific form of quasi-ordering from among all possible relations that might be used to check robustness. We further show a straightforward link between the quasi-ordering and the Gilboa-Schmeidler maxmin criterion.

In order to implement the robustness approach in practice, a set of weighting vectors must be selected. One convenient possibility is suggested by the epsilon-contamination model from decision theory: the set of all weighting vectors that can be expressed as a convex combination of the initial weighting vector and a universal set of weighting vectors, where the coefficients on each are respectively $(1 - \varepsilon)$ and ε. The coefficient $(1 - \varepsilon)$ is interpreted as the level of confidence in the initial vector; the "contamination" parameter ε is a direct measure of the size of the set around this vector. Greater confidence in the initial weighting vector is reflected in a lower level of ε-contamination and a smaller set. Our main result characterizes the ε-robustness relation and demonstrates that it has a tractable vector-valued representation: a pair of alternatives is ordered by the ε-robustness relations when the associated vector representations can be ranked using vector dominance.[3]

We then augment our approach by moving beyond zero-one tests to a continuous measure of the robustness of a given comparison. Our measure is constructed using two elements: the difference between the composite index levels of the two alternatives and the maximal "contrary" difference across all possible weighting vectors. We show that the measure has an intuitive interpretation as the maximal level of contamination ε for which the comparison is ε-robust.

An empirical example of our methods is then presented using HDI data.[4] A significant proportion of HDI comparisons across countries are found to be fully robust. An intermediate value of ε is selected and comparisons that satisfy and fail the ε-robustness test are identified. We then construct a table of robustness levels for comparisons among the top ten countries and note that many of these comparisons have limited robustness. The example shows how our techniques can be readily used for interpreting the rankings of composite indices and appropriately discounting comparisons that have minimal robustness.

The rest of the paper is structured as follows. Section 2 provides the notation and definitions that are used in the rest of the paper. The robustness quasi-orderings are

[2] See Atkinson (1970, 1987), Bawa (1975), Foster and Shorrocks (1988a,b), and Sen and Foster (1997) for related discussions.

[3] See Foster (1993, 2011) and Sen and Foster (1997, pp. 205–207) for discussions of vector-valued representations.

[4] For simplicity, we use the traditional HDI, which is a linear composite index of normalized achievements. The current HDI is nonlinear but, as discussed below, is also amenable to our robustness analysis. See UNDP (2010) for details of these indices.

defined and characterized in Section 3. In Section 4, we construct the robustness measure and demonstrate its relationship to ε-robustness. Section 5 applies this measure to the Human Development Index. Section 6 extends the analysis to certain nonlinear composite indices and explores the links with other theoretical constructs. Section 7 provides some concluding remarks and outlines future research directions.

2. Notation and Definitions

Let $X \subseteq R^D$ denote the nonempty set of alternatives to be ranked, where each alternative is represented as a vector $x \in X$ of achievements in $D \geq 2$ dimensions. For $a, b \in R^D$, the expression $a \geq b$ indicates that $a_d \geq b_d$ for $d = 1, \ldots, D$; this is the *vector dominance* relation. If $a \geq b$ with $a \neq b$, this situation is denoted by $a > b$; while $a >> b$ indicates that $a_d > b_d$ for $d = 1, \ldots, D$. Let $\Delta = \{w \in R^D: w \geq 0$ and $w_1 + \ldots + w_D = 1\}$ be the *simplex of weighting vectors*. A *composite index* C: $X \times \Delta \rightarrow R$ combines the dimensional achievements in $x \in X$ using a weighting vector $w \in \Delta$ to obtain an aggregate achievement level $C(x;w) = w \cdot x = w_1x_1 + \ldots + w_Dx_D$. In what follows, it is assumed that an *initial weighting vector* $w^0 \in \Delta$ has already been chosen; this fixes the specific composite index C_0: $X \rightarrow R$ defined as $C_0(x) = C(x;w^0)$ for all $x \in X$. The associated ordering of achievement vectors will be denoted by $\boldsymbol{C_0}$, so that $x \; \boldsymbol{C_0} \; y$ holds if and only if $C_0(x) \geq C_0(y)$. For every $d \in \{1,\ldots,D\}$, we denote by e_d the usual basis vector, whose d^{th} element is equal to one and the remaining entries are zero, e.g., $e_2 = (0,1,0,\ldots,0)$.

3. Robust Comparisons

We construct a general criterion to determine when a given comparison $x \; \boldsymbol{C_0} \; y$ is robust. Let $W \subseteq \Delta$ be a nonempty set of weighting vectors containing initial vector w^0. Define the *robustness relation* $\boldsymbol{R_W}$ on X by

$$x \; \boldsymbol{R_W} \; y \text{ if and only if } C(x;w) \geq C(y;w) \text{ for all } w \in W$$

for any pair $x,y \in X$. Whenever $x \; \boldsymbol{R_W} \; y$ is true, the inequality $C(x;w) \geq C(y;w)$ holds at $w = w^0$ and is robust to the choice of $w \in W$. When $x \; \boldsymbol{C_0} \; y$ holds but $x \; \boldsymbol{R_W} \; y$ does not, this means that $C(x;w) < C(y;w)$ for some $w \in W$, and hence the initial ranking is *not* robust. Let $\boldsymbol{I_W}$ and $\boldsymbol{P_W}$ denote the symmetric and asymmetric parts of $\boldsymbol{R_W}$. Notice that $\boldsymbol{R_W}$ is typically not a complete ordering but is reflexive and transitive.

Relation $\boldsymbol{R_W}$ is closely linked with other dominance criteria, including Sen's (1970a,b) approach to partial comparability in social choice and Bewley's (1986, 2002) multiple prior model of Knightian uncertainty. Bewley's presentation, in particular, suggests a natural characterization of $\boldsymbol{R_W}$ among all binary relations $\boldsymbol{R}$ on X. Consider the following properties, each of which is satisfied by $\boldsymbol{R_W}$.

Quasi-ordering (Q): $\boldsymbol{R}$ is reflexive and transitive.
Monotonicity (M): (i) If $x > y$ then $x \; \boldsymbol{R} \; y$; (ii) if $x >> y$ then $y \; \boldsymbol{R} \; x$ cannot hold.
Independence (I): Let $x, y, z, y', z' \in X$ where $y' = \alpha x + (1-\alpha)y$ and $z' = \alpha x + (1-\alpha)z$ for $0 < \alpha < 1$. Then $y \; \boldsymbol{R} \; z$ if and only if $y' \; \boldsymbol{R} \; z'$.

Continuity (*C*): The sets $\{x \in X \mid x\, \boldsymbol{R}\, z\}$ and $\{x \in X \mid z\, \boldsymbol{R}\, x\}$ are closed for all $z \in X$.

Axiom *Q* allows $\boldsymbol{R}$ to be incomplete. Axiom *M* ensures that $\boldsymbol{R}$ follows vector dominance when it applies and rules out the converse ranking when vector dominance is strict. Axiom *I* is a standard independence axiom, which requires the ranking between *y* and *z* to be consistent with the ranking of *y*' and *z*' obtained from *y* and *z*, respectively, by a convex combination with another vector *x*. Finally, Axiom *C* ensures that the upper and lower contour sets of $\boldsymbol{R}$ contain all their limit points. We have the following characterization, the proof of which is given in the appendix.

Theorem 1: Suppose that *X* is closed, convex, and has a nonempty interior. Then a binary relation $\boldsymbol{R}$ on *X* satisfies axioms *Q, M, I,* and *C* if and only if there exist a non-empty, closed, and convex set $W \subseteq \Delta$ such that $\boldsymbol{R} = \boldsymbol{R}_W$.

Proof: In the Appendix.

Consequently, any relation satisfying the four axioms can be generated by pair-wise comparisons of the composite index over some fixed set *W* of weighting vectors.

