Robust human development rankings[1]

Laurens Cherchye[2], Erwin Ooghe[3], and Tom Van Puyenbroeck[4]

October, 2005

[1]We are grateful to Isabel Günther, Serge Kolm and other participants in the "First Meeting of the Society for the Study of Economic Inequality (ECINEQ)" (Palma de Mallorca; July, 2005) for useful comments and suggestions. This paper is an offshoot of the KEI-project (contract n° 502529) that is part of priority 8 of the policy orientated research under the European Commission's Sixth Framework Programme (see http://kei.publicstatistics.net/).

[2]University of Leuven Campus Kortrijk and Center for Economic Studies, E. Sabbelaan 53, B-8500 Kortrijk, Belgium, e-mail: laurens.cherchye@kulak.be. Postdoctoral Fellow of the Fund for Scientific Research - Flanders (FWO-Vlaanderen).

[3]University of Leuven Center for Economic Studies, Naamsestraat 69, B-3000 Leuven, Belgium, e-mail: erwin.ooghe@econ.kuleuven.be. Postdoctoral Fellow of the Fund for Scientific Research - Flanders (FWO-Vlaanderen).

[4]European University College Brussels (Ehsal) and Center for Economic Studies, Stormstraat 2, B-1000 Brussels, Belgium, e-mail: tom.vanpuyenbroeck@ehsal.be.

Abstract

The United Nations' Human Development Index (HDI) takes several dimensions —income, school enrolment and literacy rate, longevity— and combines them into a single figure that measures the degree of development of a given country. However, there is disagreement about (i) how to normalize the scores for the different criteria to make them comparable and (ii) how to aggregate the (normalized) scores over the different criteria. At the risk of stressing the obvious, changes in normalization and/or aggregation will affect the country rankings. First, we focus on robust rankings, i.e., rankings which hold for a wide set of normalization and/or aggregation procedures. Second, we show that all proposed ranking procedures can be implemented via linear programming techniques. Third, we illustrate how our methodology can prove useful in assessing the robustness of the human development country ranking/classification (produced annually by the United Nations) in a descriptive and statistical way.

JEL-codes: C61, D63, I31, O10.

Keywords: Human Development Index, Lorenz Dominance, Linear programming, Robustness Analysis.

"It is better to be vaguely right, than precisely wrong" (Wildon Carr)

1 Introduction

Gauging human development is a much needed, yet very challenging and intricate enterprise. The epitome of this endeavour is the series of Human Development Reports (HDR) commissioned by the United Nations Development Programme (UNDP) since 1990. Ever since their inception, these reports have stressed that human development can not be assessed by looking at income alone, e.g., by analyzing the evolution of a country's GDP per capita over time. Upholding the view that 'people are the real wealth of nations' (HDR 2004, p. 217), this measure can indeed hardly be regarded as an end in itself. Heavily influenced by the work of Amartya Sen, the HDR's have instead continuously endorsed and promoted the idea that people's 'capabilities' are the prime objectives of human development. Specifically, this means that human development ought to be regarded as a process of expanding peoples' choices, i.e., of taking away obstacles to the things a person could do and be in his life. Thus income continues to be an important yardstick —since it is indicative of a person's "command over resources" (Anand and Sen, 2000)— but has been complemented by numerous other indicators. In fact, the typical HDR is composed of dozens of tables and figures, reporting on hundreds of indicators that are all related in some way or another to the overarching concept of human development.

Unfortunately, a plethora of statistics is ill-suited as a tool for public communication. Grabbing the attention of various stakeholders and triggering public debate are major reasons for the fact that, since the beginning, the reports have also provided a summary 'Human Development Index' (HDI). This index singles out (a) leading a long and healthy life, (b) being knowledgeable, and (c) enjoying a decent standard of living as three key aspects of human development. The index surely succeeds in attracting public interest to the HDR's, as witnessed by the media coverage that accompanies the release of a new HDR. In point of fact, the media headlines quite often relate to the country ranking that results from the HDI.

Evidently, by its very nature a summary index is bound to conceal important aspects of a complex phenomenon such as human development. The HDI's authors are actually the first to stress that the index is by no means a complete measure of a country's level of development.[1] In fact, the au-

[1] As a matter of fact, this statement can be found as such on hdr.undp.org/statistics/faq/.

thors have always been very meticulous in spelling out their methodology, in making the underlying data accessible, and in fostering the transparency of the HDI's construction by including interactive DIY-versions on their official website (hdr.undp.org/statistics/indices). As a result, we will not give a detailed explanation of its construction here. It will suffice to point here at the three steps that lead up to the HDI. First, each of the three aforementioned aspects is associated with particular quantitative indicators (viz. life expectancy at birth; adult literacy rate and the school enrolment ratio; GDP per capita as measured in PPP US$). In their original form these individual indicators are incommensurable. Therefore, in a second step, the original data are normalized, i.e., re-expressed on a scale from 0 to 1. Once this has been accomplished, the normalized variables are aggregated via an (equally) weighted sum. Even this briefest of accounts illustrates that the HDI indeed errs on the side of simplicity: methodological transparency and substantive precision are not exactly synonymous.[2]

Now, precisely in view of the fact that the HDI is scrupulously presented as only a part of the overall picture, attacking its method of construction seems like rather an easy criticism. It may lead one to wonder why so many researchers make every effort to analyze, criticize and try to improve that construction methodology. We think at least two reasons can be given. First, the HDR's authors have proven to be receptive to such criticisms; its construction has been slightly modified over the years and these modifications have at least partially been inspired by critical annotations from the scientific community (see, e.g., Fukada-Parr, 2001; Fukada-Parr and Shiva Kumar, 2003). Second, and more importantly, precisely because the HDI has such a high media profile, it seems to have acquired a similarly eye-catching 'test case'-position in the broader debate on whether and how a multi-dimensional phenomenon, which builds on different and often incommensurable primary components, ought to be summarized in a composite indicator.

Booysen (2002, p. 131) summarizes the debate on composite indicators by noting that "not one single element of composite indexing is above criticism." And indeed, the HDI is a well-known case in point. Some authors have proposed the inclusion of other indicators in the index (e.g., Sagar and Najam, 1998; Dar, 2004), while others have suggested indicators with a spe-

[2]Evidently, we fully agree with the view that full accessibility to both the basic data and the specific construction methodology are good qualities: they can only enhance the credibility of a composite indicator. For similar remarks, see the OECD-European Commission's Handbook on Constructing Composite Indicators (Nardo et al., 2005).

cific focus (such as poverty or gender; e.g., Klasen and Bardhan, 1999). As regards the pure measurement aspects, several writers have forwarded different alternatives to the specific, 'arbitrary' normalization procedures used by UNDP (e.g., Luchters and Menkhoff, 2000; Panigrahi and Sivramkrishna, 2002; Mazumdar, 2003; Chakravarty, 2003), whereas others have primarily focused on the method of aggregating/weighting the different components (e.g., Noorbakhsh, 1988; Sagar and Najam, 1998; Mahlberg and Obersteiner, 2001; Dowrick et al., 2003; Lind, 2004; Despotis, 2005; Chatterjee, 2005). Again, many of the aforementioned authors particularly emphasize the effects of changes in this procedure on the resulting country ranking.

In this paper, we will also be concerned with the pure normalization/aggregation aspects and their effects on eventual country rankings. We will hence take the original quantitative indicators as given, and will sidestep the controversy over whether these are effectively suited to capture all fundamental, universally shared and measurable dimensions of human development. Although we will propose a different normalization procedure than the one that is actually used, our primary purpose is not to add yet another remeasured HDI to an already impressive list of variants. Instead, we present on the following pages a methodology that intends precisely to capture the fact that there is wide disagreement about the normalization and aggregation of the HDI's components. In a specific sense, such disagreement can even be taken as a defining characteristic of this general type of composite indicator: one does agree that there is a complex multi-faceted issue such as, in this case, human development, and one does acknowledge that many sub-indicators, related to this overarching concept, can be discerned. But the exact (inter)relationship between these sub-indicators and that 'vague' global phenomenon is too difficult to grasp exactly; hence the real possibility of dispute, even among experts.

We will start in section 2 by taking each of these two well-known origins of dispute in turn. We start with the aggregation issue. In section 2.1, we propose the aggregation of (normalized) scores using the class of increasing and S-concave indices. Thus, our multi-dimensional ranking criterion is equivalent to the idea of generalized Lorenz dominance. Proceeding as such, we seek to give foundation to the concern that it may not solely be the average, but also the distribution of the (normalized) scores in the different dimensions that must be taken into account when comparing countries.[3] In

[3] Maassoumi and Nickelsburg (1988, p. 328) were among the first to point out that "without this diversity in distributions there would be no need for a multi-dimensional measure." We stress that the diversity pertains to the 'inequality' among indicators and

section 2.2, we adddress the normalization issue. Specifically, we propose a way to summarize conflicts of opinion that pertain to the question whether and how, e.g., a relatively high level of life expectancy outweighs a comparatively low literacy rate. Essentially, we allow for 'zones of agreement' (resp. 'zones of disagreement') in which there is consensus (resp. no consensus) among experts that a score in one dimension outweighs a score in another dimension. Next, in section 3, we look at alternative methods of progressing from individual opinions to an aggregate decision rule. In particular, we suggest (i) a strong dominance criterion: country x is at least as good as y according to all (mutually agreed) normalization schemes and all increasing S-concave indices, and (ii) a weak dominance criterion: country x is at least as good as y if there exists at least one normalization scheme within the mutually agreed set of possible schemes that results in x outperforming y (on the basis of all increasing S-concave indices), while no other normalization scheme leads to the opposite result. In particular, the core of our paper is concerned with identifying workable rules for the implementation of both decision criteria. Specifically, we demonstrate how these criteria translate into linear programming problems. By the very nature of these two criteria, it is evident that in several cases one may end up with incomplete rankings. But what is gained is precisely the increased robustness of the comparisons that still can be made. That is, even if the scientific community establishes only a limited consensus on how to gauge human development, the tools to be presented can still be useful to answer the question whether one country indeed 'outperforms' another one (or itself, if one uses time series data). In section 4 we look at the results of such exercises, using the 2002 UN dataset, and illustrate the usefulness of our methodology for investigating (statistically as well as descriptively) the robustness of the United Nations' human development ranking. A final section 5 concludes.

---

not about inter-individual inequality, e.g., of income. The same issue has been taken up recently by Chatterjee (2005).

2 Notation

A country is characterized by a score vector $x = (x_1, \ldots, x_j, \ldots, x_n) \in R^n_{++}$, containing one score in each 'human development'-increasing dimension $j \in J = \{1, \ldots, n\}$.[4] There is disagreement about (i) which benchmark to use in each of the dimensions to normalize scores, and (ii) how to aggregate the resulting normalized scores. Because it is necessary to understand the aggregation procedure in order to understand the implications of the normalization procedure, we take the former issue first.

2.1 Aggregation

We briefly explain the method of aggregation for a (normalized) score vector $\mathbf{s} = (s_1, \ldots, s_n)$. To aggregate (normalized) scores, one typically uses an index, which maps (normalized) score vectors into a comparable number, representing the level of human development in the different countries, formally, $I : \mathbb{R}^n_{++} \to \mathbb{R} : \mathbf{s} \mapsto I(\mathbf{s})$. There can be disagreement whether (i) a country is better than another merely because the sum of its normalized scores is the higher of the two, (ii) when two countries have the same total, a country with less dispersed values should be regarded as equal to a country with more dispersed values, and so on. The issue of disagreement here is, as we have stated in the introduction, whether and to what extent the eventual index should be responsive to 'inequality' among the (normalized) components. To obtain a robust criterion, we will therefore focus on all indices which are (i) increasing and (ii) S-concave,[5] which we collect in a set $\mathbb{I} = \{I \mid I \text{ is increasing and } S\text{-concave}\}$. This family of indices is the basis for generalized Lorenz dominance (also called second-order stochastic dominance):

DEFINITION 1: $\mathbf{s}$ *generalized Lorenz dominates* $\mathbf{s}'$ *if and only if*

$$(1)\ \forall I \in \mathbb{I} : I(\mathbf{s}) \geq I(\mathbf{s}') \textit{ and } (2)\ \exists I \in \mathbb{I} : I(\mathbf{s}) > I(\mathbf{s}').$$

This dominance criterion can be easily implemented. Denote with $(s_{\langle 1\rangle}, \ldots, s_{\langle n\rangle})$ a permutation of $\mathbf{s}$ such that $s_{\langle 1\rangle} \leq s_{\langle 2\rangle} \leq \ldots \leq s_{\langle n\rangle}$. We get:

[4] Our methodology is sufficiently general to cope with (i) 'human development'-decreasing dimensions, e.g., the number of suicides in a country, and (ii) negative scores in some dimensions. In particular, such instances can be accomodated through straightforward modifications of the normalization procedure presented in Section 2.2.