The relation $\boldsymbol{R}_W$ has an interesting interpretation in terms of the maxmin criterion of Gilboa and Schmeidler (1989) for multiple priors. Suppose that $x\, \boldsymbol{R}_W\, y$ for some nonempty, closed set $W \subseteq \Delta$. By linearity of the composite index, this can be expressed as $C(x - y;w) \geq 0$ for all $w \in W$, or as $\min_{w\in W} C(x - y;w) \geq 0$. The Gilboa-Schmeidler evaluation function $G_W(z) = \min_{w\in W} C(z;w)$ represents the maxmin criterion, which ranks a pair of options *x* and *y* by comparing $G_W(x)$ and $G_W(y)$, or the respective minimum values of the composite indicator on the set *W*. Our robustness ranking $x\, \boldsymbol{R}_W\, y$ is obtained by applying G_W to the *net vector* $(x - y)$ and checking whether the resulting value is nonnegative. Indeed, $x\, \boldsymbol{R}_W\, y$ if and only if $G_W(x - y) \geq 0$.[5]

With Theorem 1, the selection of a robustness criterion reduces to the choice of an appropriate set *W* of weighting vectors. But which *W* should be used? As we argue below, the answer depends in part on the confidence one places in the initial weighting vector w^0. If one has confidence that w^0 is the most appropriate weighting vector, then this would be reflected in the specification of a smaller set *W* containing w^0. The limiting case of $W = \{w^0\}$ indicates utmost confidence in w^0 and hence no additional robustness test is required: $x\, \boldsymbol{C}_0\, y$ is equivalent to $x\, \boldsymbol{R}_W\, y$. On the other hand, a larger *W* would suggest less confidence in w^0, a more demanding robustness test $\boldsymbol{R}_W$, and correspondingly fewer robust comparisons. Clearly $\boldsymbol{R}_{W'}$ is a sub-relation of $\boldsymbol{R}_W$ whenever $W \subseteq W'$. We now further investigate $\boldsymbol{R}_W$ for some natural specifications of *W*.

[5] The maxmin criterion applies when $G_W(x) - G_W(y) \geq 0$, while our robustness criterion holds when $G_W(x\text{-}y) \geq 0$. The maxmin criterion generates a complete relation, but requires comparisons of $C(x,w)$ with $C(y,w')$ for some $w \neq w'$, which is not so easily interpreted in the present context. See Ryan (2009) for related discussions of Bewley (2002) and Gilboa and Schmeidler (1989).

Full Robustness

We begin with the limiting case where $W = \Delta$, the set of all possible weighting vectors, and denote the associated robustness relation by $\boldsymbol{R_1}$. When $x\ \boldsymbol{R_1}\ y$ holds, we say the comparison $x\ \boldsymbol{C_0}\ y$ is *fully robust* since it is never reversed at *any* configuration of weights. Requiring unanimity over Δ is quite demanding and consequently $\boldsymbol{R_1}$ is the least complete among all such relations; however, when it *does* hold, the associated ranking of achievement vectors is undoubtedly robust.

Consider the vertices of Δ, given by $v_d = e_d$ for $d = 1,\ldots,D$, where e_d places full weight on the single achievement d. Clearly $C(x;v_d) = x_d$, which suggests a link between the $\boldsymbol{R_1}$ and vector dominance. We have the following characterization.

Theorem 2: Let $x, y \in X$. Then $x\ \boldsymbol{R_1}\ y$ if and only if $x \geq y$.

Proof: Suppose that $x\ \boldsymbol{C_0}\ y$ is true. If $x \geq y$ holds, then clearly $C(x;w) = w{\cdot}x \geq w{\cdot}y = C(y;w)$ for all $w \in \Delta$, and thus $x\ \boldsymbol{R_1}\ y$. Conversely, if $x\ \boldsymbol{R_1}\ y$ holds, then setting $w = v_d$ in $C(x;w) \geq C(y;w)$ yields $x_d \geq y_d$ for all d, and hence $x \geq y$. ■

In order to check whether a given ranking $x\ \boldsymbol{C_0}\ y$ is fully robust, one need only verify that the achievement levels in x are at least as high as the respective levels in y. It follows that completely robust indifference $x\ \boldsymbol{I_1}\ y$ is equivalent to $x = y$, while the strict robustness ranking $x\ \boldsymbol{P_1}\ y$ is equivalent to $x > y$.

One interesting implication of Theorem 2 is that judgments made by $\boldsymbol{R_1}$ are "meaningful" even when variables are ordinal and no basis of comparison between them has been fixed.[6] Suppose that each variable x_d in x is independently altered by its own monotonically increasing transformation $f_d(x_d)$ and let $x' = (f_1(x_1),\ldots, f_D(x_D))$ be the resulting transformed achievement vector. It is clear that $x > y$ if and only if $x' > y'$, and consequently, by Theorem 2 we have $x\ \boldsymbol{R_1}\ y$ if and only if $x'\ \boldsymbol{R_1}\ y'$. In other words, if $\boldsymbol{R_1}$ holds for any given cardinalization of the ordinal variables, it holds for all cardinalizations. Note that while $\boldsymbol{C_0}$ on its own is *not* meaningful in this context (as $y'\ \boldsymbol{C_0}\ x'$ is entirely consistent with $x\ \boldsymbol{C_0}\ y$), the fully robust relation $\boldsymbol{R_1}$ is preserved and hence is appropriate for use with ordinal, non-commensurate variables.

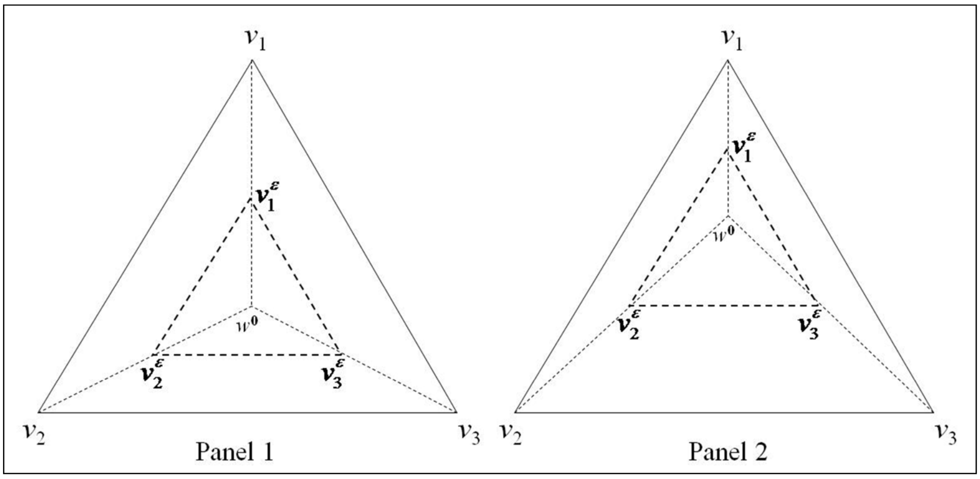
Epsilon Robustness

Now consider $\Delta_\varepsilon \subseteq \Delta$ defined by $\Delta_\varepsilon = (1-\varepsilon)\{w^0\} + \varepsilon\Delta$ for $0 \leq \varepsilon \leq 1$, which is made up of vectors of the form $(1-\varepsilon)w^0 + \varepsilon w$, where $w \in \Delta$. Parameter value $\varepsilon = 0$ yields $\Delta_0 = \{w^0\}$ and hence the "no robustness" case, while $\varepsilon = 1$ yields $\Delta_1 = \Delta$ or full robustness. Each Δ_ε with $0 < \varepsilon < 1$ is a rescaled version of Δ in which the relative position of w^0 is unchanged. Figure ***1*** provides examples of Δ_ε for the case of $D = 3$ and $\varepsilon = ¼$, where Panel 1 has $w^0 = (1/3,1/3,1/3)$ and Panel 2 has $w^0 = (3/5,1/5,1/5)$. As noted in the figure, ε is a measure of the relative size of Δ_ε. For a given w^0 the sets are nested in such a way that $\Delta_\varepsilon \supset \Delta_{\varepsilon'}$ whenever $\varepsilon > \varepsilon'$.

[6] For a formal discussion of "meaningful statements" see Roberts (1979).

The set Δ_ε of weighting vectors can be motivated using the well-known epsilon contamination model of multiple priors commonly applied in statistics and decision theory.[7] In that context, w^0 corresponds to an initial subjective distribution and Δ_ε contains all probability distributions that are convex combinations of w^0 and a distribution from the set Δ of all objectively possible distributions, where $(1-\varepsilon)$ represents the decision maker's level of confidence in w^0 and ε is the extent of the "perturbation" from w^0. The Gilboa-Schmeidler evaluation function G_W then reduces to a form invoked by Ellsberg (1961), namely $G_\varepsilon(z) = (1-\varepsilon)C(z;w^0) + \varepsilon \min_{w\in\Delta} C(z;w)$ using our notation.[8]

Figure 1: Examples of Δ_ε



Substituting Δ_ε into the definitions of $\boldsymbol{R}_W$ yields the *ε-robustness relation* $\boldsymbol{R}_\varepsilon$ for ε in [0,1]. Since the sets Δ_ε are nested for a given w^0, it follows that $x\ \boldsymbol{R}_\varepsilon\ y$ implies $x\ \boldsymbol{R}_{\varepsilon'}\ y$ whenever $\varepsilon > \varepsilon'$. The rankings require $C(x;w) \geq C(y;w)$ for all w in Δ_ε and hence at each of its vertices $v_d^\varepsilon = (1-\varepsilon)w^0 + \varepsilon v_d$. Define $x^\varepsilon = (x_1^\varepsilon,\ldots,x_D^\varepsilon)$ where $x_d^\varepsilon = C(x; v_D^\varepsilon) = v_D^\varepsilon \cdot x$, and let y^ε be the analogous vector derived from y. The following result characterizes $\boldsymbol{R}_\varepsilon$.

Theorem 3: Let $x, y \in X$. Then $x\ \boldsymbol{R}_\varepsilon\ y$ if and only if $x^\varepsilon \geq y^\varepsilon$.

Proof: We need only verify that $x^\varepsilon \geq y^\varepsilon$ implies $x\ \boldsymbol{R}_\varepsilon\ y$. Pick any $w \in \Delta_\varepsilon$, and note that since Δ_ε is the convex hull of its vertices, w can be expressed as a convex combination of $v_1^\varepsilon,\ldots,v_D^\varepsilon$, say $w = \alpha_1 v_1^\varepsilon+\ldots+\alpha_D v_D^\varepsilon$ where $\alpha_1+\ldots+\alpha_D = 1$ and $\alpha_d \geq 0$ for $d = 1,\ldots,D$. But then $C(x;w) = w\cdot x = \alpha_1 v_1^\varepsilon \cdot x+\ldots+\alpha_D v_D^\varepsilon \cdot x = \alpha_1 x_1^\varepsilon+\ldots+\alpha_D x_D^\varepsilon$, and similarly $C(y;w) = \alpha_1 y_1^\varepsilon+\ldots+\alpha_D y_D^\varepsilon$; therefore $x^\varepsilon \geq y^\varepsilon$ implies $C(x;w) \geq C(y;w)$. Since w was an arbitrary element of Δ_ε, it follows that $x\ \boldsymbol{R}_\varepsilon\ y$. ■

[7] See for example, Carlier, Dana, and Shahidi (2003); Chateauneuf, Eichberger, and Grant (2006); Nishimura and Ozaki (2006); Carlier and Dana (2008); Asano (2008); and especially Kopylov (2009).

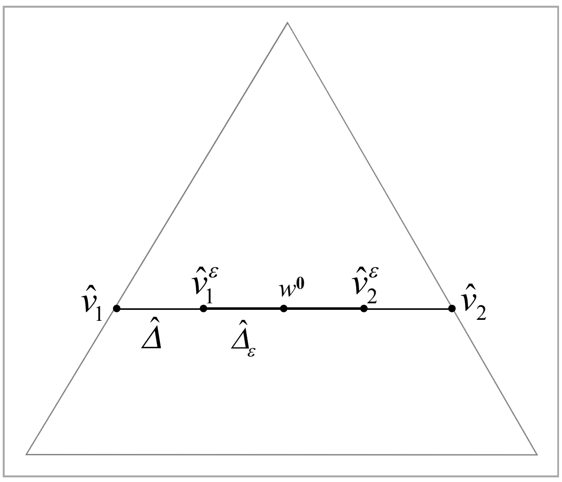
[8] See Kopolov (2009).

Theorem 3 shows that to evaluate whether a given comparison x $\boldsymbol{C_0}$ y is ε-robust, one need only compare the associated ε-robustness vectors x^ε and y^ε. If each component of x^ε is at least as large as the respective component of y^ε, then the comparison is ε-robust; if any component is larger for y^ε than x^ε, then the comparison is not. Checking whether the x^ε vector dominates y^ε is equivalent to requiring the inequality $C(x;w) \geq C(y;w)$ to hold for each vertex $w = v_D^\varepsilon$ of the set Δ_ε. Note further that x^ε is a convex combination of the vectors $(C_0(x),\ldots,C_0(x))$ and x, namely, $x^\varepsilon = (1-\varepsilon)(C_0(x),\ldots,C_0(x)) + \varepsilon x$, so that when $\varepsilon = 1$ we obtain the condition $x \geq y$ in Theorem 2, while for $\varepsilon = 0$, the condition reduces to a simple comparison of $C_0(x)$ and $C_0(y)$.

Restricted Robustness

The full robustness criterion in Theorem 2 and the epsilon-robustness criterion in Theorem 3 both use the entire simplex Δ as the universal set of weighting vectors. In certain cases, however, it may be natural to restrict the universal set to be a nonempty subset $\hat{\Delta}$ of Δ. For example, let $D = 3$ and consider the restricted set $\hat{\Delta} = \{w \in \Delta: w_1 \leq 1/3\}$. In the inequality constraint, the first weight is no more than one-third and hence the sum of the other two weights is no less than two-thirds. Alternatively, the restricted set $\hat{\Delta}' = \{w \in \Delta: w_1 = 1/3\}$ adds a second inequality $w_1 \geq 1/3$ and requires the first weight to be one-third and leaves the second and third weights free to vary, except that they must sum to two-thirds.[9] Both of these examples yields a nonempty universal set that is delimited by a finite number of linear inequality constraints and hence is a *convex polytope*; such a set can be equivalently seen as the convex hull of a finite number $n \geq 1$ of vertices $\{\hat{v}_1,\ldots,\hat{v}_n\}$. See Figure 2.

Figure 2: An Example of $\hat{\Delta}_\varepsilon$



[9] This example is analogous to an example in Ellsberg (1961): a ball is to be drawn from an urn having one-third red balls and two-thirds yellow or black balls, with the frequency of red balls being known but the frequency of yellow (or black) balls being uncertain.

When $w^0 \in \hat{\Delta}$, we define $\hat{\Delta}_\varepsilon = (1-\varepsilon)\{w^0\} + \varepsilon\hat{\Delta}$ for $0 \le \varepsilon \le 1$ and consider the generalized ε-robustness relation $\hat{\boldsymbol{R}}_\varepsilon$ obtained by setting $W = \hat{\Delta}_\varepsilon$ in the definition of $\boldsymbol{R_W}$. The i^{th} vertex of $\hat{\Delta}_\varepsilon$ is constructed as $\hat{v}_i^\varepsilon = (1-\varepsilon)w^0 + \varepsilon\hat{v}_i$ for $i = 1,\ldots,n$. Define $\hat{x}^\varepsilon = (\hat{x}_1^\varepsilon, \ldots, \hat{x}_n^\varepsilon)$ where $\hat{x}_i^\varepsilon = C(x; \hat{v}_i^\varepsilon)$ or the value of the composite index of x evaluated at weighting vector $\hat{v}_i^\varepsilon$ for all i, and let $\hat{y}^\varepsilon$ be the analogous n-dimensional vector derived from y. We have the following result, whose proof is entirely analogous to the proof of Theorem 3.