[5] An index I is S-concave if and only if $I(B\mathbf{s}) \geq I(\mathbf{s})$, with B a bistochastic matrix, i.e., rows and columns sum up to 1 and all entries are positive.

PROPOSITION 1: $\mathbf{s}$ *generalized Lorenz dominates* $\mathbf{s}'$ *if and only if*

$$\forall j \in J : \sum_{k=1}^{j} s_{\langle k \rangle} \geq \sum_{k=1}^{j} s'_{\langle k \rangle} \text{ and } \exists j \in J : \sum_{k=1}^{j} s_{\langle k \rangle} > \sum_{k=1}^{j} s'_{\langle k \rangle}.$$

For later purposes, we mention some properties of the generalized Lorenz dominance criterion. It satisfies increasingness (a higher (normalized) score in each dimension improves human development) and anonymity (the names of the dimensions do not matter, or, permuting the scores over dimensions does not alter the level of human development). Both properties are common to all indices in $\mathbb{I}$ and therefore inherited by the generalized Lorenz dominance criterion. Furthermore, the generalized Lorenz dominance criterion also satisfies ratio-scale invariance: dividing the (normalized) scores in all dimensions by the same strictly positive scalar leaves the ranking of two countries unchanged.

At this point, it should be emphasized that the final outcome of our robustness tests will build on pairwise dominance comparisons rather than on checking the robustness of a country's 'index number'. As is evident from the foregoing, this is precisely because we want to check a necessary and sufficient dominance condition for a complete family of (increasing and S-concave) indices.[6]

2.2 Normalization

The fundamental reason for choosing to normalize the HDI's components is the lack of a 'known scientific relationship' which would otherwise serve to accommodate for such comparability problems *between* the dimensions under consideration, rather than the fact that the components were originally measured in different units (this point is also stressed by Ebert and Welsch, 2004). Explicitly taking this position towards the normalization problem is, admittedly, a rather nonstandard approach. The more usual approach is to present the same problem 'vertically', i.e., dimension-wise: one eventually gets rid of original measurement units by dividing each original sub-indicator value by another parameter measured on the same scale, to aggregate these pure numbers afterwards. In particular, the HDI's specific use

[6] Our approach thus differs from the robustness tests for composite indicators as discussed by Saisana *et al.* (2005). Essentialy, these authors consider combinations of different normalization and aggregation scenarios, and use Monte Carlo techniques to check how a country's index number changes conditional upon a change in the scenario. The end-product of such an approach is that the original index number is complemented by a confidence interval.

of the affine transformation, based on exogenous maximum and minimum goalposts for each of the original sub-indicators, highlights the benchmarking quality of the normalization step *within* each dimension. That is: the normalized scores are constructed in such a way as to tell the user how much better or worse a country is, in a specific dimension, as compared to some dimension-specific goalposts.

As stated in the introduction, there has already been considerable debate about the most appropriate technique for this normalization. Should one use exogenous goalposts, or the sample maxima and minima? Should one allow for a concave transformation, expressing the idea that an increase in the value of some sub-indicator is somehow better starting from a low level than the same increase when starting from a high level (e.g., Chakravarty, 2003)? Should this only be for GDP per capita, as is the case today? And, even if so, should GDP per capita indeed be measured in logarithms to start with (e.g. Anand and Sen, 2000)? Moreover, it has by now been documented by several of the aforementioned authors that a change in goalposts, let alone a more general change in the normalization method, may well lead to changes in the eventual country rankings.

The last observation is telling, as it immediately reveals that the choices taken in the preliminary normalization stage already add considerable 'horizontal' structure to the multi-dimensional comparison. Some authors (e.g., Panigrabi and Sivramkrishna, 2002; Lind, 2004) have specifically emphasized that the HDI's 'explicit', seemingly straightforward (equal) weighting scheme in fact conceals 'implicit' assumptions about the trade-offs between the sub-components that originate from the normalization stage. The complete normalization/aggregation problem is therefore more intricate than one is often inclined to think. Accordingly, when spelling out our normalization approach, we will pay special attention to clarifying its 'horizontal' implications.

Our main concern is with the fact that, for understandable reasons, there are differences of opinion about the comparability problem: some would say that the monetary value of 1 life year is at least x US\$, others would name a different figure, and still others would refrain from making such comparisons altogether. We now develop a formal approach to this disagreement problem. Our starting point is that such disagreement regarding the normalization issue essentially boils down to differences in opinion regarding the appropriate (dimension-specific) benchmarks. Of course, this intra-dimensional disagreement will carry over to the inter-dimensional comparibility issue, as we will subsequently demonstrate.

We will start from the thought experiment in which a set of possible benchmarks is considered for each dimension. Specifically, consider a 'percentile' p, with $0 \leq p \leq \frac{1}{2}$. For each dimension $j \in J$, define the set $B_j^p = \left[b_j^p, b_j^{1-p}\right]$ with b_j^p the p-th quantile of all scores in dimension j. Two extreme cases emerge. First, $p = \frac{1}{2}$ implies that everybody agrees to use the median score (in each dimension) as the only acceptable benchmark (for that dimension). The other extreme, i.e. choosing $p = 0$, entails that all values in between the minimal and maximal score can be regarded as plausible benchmarks. A benchmark vector is denoted by $b = (b_1, \ldots, b_n)$, which belongs to the bounded set $B^p = \left\{\mathbf{b} \mid \mathbf{b} \in \times_{j \in J} B_j^p\right\}$. Two points are worth stressing regarding this set-up. First, while we model consensus by choosing $p = \frac{1}{2}$, i.e., in each dimension the median score serves as the benchmark, it should be obvious that our methodology can also handle other choices of commonly agreed benchmark. Second, we choose p to be the same in each dimension mainly to facilitate the exposition, but -of course- one could also use p values that are specific for each dimension.