Theorem 4: Suppose that $\hat{\Delta}$ is a convex polytope and let $x, y \in X$. Then $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ if and only if $\hat{x}^\varepsilon \ge \hat{y}^\varepsilon$.

The logic of the previous two results still applies in the case where additional linear inequality constraints have been imposed and the resulting universal set is an arbitrary convex polytope. Verifying $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ amounts to checking whether the inequality $C(x;w) \ge C(y;w)$ holds for a finite number of weighting vectors, namely $w = \hat{v}_i^\varepsilon$; the robustness quasi-ordering is represented by the associated n-dimensional vector ordered by vector dominance.

As an example, consider the restricted universal set $\hat{\Delta} = \{w \in \Delta: w_1 = 1/3\}$ and the initial weighting vector $w^0 = (1/3, 1/3, 1/3)$ depicted in Figure 2. The vertices of $\hat{\Delta}$ are given by $\hat{v}_1 = (1/3, 0, 2/3)$ and $\hat{v}_2 = (1/3, 2/3, 0)$, so that the weight on dimension 1 remains fixed at 1/3, but the weights on the other two dimensions are allowed to vary. Full robustness $x\ \hat{\boldsymbol{R}}_1\ y$ in this case requires $n = 2$ comparisons, one for each vertex, which can be succinctly expressed as $\hat{x} \ge \hat{y}$ where

$$\hat{x} = (x_1+2x_3, x_1+2x_2)/3 \text{ and } \hat{y} = (y_1+2y_3, y_1+2y_2)/3.$$

Now for $\varepsilon = 1/2$, the relevant vertices of $\hat{\Delta}_\varepsilon$ are given by $\hat{v}_1^\varepsilon = (1/3, 1/6, 1/2)$ and $\hat{v}_2^\varepsilon = (1/3, 1/2, 1/6)$ and the conditions for $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ are less stringent, namely $\hat{x}^\varepsilon \ge \hat{y}^\varepsilon$, where

$$\hat{x}^\varepsilon = (2x_1+x_2+3x_3, 2x_1+3x_2+x_3)/6 \text{ and } \hat{y}^\varepsilon = (2y_1+y_2+3y_3, 2y_1+3y_2+y_3)/6$$

are the associated ε-robustness vectors.

4. Measuring Robustness

Our method of evaluating the robustness of comparison $x\ \boldsymbol{C_0}\ y$ fixes a set $\hat{\Delta}_\varepsilon$ of weighting vectors and confirms that the ranking at w^0 is not reversed at any other $w \in \hat{\Delta}_\varepsilon$, in which case the associated ε-robustness condition $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ applies. Theorem 4 provides simple conditions for checking when $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ holds. This section augments this approach by formulating a measure that identifies a robustness level $r \in [0,1]$ for any comparison $x\ \boldsymbol{C_0}\ y$, where levels $r = 0$ and $r = 1$ correspond to the absence of robustness and to full robustness, respectively, and each $r \in (0,1)$ is some level between the two extremes.

We construct r using two statistics – one that might be expected to move in line with robustness and another that is likely to work against it. The first of these is

$$A = C_0(x) - C_0(y)$$

or the difference between the composite value of x and the composite value of y at the initial weighting vector w^0. Intuitively, A is an indicator of the strength of the dominance of x over y at the initial weighting vector. It is nonnegative whenever $x\ \boldsymbol{C_0}\ y$. The second is

$$B = \begin{cases} \max_{w\in\hat{\Delta}}[C(y;w) - C(x;w)] & \text{for } \max_{w\in\hat{\Delta}}[C(y;w) - C(x;w)] > 0 \\ 0 & \text{otherwise} \end{cases}$$

or the maximal "contrary" difference between the composite values of y and x if this quantity is positive, and zero otherwise. Note that when the original comparison is fully robust over $\hat{\Delta}$, then $C(y;w) - C(x;w) \leq 0$ for all $w \in \hat{\Delta}$ and there is no contrary difference. This is the case where $B = 0$. On the other hand, when the comparison is not fully robust, then $C(y;w) - C(x;w) > 0$ for some $w \in \hat{\Delta}$, and hence $B = \max_{w\in\hat{\Delta}}[C(y;w) - C(x;w)] > 0$. B is then the worst-case estimate of how far the original difference at w^0 could be reversed at some other weighting vector.

We propose the following as our measure of robustness:

$$r = \begin{cases} \frac{A}{A+B} & \text{for } B > 0 \\ 1 & \text{for } B = 0 \end{cases}$$

Notice that when the initial comparison $x\ C_0\ y$ is fully robust over $\hat{\Delta}$, we have $B = 0$ and hence $r = 1$, as desired. Alternatively, when the initial comparison is not fully robust and $B > 0$, the measure r is strictly increasing in the magnitude of the initial comparison A and strictly decreasing in the magnitude of the contrary worst-case evaluation B. These characteristics accord well with an intuitive understanding of how A and B might affect robustness.[10]

Practical applications of r may be hampered by the fact that it requires a maximization problem to be solved, namely $\max_{w\in\hat{\Delta}}[C(y;w) - C(x;w)]$. However, by the linearity of $C(y;w) - C(x;w) = (y - x)\cdot w$ in w, the problem has a solution at some vertex $\hat{v}_i$ where the difference $C(y;w) - C(x;w)$ takes on the simplified expression $\hat{v}_i\cdot(y - x) = \hat{y}_i$ –

[10] The limiting case where A falls to 0 is of particular interest. If it occurs in the presence of a fixed $B > 0$, the measure of robustness r tends to 0, which is its value for the case $A = 0$ and $B > 0$; if it occurs when $B = 0$, the measure r tends to 1, which is its value for the case $A = 0$ and $B = 0$.

$\hat{x}_i$. Consequently, B depends on the maximum coordinate of the vector $\hat{y} - \hat{x}$, which in the case of $\hat{\Delta} = \Delta$, becomes the maximum coordinate of $y - x$. The measure r is readily derived using this equivalent expression.

Now what is the relationship between the robustness measure r and the relation $\hat{\boldsymbol{R}}_\varepsilon$ developed in the previous section? The following theorem provides the answer.

Theorem 5: Let r be the robustness level associated with comparison $x\ C_0\ y$ for $x, y \in X$. Then $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ holds if and only if $0 \leq \varepsilon \leq r$.

Proof: Select any comparison $x\ C_0\ y$ and let $r \geq 0$ be its robustness level. We consider first the case of $0 < r < 1$. If $0 \leq \varepsilon \leq r$, then by the definition of r, we have $\varepsilon \leq A/(A + B)$ and hence $\varepsilon B \leq (1 - \varepsilon)A$. Pick any $i = 1, \ldots, n$. From the definitions of A and B, we see that $\varepsilon(\hat{y}_i - \hat{x}_i) \leq (1 - \varepsilon)(w^0 \cdot x - w^0 \cdot y)$ and hence $\varepsilon \hat{v}_i \cdot y + (1 - \varepsilon)w^0 \cdot y \leq \varepsilon \hat{v}_i \cdot x + (1 - \varepsilon)w^0 \cdot x$. Consequently, $\hat{v}_i^\varepsilon \cdot y \leq \hat{v}_i^\varepsilon \cdot x$, and since this is true for all i, it follows that $\hat{x}^\varepsilon \geq \hat{y}^\varepsilon$ and thus $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ by Theorem 4. Conversely, suppose that $r < \varepsilon \leq 1$. Then $(1 - \varepsilon)A < \varepsilon B$ so that $(1 - \varepsilon)(w^0 \cdot x - w^0 \cdot y) < \varepsilon(\hat{y}_i - \hat{x}_i)$ for some i, and hence $\hat{v}_i^\varepsilon \cdot y > \hat{v}_i^\varepsilon \cdot x$ or $\hat{y}_i^\varepsilon > \hat{x}_i^\varepsilon$ for this same i. It follows, then, that $\hat{x}^\varepsilon \geq \hat{y}^\varepsilon$ cannot hold, and neither can $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ by Theorem 4.

Consider next the case of $r = 0$. If $0 \leq \varepsilon \leq r$, it follows that $\varepsilon = 0$. Clearly $x\ \boldsymbol{C}_0\ y$ immediately ensures that $x\ \hat{\boldsymbol{R}}_0\ y$. Conversely, suppose that $r < \varepsilon \leq 1$. By the definition of r, we have $A = 0$ and $B > 0$, and hence the vector $\hat{x} - \hat{y}$ has both positive and negative entries, as must $\hat{x}^\varepsilon - \hat{y}^\varepsilon$. Consequently $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ cannot hold. Finally, consider the case of $r = 1$. Clearly $B = 0$ and hence the comparison $x\ C_0\ y$ is fully robust over $\hat{\Delta}$. Thus $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ for all $0 \leq \varepsilon \leq 1$, which completes the proof. ■

As ε rises, the robustness criterion becomes more demanding and $\hat{\boldsymbol{R}}_\varepsilon$ less complete. Theorem 5 identifies r as the *maximal* ε for which $x\ \hat{\boldsymbol{R}}_\varepsilon\ y$ holds, and hence $\hat{\Delta}_r$ is the largest set $\hat{\Delta}_\varepsilon$ over which the original comparison is not reversed. Alternatively, r is the largest ε for which the Gilboa-Schmeidler (or Ellsberg) evaluation function of the net achievement vector $(x–y)$ is nonnegative; i.e., $\hat{G}_\varepsilon(x - y) = (1-\varepsilon)C(x-y;w^0) + \varepsilon\min_{w\in\hat{\Delta}} C(x-y;w) \geq 0$. Note that reducing the size of the universal set $\hat{\Delta}$ reduces the size of $\hat{\Delta}_\varepsilon$ and allows additional comparisons to be made by $\hat{\boldsymbol{R}}_\varepsilon$. This implies that for a given comparison $x\ \boldsymbol{C}_0\ y$ the measure of robustness r will rise or at least will not fall. Exactly how much r is affected depends on the extent to which the maximum contrary valuation B falls as the set of weights shrinks.

5. Empirical Illustration

To illustrate our methods, we now apply them to the well-known Human Development Index (HDI) in its traditional version as an equally weighted linear composite index over the three dimensions of health (H), education (E), and standard of living (SL).[11] For simplicity, we take Δ to be the universal set of weights, which has the initial weighting vector $w^0 = (1/3,1/3,1/3)$ at its center.

Table ***1*** lists the top ten countries by HDI value. While the associated $\boldsymbol{C}_0$ ranking is a complete ordering, it says nothing about the robustness of a given judgment to changes in weights.

Table 1: The Top 10 HDI Countries in 2004

Rank	Country	HDI
1	Norway	0.965
2	Iceland	0.960
3	Australia	0.957
4	Ireland	0.956
5	Sweden	0.951
6	Canada	0.950
7	Japan	0.949
8	United States	0.948
9	Switzerland	0.947
10	Netherlands	0.947

Table 2 evaluates the robustness of three specific comparisons. The first set of columns restates the HDI information from Table 1. The next three columns provide the dimensional achievements x_1, x_2, and x_3 needed to ascertain whether the full robustness $\boldsymbol{R}_1$ is obtained. It is evident that the achievement vector of Australia dominates the achievement vector of Sweden and hence by Theorem 2 this comparison is fully robust. However, for the comparison between Iceland and the US there is a reversal in the standard of living dimension, while the Ireland/Canada comparison has a reversal in health, and so neither of these comparisons is fully robust. Observe that the HDI margin between Australia and Sweden (0.006) is identical to the margin for Ireland and Canada,

[11] As noted in Section 6 below, our methods are equally applicable to the new HDI introduced in 2010, which is a geometric mean; an empirical example can be obtained from the authors upon request. Data for 2004 were obtained directly from the UNDP and allow greater precision than the rounded off published figures in the 2006 *Human Development Report*. In particular, Switzerland has a slightly higher HDI level than the Netherlands, although the rounded off levels are identical. For more extensive empirical applications of our method, see Permanyer (2011) and Foster, McGillivray and Seth (2012).

and yet the robustness characteristics of the two comparisons are quite different: the Australia/Sweden comparison is fully robust; whereas the Ireland/Canada comparison is not. Also notice that the HDI margin between Iceland and USA is twice as large (0.012) and yet it too is not fully robust.

Table 2: Three HDI Comparisons

Rank	Country	HDI	H x_1	E x_2	SL x_3	H $x_1^{0.25}$	E $x_2^{0.25}$	SL $x_3^{0.25}$
3	Australia	0.957	0.925	0.993	0.954	0.949	0.966	0.956
5	Sweden	0.951	0.922	0.982	0.949	0.944	0.959	0.951
2	Iceland	0.960	0.931	0.981	0.968	0.953	0.965	0.962
8	USA	0.948	0.875	0.971	0.999	0.930	0.954	0.961
4	Ireland	0.956	0.882	0.990	0.995	0.937	0.964	0.966
6	Canada	0.950	0.919	0.970	0.959	0.942	0.955	0.952

The final three columns of Table 2 report the entries of the associated ε-robustness vectors for $\varepsilon = 0.25$ to check the ε-robustness of the comparisons. Recall that each element of the ε-robustness vector is the composite index evaluated at the respective vertex of Δ_ε. A quick evaluation in terms of vector dominance reveals that both the Australia/Sweden and the Iceland/USA comparisons are ε-robust by Theorem 3, but the absence of vector dominance in the Ireland/Canada comparison implies that ε-robustness does not hold for this ranking when $\varepsilon = 0.25$. In other words, there are weighting vectors in Δ_ε at which Canada's composite index is larger than that of Ireland.

The analysis in the previous paragraph shows whether the comparisons are robust for a particular robustness level. We can also calculate the levels of robustness for each of these comparisons. The Australia/Sweden comparison is fully robust, with $A = 0.006$ and $B = 0$, and hence $r = 100\%$. The Iceland/USA comparison has $A = 0.012$ and $B = 0.031$, and hence $r = 27.9\%$. In contrast, the Ireland/Canada ranking has $A = 0.006$ and $B = 0.037$, and therefore $r = 14.3\%$.[12] Table 3 presents the level of robustness of pair-wise comparisons for the top ten countries. For every cell below the diagonal, the "column country" of the cell has a higher ranking according to $\boldsymbol{C_0}$ than the "row country". The number in the cell indicates the level of robustness of the associated comparison, expressed in percentage terms. Out of the 45 pair-wise comparisons involving the top ten countries, four are fully robust, while 20 of them have robustness levels of 25% or higher. However, for the entire dataset of 177 countries for the same year, we find that 69.7% of the pair-wise comparisons are fully robust while about 92% have robustness levels of 25% or higher.[13]

[12] The robustness measures were computed from the more precise underlying data and hence may differ slightly from computations using numbers reported in Table 2.

[13] We are ignoring the reflexive comparisons as any comparison of a country with itself is trivially fully robust. For an extended discussion of the prevalence of robust comparisons in a given dataset, and how it relates to positive association among dimensions, see Foster, McGillivray, and Seth (2012).