The next step completes the normalization stage: we "get rid of measurement units" by dividing each original score by the corresponding benchmark.[7] Thus, each eligible benchmark vector $b = (b_1, \ldots, b_n)$ transforms the original score vector $x = (x_1, \ldots, x_n)$ into a normalized score vector $s = (s_1, \ldots, s_n) \equiv \left(\frac{x_1}{b_1}, \ldots, \frac{x_n}{b_n}\right)$. Yet, as we have stressed above, this stage in fact entails more then just getting rid of measurement units. It is here that original values are rendered commensurable, and so the implications of our approach with regard to the inter-dimensional comparibility issue will be spelled out in more detail.

Again, consider first the case of $p = 1/2$. This is the full agreement case, as it implies that in each dimension the median value is commonly considered as the appropriate benchmark value. In turn, this agreement carries over to the way original values are transformed into normalized values. For example, in the case of the sample of our own empirical application (see section 4), a country with a life expectancy equal to 35 years, which is 50% of the sample's median life expectancy, obtains a score of 0.5 in this dimension, while a country with a GDP per capita equal to 10600 US$,

[7] Although we use a ratio scale transformation in this paper, it is worth noting that our methodologoy can also cope with more general affine transformations. This could be useful for applications with possibly negative scores. In such instances, empirical implementation of our robust outranking criteria requires a rather straightforward modification of the linear programming tests introduced in the next section.

twice the median value in our sample, obtains a score of 2. Consider then the class of indices we discussed in section 2.1. In view of the ratio-scale invariance property of this class, the foregoing also implies agreement about the latter achievement being four times higher than the former. Finally, and importantly, consider what happens if both scores are obtained by the same country and if we proceed to the aggregation stage. As all composite indices in the set $\mathbb{I}$ are increasing and anonymous, everyone would in that case agree that the achievement in GDP per capita must outweigh the achievement in life expectancy.

In fact, the interdimensional comparibility issue can also be presented differently. Because of ratio-scale invariance, one can also use one dimension, say dimension n, as the numeraire and normalize the scores in dimension $j = 1, \ldots, n$ by multiplying them with $b_n^{0.5}/b_j^{0.5}$. In terms of our example, and choosing GDP per capita as the numeraire, the number $b_n^{0.5}/b_j^{0.5}$ is the implicit price of dimension j, i.e., the value (in PPP US$ per capita) of 1 unit of dimension j. In our sample, e.g., we get a figure of about 75 US$ per person, per life year. As a consequence, an achievement in life expectancy equal to x outweighs an achievement in GDP of y if and only if $75x$ is larger than y.[8]

So far, we have focused on the extreme 'full agreement' scenario. Following the above reasoning, other (and often more realistic) disagreement scenarios can be captured by decreasing p. More specifically, we select lower values for p to introduce doubt about the exact implicit price of the different dimensions. For example, choosing $p = \frac{1}{4}$ (the inter-quartile range), leads to a situation where the implicit price of dimension j belongs to the interval $\left[b_n^{0.25}/b_j^{0.75}, b_n^{0.75}/b_j^{0.25}\right]$. For our sample, this means that the price of one life year must then lie somewhere in $[27, 214]$. As a consequence, there is no longer full agreement regarding the exact trade-off between life expectancy achievement and GDP performance.

[8] We should emphasize at this point that this figure is the implicit trade-off as it results from the (undoubtedly normative) normalization options taken by the analyst, rather than the "true" trade-offs that may possibly be retrieved from a positive model. See Dowrick et al. (2003) for a method of retrieving such positive information. Furthermore, the specific figure stated in the main text relates to the analysts' trade-off as it results from our proposed framework in the extreme universal agreement case. Lind (2004) spells out the implicit trade-offs that are a consequence of the HDI's actual (normative) modelling options.

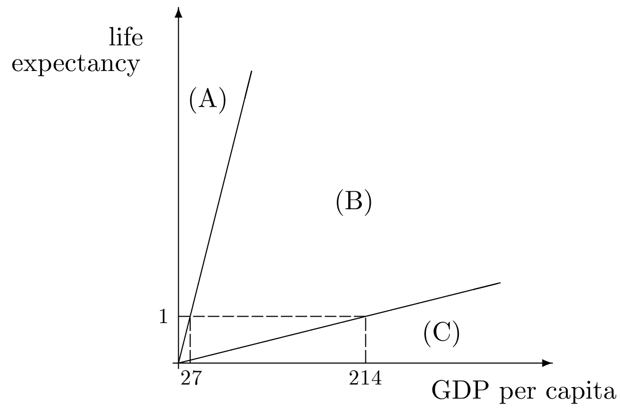

Figure 1: Partial comparability between different dimensions in case of benchmarking.

Figure 1 illustrates the dispute. We put GDP per capita on the horizontal and life expectancy on the vertical axis. There are three zones: zone A presents all combinations of life expectancy and GDP per capita where, even with disagreement about the exact trade-offs to be used, the former still contributes more to overall human development. We have the reverse situation in zone C. In zone B there is disagreement: according to some benchmark combinations (implicit prices) GDP outweighs life expectancy, while according to other benchmark combinations life expectancy outweighs GDP.

For later reference, we point out that such disagreement may still lead to one country dominating another country in the generalized Lorenz sense. To see this, note that for a country α finding itself in the disagreement zone the formal implication is that $s^{\alpha}_{Life\ years} \leq s^{\alpha}_{GDP/capita}$ for some benchmarks and $s^{\alpha}_{GDP/capita} \leq s^{\alpha}_{Life\ years}$ for some other vectors in B^p. (The same may well be true for a country β). Now, as stated in section 2.1, checking whether α dominates β implies checking ordered vector inequalities of the form $\sum_{k=1}^{j} s^{\alpha}_{\langle k \rangle} \geq \sum_{k=1}^{j} s^{\beta}_{\langle k \rangle}$. In terms of these inequalities, disagreement implies that there are several possibilities for ordering the normalized scores on the left hand side (if it pertains to country α's achievements) and/or on the right hand side (if it pertains to country β's achievements). To check for dominance means checking whether the cumulative ordered sum totals for country α would not be lower (and at least once higher) than those of country β. Such a scenario is certainly possible, not the least in the obvious case where country α outperforms β in each (normalized) dimension (see e.g., the next sections).[9]

The size of the disagreement zone crucially depends on p. If $p = 0.5$ both lines coincide and the disagreement zone B disappears. Decreasing p increases zone B. In the end, one may set $p = 0$ for all dimensions, the value that summarizes complete disagreement on the normalization issue. In that case, we consider all values in between the minimal and maximal (sample) score as plausible benchmarks for a specific dimension.