Table 3: Robustness Levels (Percentage)

Country		NOR	ISL	AUS	IRL	SWE	CAN	JPN	USA	CHE
	Rank	1	2	3	4	5	6	7	8	9
Norway (NOR)	1									
Iceland (ISL)	2	19.5								
Australia (AUS)	3	34.8	18.9							
Ireland (IRL)	4	85.7	14.0	3.6						
Sweden (SWE)	5	53.1	94.0	100	10.5					
Canada (CAN)	6	61.3	100	60.5	14.3	13.8				
Japan (JPN)	7	27.6	34.0	23.0	9.1	7.1	2.5			
USA (USA)	8	76.6	27.9	16.6	67.3	5.3	3.0	0.7		
Switzerland (CHE)	9	49.3	100	41.1	15.6	16.7	19.5	6.3	1.9	
Netherlands (NLD)	10	100	67.5	56.6	47.1	24.7	12.8	3.7	7.1	0.6

6. Extensions, Links and Further Applications

The analytical results in this paper can be applied to certain composite indices that are not linear but can be transformed into linear composite indices. For example, expected utility analysis typically makes use of nonlinear utility transformations $x_d = u_d(s_d)$ where s_d is an underlying source variable. Another example is provided by the class of alternative human development indices proposed by Chakravarty (2003), for which $x_d = s_d^{\alpha}/D$, with for $0 < \alpha < 1$. In these cases, robustness can be analyzed for the transformed variables x_d using the methods presented above.[14] If a composite index can be expressed as a monotonic transformation f of a linear composite index, so that $F(x;w) = f(C(x;w))$, then the robustness analysis for F can be conducted in terms of the underlying C. The current Human Development Index (UNDP 2010), which takes the form of a geometric mean, provides an example having both forms of transformation, namely, $f(C(x;w^0))$ where $x_d = \ln s_d$ and $f(t) = e^t$. The robustness analyses for the new measure can be conducted on $C(x;w^0)$ using the standard methods. Thus, our approach is not limited to linear composite indices but also applicable to a wide range of nonlinear composite indices such as the new Human Development Index (UNDP 2010), Human Poverty Index, and the Inequality adjusted HDI.[15]

Our robustness analysis is closely related to constructs found in other areas of economic theory, and these links might well suggest directions for future investigation. We now describe several of these links – to decisions under uncertainty, to partial comparability in social choice, and to opportunity freedom.

[14] Note that in the case of full robustness, the criterion can be equivalently expressed in terms of the underlying variables s_d. See the discussion after Theorem 2 in Section 3.

[15] The Inequality-adjusted HDI can be expressed as a composite index due to its property of path independence. See Foster, Lopez-Calva and Szekely (2005) and Alkire and Foster (2010). Other examples of applicable nonlinear composite indices can be found in Hicks (1997).

First, our theoretical framework mirrors the structure used in decision making under uncertainty, where x is interpreted as the state-specific utility levels associated with an act, w is a probability vector, and $C(x;w)$ is the expected utility. Decision making under risk can then be seen as the special case where w^0 is given and the complete ranking $\boldsymbol{C_0}$ is available. Our analysis evaluates the robustness of choice under risk of uncertainty, where the latter takes the form of ε-contamination relative to a universal set of probability vectors. A finding of $x\ \hat{\boldsymbol{R}}_1\ y$ (and hence $r = 1$) ensures that the selection of x over y is completely robust, whereas a very low level of r may lead one to re-evaluate the decision. For example, consider the choice between $x = (2,3,3)$, and $y = (2,4,4)$, given $w^0 = (1,0,0)$ and let the universal set be $\hat{\Delta} = \Delta$. Clearly, $x\ \boldsymbol{C_0}\ y$ holds, which is consistent with x being chosen over y and, indeed, this weak ranking is confirmed by the maxmin criterion of Gilboa and Schmeidler (1989). The robustness level of $x\ \boldsymbol{C_0}\ y$, though, is $r = 0$; if there were any lack of confidence in w^0, and some other w were used, the expected utility ranking would be strictly reversed and the choice of x over y would be seen as wrong. In contrast, the robustness level of the converse comparison $y\ \boldsymbol{C_0}\ x$ is $r = 1$, indicating the robust priority of y over x in this case.[16]

Second, our robustness approach and indeed the recent constructs from uncertainty and ambiguity analysis have certain parallels in the theory of social choice. Especially relevant is Sen's (1970a, 1970b) analysis of partial comparability in utilitarian evaluations of welfare. Options x and y are now vectors of "basic" utility levels across the population for two social states, while the utilitarian welfare level is just the sum of entries. The comparability of utilities of different persons is modeled in this context as the set of allowable transformation vectors, each containing coefficients that independently rescale each person's utility up or down before summing. Partial comparability is modeled using a cone of coefficient vectors that is larger than a single ray (complete comparability) and smaller than the strictly positive orthant (non-comparability), with a larger set indicating less comparability. Sen explores whether a given social choice comparison is robust to the changes allowed under partial comparability and defines the associated dominance quasi-ordering on the set of social states. Without loss of generality we can normalize the utility coefficients to sum to one, in which case his structure maps perfectly into our framework, with w being the normalized vector of coefficients, $C(x;w)$ being the utilitarian welfare given w, and $\boldsymbol{R_W}$ being his dominance quasi-ordering given the set W of allowable (normalized) coefficient vectors. Sen even provides an indicator of partial comparability between two social states that is similar in spirit, but different in orientation, to our measure of robustness.

Third, the analysis is also relevant to the evaluation of the freedom (or flexibility) inherent in opportunity sets or menus of alternatives.[17] A decision maker has two

[16] Note that w^0 is the probability vector in $W = \Delta$ at which the minimum level of expected utility is achieved for both x and y, and this minimum level is the same for both. Hence, $G_W(x) = G_W(y)$, and the maxmin criterion is unable to discern between dominated x and dominant y. The criterion leads the decision maker to evaluate x and y through the lens of w^0, which suppresses information that is arguably relevant in an uncertain environment. Put differently, if there were the least bit of doubt that w^0 would be the correct initial weight, then y should be the unambiguous choice.

[17] See for example Kreps (1979), Foster (1993, 2011), Arrow (1995), and Sen (2002).

decisions: in period one, an opportunity set must be selected; in period two an alternative must be selected from the chosen opportunity set. The decision maker is endowed with a set of potential utility functions for evaluating alternatives in the second period. However, in the first period, the decision maker is uncertain which utility function will be relevant in the second period and must try to select an opportunity set that offers enough freedom or flexibility to provide reasonable alternatives at the second stage for every possible utility function. If the decision maker has a single utility function, and hence knows period two utility in period one, then the sets can be ranked using an indirect utility approach in which the value of a set is the utility of its best element. If the decision maker has *plural* utilities and uncertainty about which of them might arise, then each opportunity set maps to a *vector* of indirect utilities – one for each of the possible utility functions. Kreps (1979) and Arrow (1995) posit the existence of a probability vector w^0 that indicates the relative likelihood of each utility function and leads to a complete ranking $\boldsymbol{C_0}$ of opportunity sets via the expected utility function $C(x;w^0)$. The counting approach of Pattanaik and Xu (1990) can be viewed as a special case where the decision maker has equal probabilities over a set of utilities that represent all logically possible preferences and thus has very little information with which to differentiate among alternatives. At the other extreme is the case of full robustness where any probability vector in Δ could arise and the relation $\boldsymbol{R}_1$ becomes the effective freedom ranking of Foster (1993, 2011), which in turn can be represented as vector dominance over the utility vectors. Our robustness analysis could provide a method for bridging these two extremes and identifying new intermediate measures of freedom.

7. Concluding Remarks

Composite indices are commonly used in economics and other disciplines to order and rank alternatives. The information provided by these rankings is often of great importance and influence, yet by definition is contingent on an initial vector of weights. This paper has presented methods for evaluating the extent to which comparisons are robust to changing weights. We began with a general robustness quasi-ordering requiring dominance or unanimous comparisons for a set of weighting vectors and applied a result from the Bewley (1986) model of Knightian uncertainty to characterize it. We then focused on particular sets of weighting vectors suggested by the epsilon-contamination model of ambiguity reflecting one's "degree of confidence" in the initial weighting vector, as in Ellsberg (1961). Practical vector-valued representations of the resulting epsilon-robustness quasi-orderings were provided, and a numerical measure to gauge the robustness of a given comparison was proposed and characterized in terms of the quasi-orderings.