[9] Conversely, agreement does evidently not guarantee dominance (i.e., the vector inequalities in proposition 1 may not hold even if there is agreement on the normalized score vectors).

3 Robust outranking criteria

The dominance criterion of section 2.1. is based on pairwise country comparisons of (sums of) ordered normalized scores. The approach in section 2.2. essentially boils down to the statement that we allow for several possible scenarios regarding the magnitude, as well as the summation order, of these normalized scores. We can now put the different pieces of the puzzle together. At this final stage, we explicitly need to define a decision rule that takes the diversity of opinions into account. We will here define two dominance criteria, a strong one and a weak one, and will subsequently indicate how these can be implemented.

The strong criterion is inspired by Fleurbaey et al.'s (2004) generalized Lorenz dominance criterion with bounded equivalence scales to compare income distributions over households that differ in needs. We here adapt it to multi-dimensional comparisons of human development performances. In words, we say that a country (generalized Lorenz) dominates another country in a strongly robust sense if it does so *for every possible normalization* (while the former country does not dominate the latter country for some normalization). By a 'possible normalization' we mean a normalization that builds on benchmark vectors belonging to the set B^p. Intuitively, this means that it will be harder for a country to outperform another one as the disagreement about 'possible normalizations' increases. Indeed, as we have already stated in the introduction, one is likely to end up with an incomplete ordering of countries. The same idea can be stated alternatively as follows: as zones of disagreement get wider, a country's raw data rather than the analyst's normative modelling options become increasingly important in assessing that country's relative position in the sample. This is indeed a plausible way to think about the robustness problem.

We also consider a different approach. The robustness issue is ultimately rooted in the uncertainty surrounding the index' construction options, as reflected by allowing for a set of different normalizations. It is therefore arguable whether the strong criterion itself is convincing in all cases. For instance, while one country might not dominate another one for every possible normalization, it might do so for a subset of those normalizations, without a similar but reverse relationship holding. Carrying this reasoning to its extreme, we will thus also present a weak variant, which states that a country dominates another in a weakly robust sense if it does so for at least one possible normalization (while the opposite dominance relationship does not occur). In contrast with the strong dominance criterion, the weak version looks for support rather than unanimity in terms of the benchmark selection/possible normalizations. We may regard this weak criterion as conservative, in that it maintains dominance of the former country over the latter as long as there exists at least one possible normalization that supports this conclusion. In a sense, it allows countries with a different policy mix, and thus an outstanding performance in one of the dimensions, to put more weight on its best dimension in order to outperform other countries. Note that one can also refer to the robustness concern here: if a country is judged to have been outperformed on the basis of this weak criterion, it is far more due to its (feeble) raw data rather than to the composite index's artificiality.[10]

We first define the strong criterion. As introduced informally above, it states that a country with score vector $\mathbf{x}$ strongly dominates another country with score vector $\mathbf{y}$ if and only if (1) $\mathbf{x}$ is at least as good as $\mathbf{y}$ according to all (mutually agreed) benchmarks (in $\mathbb{B}^p$, with $0 \leq p \leq \frac{1}{2}$) and all indices (in $\mathbb{I}$) and (2) not the other way around, or, there exists at least one benchmark and one index such that $\mathbf{x}$ is strictly better than $\mathbf{y}$. We get:

DEFINITION 2: $\mathbf{x}$ *strongly dominates* $\mathbf{y}$ *for some* p, $0 \leq p \leq \frac{1}{2}$, *denoted* $\mathbf{x} \succ_*^p \mathbf{y}$, *if and only if*

$$(1)_*^p \quad \forall \mathbf{b} \in \mathbb{B}^p \, and \, \forall I \in \mathbb{I} : I\left(\frac{x_1}{b_1}, \ldots, \frac{x_n}{b_n}\right) \geq I\left(\frac{y_1}{b_1}, \ldots, \frac{y_n}{b_n}\right),$$

$$(2)_*^p \quad \exists \mathbf{b} \in \mathbb{B}^p \, and \, \exists I \in \mathbb{I} : I\left(\frac{x_1}{b_1}, \ldots, \frac{x_n}{b_n}\right) > I\left(\frac{y_1}{b_1}, \ldots, \frac{y_n}{b_n}\right).$$

Notice again that decreasing p increases the benchmark vector set and makes the criterion more robust, but at the cost of completeness. More precisely, for all p, q with $0 \leq p \leq q \leq \frac{1}{2}$, and for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_{++}$ we have $\mathbf{x} \succ_*^p \mathbf{y} \Rightarrow \mathbf{x} \succ_*^q \mathbf{y}$. In the end, $\succ_*^0$ is equal to the unanimity criterion: a country dominates another country if and only if the former dominates the latter in each dimension separately.

The strong dominance criterion $\succ_*^p$ can be implemented via linear programming techniques, verifying generalized Lorenz dominance (see proposition 1) for the 'least favorable' benchmark selection. Define $S_j(J)$ as the set containing all subsets $K \subseteq J$ with cardinality j. We get (the case for $(2)_*^p$

[10] This weak version is inspired by Cherchye et al. (2004). Specifically, focussing on a linear aggregation index, these authors suggest the endogenous selection of the most favorable linear aggregation weights. Here, we generalize this idea towards S-concave aggregation indices.

is analogous):

$$\begin{aligned}
(1)^p_* \text{ holds} \;&\Leftrightarrow\; \min_{\mathbf{b}\in\mathbb{B}^p}\min_{j\in J}\left[\min_{K\in S_j(J)}\left(\sum_{k\in K}\frac{x_k}{b_k}\right)-\min_{L\in S_j(J)}\left(\sum_{\ell\in L}\frac{y_\ell}{b_\ell}\right)\right]\geq 0\\
&\Leftrightarrow\; \max_{\mathbf{b}\in\mathbb{B}^p}\max_{j\in J}\left[\min_{L\in S_j(J)}\left(\sum_{\ell\in L}\frac{y_\ell}{b_\ell}\right)-\min_{K\in S_j(J)}\left(\sum_{k\in K}\frac{x_k}{b_k}\right)\right]\leq 0\\
&\Leftrightarrow\; \max_{\mathbf{b}\in\mathbb{B}^p}\max_{j\in J}\max_{K\in S_j(J)}\left[\min_{L\in S_j(J)}\left(\sum_{\ell\in L}\frac{y_\ell}{b_\ell}\right)-\sum_{k\in K}\frac{x_k}{b_k}\right]\leq 0\\
&\Leftrightarrow\; \max_{j\in J}\max_{K\in S_j(J)}\max_{\mathbf{b}\in\mathbb{B}^p}\left[\min_{L\in S_j(J)}\left(\sum_{\ell\in L}\frac{y_\ell}{b_\ell}\right)-\sum_{k\in K}\frac{x_k}{b_k}\right]\leq 0.
\end{aligned}$$

This can be implemented as follows:

PROPOSITION 2: $\mathbf{x}$ *strongly dominates* $\mathbf{y}$ *for some* p*,* $0 \leq p \leq \frac{1}{2}$*, denoted* $\mathbf{x} \succ^p_* \mathbf{y}$*, iff*

$(1)^p_* \quad \varepsilon^p_*(\mathbf{x},\mathbf{y}) \leq 0 \quad$ *and* $\quad (2)^p_* \quad \varepsilon^p_*(\mathbf{y},\mathbf{x}) > 0$*, with*

$\varepsilon^p_*(\mathbf{x},\mathbf{y}) = \max_{j\in J}\max_{K\in S_j(J)}\max_{\mathbf{b}\in\mathbb{B}^p} \alpha \quad$ *subject to*

$$\beta \leq \sum_{\ell\in L}\frac{y_\ell}{b_\ell}, \forall L \in S_j(J) \qquad \textit{and} \qquad \alpha \leq \beta - \sum_{k\in K}\frac{x_k}{b_k}.$$

Formally, the weak 'conservative' dominance criterion states that a country with score vector $\mathbf{x}$ weakly dominates another country with vector $\mathbf{y}$ if and only if there exists a benchmark (in $\mathbb{B}^p$, with $0 \leq p \leq \frac{1}{2}$) such that $\mathbf{x}$ is at least as good as $\mathbf{y}$ for all indices (in $\mathbb{I}$) and (ii) not the other way around, or there exists, for each benchmark at least one index such that $\mathbf{x}$ is strictly better than $\mathbf{y}$. We get:

DEFINITION 3: $\mathbf{x}$ *weakly dominates* $\mathbf{y}$ *for some* p*,* $0 \leq p \leq \frac{1}{2}$*, denoted* $\mathbf{x} \succ^p \mathbf{y}$*, if and only if*

$(1)^p \quad \exists \mathbf{b} \in \mathbb{B}^p$ *such that* $\forall I \in \mathbb{I} : I\left(\frac{x_1}{b_1},\ldots,\frac{x_n}{b_n}\right) \geq I\left(\frac{y_1}{b_1},\ldots,\frac{y_n}{b_n}\right)$,

$(2)^p \quad \forall \mathbf{b} \in \mathbb{B}^p : \exists I \in \mathbb{I}$ *such that* $I\left(\frac{x_1}{b_1},\ldots,\frac{x_n}{b_n}\right) > I\left(\frac{y_1}{b_1},\ldots,\frac{y_n}{b_n}\right)$.

Notice that weak dominance extends strong dominance, i.e., for some given p, $0 \leq p \leq \frac{1}{2}$ and for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^n_{++}$ we have $\mathbf{x} \succ^p_* \mathbf{y} \Rightarrow \mathbf{x} \succ^p \mathbf{y}$. But weak dominance might also introduce cycles, i.e., a sequence of countries

$\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k$ such that $\mathbf{x}^1 \succ^p \mathbf{x}^2 \succ^p \ldots \succ^p \mathbf{x}^k \succ^p \mathbf{x}^1$. Finally given p, q with $0 \leq p \leq q \leq \frac{1}{2}$, there is no relation between $\succ^p$ and $\succ^q$ anymore.

Attractively, the weak dominance can also be implemented via linear programming techniques, but here verifying generalized Lorenz dominance for the 'most favorable' benchmark. Define $\Pi(J)$, the set containing all permutations $\pi : J \to J : k \mapsto \pi(k)$. We get (the case for $(2)^p$ is again analogous):

$$
\begin{aligned}
(1)^p \text{ holds} \quad &\Leftrightarrow \quad \max_{\mathbf{b}\in\mathbb{B}^p} \min_{j\in J} \left[\min_{K\in S_j(J)} \left(\sum_{k\in K} \frac{x_k}{b_k} \right) - \min_{L\in S_j(J)} \left(\sum_{\ell\in L} \frac{y_\ell}{b_\ell} \right) \right] \geq 0 \\
&\Leftrightarrow \quad \min_{\mathbf{b}\in\mathbb{B}^p} \max_{j\in J} \left[\min_{L\in S_j(J)} \left(\sum_{\ell\in L} \frac{y_\ell}{b_\ell} \right) - \min_{K\in S_j(J)} \left(\sum_{k\in K} \frac{x_k}{b_k} \right) \right] \leq 0 \\
&\Leftrightarrow \quad \min_{\mathbf{b}\in\mathbb{B}^p} \min_{\pi\in\Pi(J)} \max_{j\in J} \left[\sum_{\ell=1}^{j} \frac{y_{\pi(\ell)}}{b_{\pi(\ell)}} - \min_{K\in S_j(J)} \left(\sum_{k\in K} \frac{x_k}{b_k} \right) \right] \leq 0 \\
&\Leftrightarrow \quad \min_{\pi\in\Pi(J)} \min_{\mathbf{b}\in\mathbb{B}^p} \max_{j\in J} \left[\sum_{\ell=1}^{j} \frac{y_{\pi(\ell)}}{b_{\pi(\ell)}} - \min_{K\in S_j(J)} \left(\sum_{k\in K} \frac{x_k}{b_k} \right) \right] \leq 0,
\end{aligned}
$$

where the crucial third step comes from the observation that $\forall \pi \in \Pi(J), \forall j \in J:$

$$
\min_{L\in S_j(J)} \left(\sum_{\ell\in L} \frac{y_\ell}{b_\ell} \right) - \min_{K\in S_j(J)} \left(\sum_{k\in K} \frac{x_k}{b_k} \right) \leq \sum_{\ell=1}^{j} \frac{y_{\pi(\ell)}}{b_{\pi(\ell)}} - \min_{K\in S_j(J)} \left(\sum_{k\in K} \frac{x_k}{b_k} \right),
$$

with an equality sign for at least one permutation $\pi \in \Pi(J)$. This can be implemented as follows:

PROPOSITION 3: $\mathbf{x}$ *weakly dominates* $\mathbf{y}$ *for some* p, $0 \leq p \leq \frac{1}{2}$, *denoted* $\mathbf{x} \succ^p \mathbf{y}$, *iff*

$(1)^p \qquad \varepsilon^p(\mathbf{x}, \mathbf{y}) \geq 0 \qquad$ *and* $\qquad (2)^p \qquad \varepsilon^p(\mathbf{y}, \mathbf{x}) < 0$, *with*

$\varepsilon^p(\mathbf{x}, \mathbf{y}) = \min_{\pi\in\Pi(J)} \min_{\mathbf{b}\in\mathbb{B}^p} \alpha$ *subject to*

$$
\forall j \in J, \forall K \in S_j(J) : \beta_j \leq \sum_{k\in K} \frac{x_k}{b_k} \qquad \textit{and} \qquad \forall j \in J : \alpha \geq \sum_{\ell=1}^{j} \frac{y_{\pi(\ell)}}{b_{\pi(\ell)}} - \beta_j.
$$

4 An illustrative application to UN 2002 data

As stated in the introduction, the media are eager to focus on the HDI-country rankings on the release of a new HDR. Media headlines (and official press releases as well) often have a clear hit-parade flavor ("For the fifth year in a row, Norway leads", "Niger displaced Sierra Leone at the bottom", "the Netherlands: down seven places!",...). But to what extent do such hit-parades hinge on the specific modelling options? Are some countries 'mistreated', in the sense that their ranking relative to other coutries would be reversed conditional upon a change in such modelling options? As stated above, these kinds of questions have been the subject of considerable research. In this section, we will use the tools developed above to investigate the extent to which country comparisons can still be made if one embeds modelling uncertainty from the very beginning. Our application uses the 2002 UN-dataset, which contains the life expectancy at birth (in years), the educational achievement (in %) —a weighted average of the adult literacy rate (% with weight 2/3) and the gross school enrolment rate (% with weight 1/3)— and the adjusted gross domestic product per capita (in PPP US$) for 176 countries.[11]

Recall that our check for consistency with generalized Lorenz dominance (i.e., with a class of indices) implies that we will not produce an index number as such. What we can produce, using this general test, are pairwise dominance results: a country may outperform another one, or the reverse, or there is no robust dominance relation between the two countries considered. Rather than overwhelming the reader with huge dominance tables, in the next subsection we will explore the robustness of the HDI ranking for two specific groups of countries, namely 24 OECD and 45 sub-Saharan African countries.[12] A descriptive analysis of the pairwise dominance relationships within these two country groups, which for the most part are situated at the two extremes of the UN HDI rankings, should provide some first insight into the robustness of the HDI. The second subsection looks at the full sample of 176 countries. There we complement the former descriptive analysis by statistical tests of the HDI's robustness. We will be concerned with two

[11] These data are downloadable at http://hdr.undp.org/statistics/data/.

[12] Some summary information about dominance for the complete sample is provided in the appendix. Specifically, for the (extreme) criteria $\succ_*^0$ and $\succ^0$ we there tabulate for each country the number of observations that are (robustly) dominated by that country and the number of observations that are dominating the country. In addition, we rank the countries for each criterion on the basis of a net dominance metric, i.e., the number of dominated observations minus the number of dominating observations.

types of robustness assessment. First, we will address robustness of the HDI country ranking itself. Second, we will take a broader view and ask whether the HDI's classification of countries (into low, middle and high human development classes) is robust. In both subsections we use the strong and weak dominance criteria, and we capture disagreement by considering $p = 1/4$ (which corresponds to the inter-quartile range as the benchmark set).[13]

4.1 Weakly and strongly robust dominance: OECD and sub-Saharan Africa

We start with the OECD countries. Table 1 shows the strong pairwise dominance relationships and Table 2 the weak relationships (according to $\succ_*^{1/4}$ and $\succ^{1/4}$) for 24 OECD countries. A '1' means that the corresponding row country dominates the column country, while a '0' means that the corresponding row and column countries cannot be compared.

[Insert tables 1 and 2]

Notice that countries are ranked according to the HDI (from highest =1 to lowest = 24). Thus, whenever we find (in tables 1 and 2) a '1' below the diagonal, we have a 'rank reversal', i.e., a country with a worse HDI rank which, using our robust method, dominates a country with a better HDI rank. Whenever we find a '0' above the diagonal, we have a 'disputable ranking', i.e., a country with a worse HDI rank that, given the perspective we uphold, cannot be compared to a country with a better HDI rank. From inspection of the tables, rank reversals do not occur. However, many 'disputable HDI rankings' exist: almost 75% of all possible pairwise comparisons for the strong criterion. To recall, this means that if there is disagreement about the appropriate normalization benchmark, three out of four times there is either no (generalized Lorenz) dominance relation between the two countries, or the finding that country α outperforms country β may be reversed, conditional upon which normalization benchmark (taken from B^p) is actually used. Such disputable rankings occur in 50% of the cases for the weak criterion. Or stated verbally: we seek at least one possible normalization for which country α (generalized Lorenz) dominates β while at the

[13] Recall from our discussion in section 2.2 that $p = 1/2$ and $p = 0$ correspond to the extreme scenarios of, respectively, full agreement and full disagreement regarding the appropriate benchmark choice. We choose to focus on the intermediate scenario $p = 1/4$ given the mainly illustrative purpose of this application.