An illustration of the applicability of the paper's methods was also presented, using the well-known Human Development Index in its traditional linear version. A significant proportion — nearly 70% — of HDI comparisons for the year 2004 across countries were found to be fully robust. It was observed that the robustness of comparisons could be very different even when the HDI differences were essentially the same. The paper then provided an extension of its methods to certain nonlinear composite indices (including the current Human Development Index (UNDP 2010), which takes the form of a geometric mean) and explored links with decision theory, partial comparability in social choice, and the measurement of the freedom of choice.

One lesson to be drawn from this paper is that unless greater care and sophistication are used in the reporting of composite indices, their ability to inform will be marginalized. It is commonplace in reporting the results of econometric analysis to provide a range of diagnostic and other statistics, including *t*-ratios, so the reader can make judgments about the veracity of these results. No such analysis presently accompanies the release of composite index values and rankings, despite the ambiguities associated with the design of these indices. This paper has provided several methods and one statistic (the robustness measure) that could be reported alongside composite index values and ranks, thereby improving the information content and strengthening the interpretation of results.

We end by discussing two potential extensions of our approach – to multidimensional poverty and to price indices. A recent paper by Alkire and Foster (2011) presents a new approach to evaluating multidimensional poverty that has been taken up by various countries and international organizations. In this approach each dimensional deprivation has a value, and persons are considered poor if the sum of their deprivation values exceeds a poverty cutoff. Once the poor are identified and the nonpoor data have been censored, the "adjusted headcount ratio" measure of poverty can be defined as a linear composite index of the censored data with (normalized) deprivation values taking on the role of initial weights. A natural question is whether poverty comparisons are robust to changes in these weights. Our methods are immediately applicable to the aggregation step of the process – after the poor have been identified – to evaluate a limited form of robustness. However, in the poverty setting, weights can also affect who is seen as being poor. It would be interesting to try to formulate an analogous notion of robustness in the identification of the poor and then to combine the two stages to obtain an idea of the overall robustness of the measure.[18]

A second potential extension is to price indices and the related constructs of poverty lines and purchasing power parity (PPP) indices.[19] In an interesting example, Digby (2001) used CPI data for the British Virgin Islands for 1996 and 1997 to show how estimates of inflation are sensitive to CPI weights. In particular, he showed that if the weight on the food item were varied from 0% to 100% and the rest of the weight were equally distributed across the other six items, then inflation estimates would range from 2.9% to 3.5%. This exercise is similar to our restricted robustness setting, where commodities are grouped into food and non-food items, and then the weights on the two groups are varied. The statements being subjected to robustness tests are not, however,

[18] To get a sense of the robustness of country rankings for the Multidimensional Poverty Index (MPI) – an implementation of the adjusted headcount ratio to 109 countries – Alkire and Santos (2010) and Alkire, Roche, Santos and Seth (2011) recalculate the index for three alternative weighting vectors and compare country rankings to those of the initial weighting vector.

[19] A recent ILO manual on consumer prices explained the importance of weights in this context: "If all prices moved in the same way, weights would not matter. On the other hand, the greater the variation in price behaviour between products, the greater the role of weights in measuring aggregate price change" (ILO 2004).

only concerned with rankings (e.g, prices have risen from 1996 to 1997) but can depend on the cardinal values of the composite index (e.g., inflation is above 2.9%). It would be interesting to explore robustness criteria that could apply to these more stringent statements.

References

Alkire, S. and Foster, J.E. (2010) Designing the Inequality-Adjusted Human Development Index (IHDI). *OPHI Working Paper 37*, Oxford Poverty and Human Development Initiative, University of Oxford.

Alkire, S. and Foster, J.E. (2011) Counting and Multidimensional Poverty Measurement. *Journal of Public Economics*, 95:476–487.

Alkire, S., Roche, J.M., Santos, M.E. and Seth, S. (2011) Multidimensional Poverty Index 2011: Brief Methodological Note. Oxford Poverty and Human Development Initiative, University of Oxford. <http://www.ophi.org.uk/policy/multidimensional-poverty-index/mpi-data-methodology/>

Alkire, S. and M. E. Santos (2010) Acute Multidimensional Poverty: A New Index for Developing Countries. *OPHI Working Paper No. 38*, Oxford Poverty and Human Development Initiative, University of Oxford.

Arrow, K. J. (1995) A Note on Freedom and Flexibility. In K. Basu, P. K. Pattanaik, & K. Suzumura (Eds.), *Choice, Welfare, and Development. A Festschrift in Honor of Amartya K. Sen* (pp. 7–16). Oxford: Clarendon Press.

Asano, T. (2008) Portfolio Inertia and Epsilon-contaminations. *Theory and Decision*, 68: 341–365.

Atkinson, A. B. (1970) On the Measurement of Inequality. *Journal of Economic Theory*, 2: 244–263.

Atkinson, A. B. (1987) On the Measurement of Poverty. *Econometrica*, 2: 244–263.

Bandura R. (2008) A Survey of Composite Indices Measuring Country Performance: 2008 Update. *UNDP/ODS Working Paper*. New York: Office of Development Studies, United Nations Development Programme.

Bawa, V. (1975) Optimal Rules for Ordering Uncertain Prospects. *Journal of Financial Economics*, 2: 95–121.

Bewley, T. F. (1986) Knightian Decision Theory: Part I. *Discussion Paper No. 807*, Cowles Foundation.

Bewley, T. F. (2002) Knightian Decision Theory. Part I. *Decisions in Economics and Finance*, 25: 79–110.

Carlier, G. and Dana, R.A. (2008) Two-persons Efficient Risk-sharing and Equilibria for Concave Law-invariant Utilities. *Economic Theory*, 36: 189–223.

Carlier, G., Dana, R.A. and Shahidi, N. (2003) Efficient Insurance Contracts Under Epsilon-contaminated Utilities. *The GENEVA Papers on Risk and Insurance Theory*, 28: 59–71.

Chakravarty, S. R. (2003) A Generalized Human Development Index. *Review of Development Economics*, 7: 99–114.

Chateauneuf, A., Eichberger, J. and Grant, S. (2006) Choice under Uncertainty with the Best and Worst in Mind. *Journal of Economic Theory*, 137: 538–567.

Cherchye L., Moesen, W., Rogge, N., Puyenbroeck, T.V., Saisana, M., Saltelli, A., Liska, R. and Tarantola, S. (2007) Creating Composite Indicators with DEA and Robustness Analysis: the Case of the Technology Achievement Index. *Journal of the Operational Research Society*, 59: 239–251.

Cherchye L., Ooghe, E., and Puyenbroeck, T.V. (2008) Robust Human Development Rankings. *Journal of Economic Inequality*, 6: 287–321.

Decancq K. and M. A. Lugo (2013), "Weights in Multidimensional Indices of Well-Being: An Overview", Econometric Reviews, 32, 7-34.

Digby P. W. (2001) A Note on Weights for Consumer Price Indices: Comments on Articles by Chris Scott and Rameshwar Srivastava. *Inter-Stat*, 23: 105–111.

Ellsberg, D. (1961) Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics*, 75: 643–669.

Foster J. E. (1993) Notes on Effective Freedom. *Working Paper No. 34*, Oxford Poverty and Human Development Initiative, Oxford University.

Foster J. E. (2011) Freedom, Opportunity and Wellbeing. In *Handbook of Social Choice and Welfare*, K. Arrow, A. Sen, and K. Suzumura (Eds.), North Holland, 687–728.

Foster J. E., McGillivray, M. and Seth, S. (2012) Composite Indices: Rank Robustness, Statistical Association and Redundancy. *Econometric Reviews*.

Foster, J. E., Lopez-Calva, L.F. and Szekely, M. (2005) Measuring the Distribution of Human Development: Methodology and an Application to Mexico. *Journal of Human Development*, 6: 5–29.

Foster J. E. and Shorrocks, A.F. (1988a) Poverty Orderings and Welfare Dominance. *Social Choice and Welfare*, 5: 179–198.

Foster J. E. and Shorrocks, A.F. (1988b) Poverty Orderings. *Econometrica*, 56: 173–177.

Gilboa, I. and Schmeidler, D. (1989) Maxmin Expected Utility with Non-unique Prior. *Journal of Mathematical Economics*, 18: 141–153.

Hicks, D. A. (1997). The inequality adjusted human development index: A constructive proposal. *World Development*, 25: 1283–1298.

ILO, IMF, OECD, UNECE, Eurostat and The World Bank (2004) *Consumer Price Index Manual: Theory and Practice*. Geneva: International Labour Office.

Kopylov, I. (2009) Choice Deferral and Ambiguity Aversion. *Theoretical Economics*, 4: 199–225.

Kreps, D. M. (1979) A Representation Theorem for 'Preference for Flexibility'. *Econometrica*, 47: 565–577.

McGillivray, M and Noorbakhsh, F. (2006) Composite Indices of Human Well-being: Past, Present and Future, M. McGillivray (Ed.) Human Well-Being: Concept and Measurement. Basingstoke: Palgrave Macmillan, 113-134.

Nardo M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. (2005) *Handbook on Constructing Composite Indicators: Methodology and User's Guide*, Paris: Joint Research Centre (JRC) of the European Commission and OECD.

Nishimura, K. G. and Ozaki, H. (2006). An Axiomatic Approach to ε-contamination. *Economic Theory*, 27: 333–340.

Pattanaik, P. K. and Xu, Y. (1990) On Ranking Opportunity Sets in terms of Freedom of Choice. *Récherches Economiques de Louvain*, 56: 383–390.

Permanyer I. (2011) Assessing the Robustness of Composite Indices Rankings. *Review of Income and Wealth,* 57: 306–326.

Roberts F. S. (1979) Measurement Theory: with Applications to Decision Making, Utility, and the Social Sciences. *Encyclopedia of Mathematics and its Applications*, Vol. 7, Reading, MA: Addison-Wesley.

Ryan, M. J., (2009) Generalizations of SEU: a Geometric Tour of some Non-standard Models. *Oxford Economic Papers*, 61: 327–354.

Saisana, M., Saltelli, A. and Tarantola, S. (2005) Uncertainty and Sensitivity Analysis as Tools for the Quality Assessment of Composite Indicators. *Journal of the Royal Statistical Society*, Ser. A 168: 1–17.

Sen, A. K. (1970a) *Collective Choice and Social Welfare*, San Francisco: Holden Day.

Sen, A. K. (1970b) Interpersonal Aggregation and Partial Comparability. *Econometrica*, 38: 393–409.

Sen, A. K. (2002) *Rationality and Freedom*, Cambridge, MA: The Belknap Press of Harvard University Press.

Sen A. and Foster J. E. (1997) On Economic Inequality: After a Quarter Century. Annex to the enlarged edition of *On Economic Inequality* by Amartya Sen, Oxford: Clarendon Press.

UNDP (1990) *Human Development Report 1990*, New York: Oxford University Press.

_____ (2000) *Human Development Report 2000*, New York: Oxford University Press.

_____ (2001) *Human Development Report 2001*, New York: Oxford University Press.

_____ (2006) *Human Development Report 2006,* New York: Oxford University Press.

_____ (2010) *Human Development Report 2010*, New York: Oxford University Press.

Appendix

Proof of Theorem 1. Let $\boldsymbol{R}$ be a binary relation on a set X that is closed, convex, and has some z in its interior.

If $\boldsymbol{R} = \boldsymbol{R}_W$ for some non-empty, closed, and convex $W \subseteq \Delta$, then it is immediate that $\boldsymbol{R}_W$ satisfies Q, M, I, and C.

Conversely, suppose that $\boldsymbol{R}$ satisfies Q, M, I, and C. Define $U = \{x \in X: x\,\boldsymbol{R}\,z\}$ as the upper contour set of $\boldsymbol{R}$ at z. We know that $z \in U$ by Q and U is closed by C. Moreover, we can show that U is convex. Pick any $x, y \in U$. Let $x' = \alpha x + (1 - \alpha)y$ for some α with $0 < \alpha < 1$. Then, where $z' = \alpha z + (1 - \alpha)y$, we have $x', z' \in X$ and by axiom I it follows that $x'\,\boldsymbol{R}\,z'$. Moreover, by a second application of I, it follows from $y\,\boldsymbol{R}\,z$ that $z'\,\boldsymbol{R}\,z$. Therefore, by Q we have $x'\,\boldsymbol{R}\,z$ and so U is convex.

Since, z is in the interior of X, there exists $\lambda > 0$ such that $N_\lambda = \{x \in R^D: \|x - z\| \leq \lambda\} \subseteq X$. Define $U_\lambda = U \cap N_\lambda$ and note that it is compact, convex, and contains z, so that the set $K_\lambda = \{z\} - U_\lambda$ is compact, convex, and contains 0. Let $K = \text{Cone } K_\lambda$ be the cone generated by K_λ. It is immediate that K is closed, compact, and contains 0. We can state that K has the property that for $x, y \in X$ we have $x\,\boldsymbol{R}\,y$ if and only if $y - x \in K$. To see this, let $x, y \in X$ and select $\alpha > 0$ small enough that z' satisfying $z = \alpha y + (1 - \alpha)\, z'$ lies in N_λ and $x' = \alpha x + (1 - \alpha)\, z'$ is also in N_λ. Clearly, $z - x' = \alpha(y - x)$ for $\alpha > 0$. So if $x\,\boldsymbol{R}\,y$, we know that $x'\,\boldsymbol{R}\,z$ by I, and hence $z - x' \in K$ which implies $y - x \in K$. On the other hand, if $y - x \in K$, then since $z - x' \in K$, we have $x'\,\boldsymbol{R}\,z$ so that $x\,\boldsymbol{R}\,y$ by I, establishing the result.

Now let $P = \{p \in R^D: p{\cdot}k \leq 0 \text{ for all } k \in K\}$ be the polar cone of K, so that by standard results on polar cones, P is closed and convex. It is clear that $P \subseteq R^D_+$, since by monotonicity, we have $-v_d \in K$ and so $p{\cdot}(-v_d) \leq 0$ and $p_d \geq 0$, where v_d is the D-dimensional usual basis vector for co-ordinate d. In addition, we can show that P contains at least one element $p \neq 0$. Indeed, it is clear from M that K contains no $k >> 0$ (otherwise, we would have $x << z$ with $x\,\boldsymbol{R}\,z$). Then, $K \cap R^D_{++} = \varnothing$ and since both sets are convex, we can apply the Minkowski separation theorem to find $p^0 \neq 0$ in P. Let $W = \Delta \cap P$, so that cone $W = P$. Clearly, K is the polar cone of both P and W, hence, $K = \{t \in R^D: w{\cdot}t \leq 0 \text{ for all } w \in W\}$.

We now show that $\boldsymbol{R} = \boldsymbol{R}_W$. If $x\,\boldsymbol{R}\,y$, then $y - x \in K$ and so $w(y - x) \leq 0$ for all $w \in W$, hence $x\,\boldsymbol{R}_W\,y$. Conversely, if $x\,\boldsymbol{R}_W\,y$, then by definition we have $w(y - x) \leq 0$ for all $w \in W$, hence $x - y \in K$ or $x\,\boldsymbol{R}\,y$. ■