same time the reverse relation does not hold, but in half of the cases this normalization cannot be found. Moreover, a large part (i.e., 25%) of the remaining 'undisputable' cases is essentially trivial, as they would also pass the unanimity test (i.e., the dominating country achieves a higher score in each separate dimension). Thus, if we disregard the uncontroversial unanimity rankings, none of the HDI rankings are robust according to the strong dominance criterion $\succ_*^{1/4}$, while only 25% is considered robust when using the weak version.[14]

There is also another sense in which the HDI's complete ranking (i.e., its 'hit-parade nature') can be questioned. If we zoom in on the strong criterion, some countries never dominate another country (Ireland, Luxembourg, Italy, Greece and the Republic of Korea), while others are never dominated by another country (Norway, Sweden, Australia, Canada, the Netherlands, Iceland, Japan, Switzerland and Luxembourg). Notice that one of the OECD countries is developed 'by default'. More precisely, Luxembourg never dominates another country (due to its low school enrollment rate), but is never dominated (due to its high GDP per capita). Other countries —the United States, Japan and Ireland— are 'almost' developed 'by default' (dominated by at most one country and dominating at most one country), but for different reasons. More specifically, while the United States and Ireland attain a high GDP per capita, but a moderate life expectancy and school performance, Japan achieves a high life expectancy and a poor school performance.

Diverging circumstances or policy choices may lead to more unequal (normalized) scores. In contrast to the HDI, our methodology is sensitive to inequality and thus puts such a less balanced development at a disadvantage.

Now, let us focus on the weak criterion. Using this 'conservative' criterion, we retain only a single country (Republic of Korea) that never dominates another country, while some other countries (Norway, Sweden, Australia and Luxembourg) are never dominated. Also note that the Republic of Korea is now dominated by all other countries. Recall that the weak criterion is much more flexible compared to the strong one: it allows countries to put more weight on their best achievement over the different dimensions. Therefore, the United States, Ireland and Japan are able to dominate other countries (and are dominated by others). Only Luxembourg remains (almost) developed 'by default'.

Finally, we look at the 45 sub-Saharan African countries in our sample. In tables 3 and 4 we report the dominance tables for $\succ_*^{1/4}$and $\succ^{1/4}$. These

[14] Recalling our discussion in section 3, this means that for the OECD countries the dominance table according to $\succ_*^{1/4}$ corresponds with the unanimity ranking based on $\succ_*^{0}$.

countries are again ranked according to the HDI (from highest =1 to lowest = 45).

[Insert tables 3 and 4]

Again no rank reversals occur for the strong criterion, but now there are 81 rank reversals for the weak criterion. Let us zoom in on the extreme cases, viz. countries dominating other countries which are nonetheless ranked at least 10 positions higher according to the HDI (for the subgroup of sub-Saharan African countries). Although Guinea has a worse school performance, it is able to dominate Congo, because the weak criterion allows it to pick out one specific normalization so as to put sufficient weight on its best achievement (relative to Congo), namely its GDP per capita figure. Apparently, the opposite relationship does not hold, i.e., Congo cannot put sufficient weight on its best achievement (relative to Guinea) to outperform Guinea. The same mechanism (via a relatively high GDP per capita) allows (i) Angola to dominate both Madagascar and Nigeria and (ii) Cote d'Ivoire to dominate Madagascar.

Although we find some rank reversals (according to the weak dominance criterion), there are less disputable rankings in this subset, i.e., fewer countries that are ranked by the HDI but which turn out to be non-comparable once disagreement has been introduced. More precisely, about 59% and 97% of all pairwise comparisons are considered robust by respectively the strong and the weak dominance criterion. But again, we should remark that 21% of all pairwise comparisons are due to unanimity. Thus, disregarding the uncontroverial unanimity rankings leaves us with 38% of all comparisons that are robust according to our strong criterion, while 75.6% are robust in terms of the weak criterion.

Finally, using the strong dominance criterion, the Seychelles are never dominated, while at the other end, Burkina Faso, Mali and Niger never dominate another country. Even stronger, the Seychelles succeed in dominating all other countries, while Niger is dominated by all other countries. Moving to the weak criterion, the Seychelles again are never dominated —as they dominate all other countries— while Niger never dominates another country —as it is dominated by all other countries.

4.2 Statistical robustness analysis of the HDI

The above results for the OECD and sub-Saharan African subsamples demonstrate the usefulness of our method in descriptively assessing the robustness of pairwise dominance relationships. As a general conclusion, we may state that they question the robustness of the HDI as a ranking tool. Indeed, for both subsamples we obtain the result that a large part of the HDI country dominance relationships do not appear to be robust with respect to the strong outranking criterion; and a similar conclusion, albeit to a somewhat lesser extent, applies for the (more conservative) weak criterion. We now complement this descriptive assessment by a statistical investigation, which considers the full sample of 176 countries.

Before presenting the statistical results, we start with some general observations that allow for a rough assessment of the robustness of the HDI ranking. Obviously, the HDI ranking is complete and thus allows us to rank 100% of all possible pairwise country comparisons. Table 5 decomposes this percentage of pairwise comparisons into robust comparisons and non-robust comparisons (in columns), using the $\succ_*^{1/4}$ and $\succ^{1/4}$ dominance criteria (in rows) to check robustness.

	robust due to		non-robust due to	
	unanimity	non-unanimity	non-comparability	rank reversals
$\succ_*^{1/4}$	69.66%	4.39%	25.94%	0.01%
$\succ^{1/4}$	69.66%	23.44%	1.70%	5.19%

Table 5: decomposition of the HDI ranking into robust and non-robust comparisons.

The explanation of the table is as follows. If we focus on the strong criterion $\succ_*^{1/4}$, the percentage of robust comparisons is about 74% (= 69.66% + 4.39%). Still, the bulk of these pairwise dominance relationships (namely, 69.66%) is due to unanimity, i.e., a country which dominates another country by a higher (normalized) score in each dimension. Non-robust comparisons are mainly due to rankings which are judged non-comparable according to $\succ_*^{1/4}$, rather than rank reversals. More precisely, only two rank reversals took place: South-Africa (resp. Indonesia) is more developed than Tajikistan (resp. Equatorial Guinea) according to $\succ_*^{1/4}$, while Tajikistan (resp. Equatorial Guinea) has a higher HDI rank.[15] Using the weak criterion $\succ^{1/4}$, the picture is very different. 93.1% of all pairwise comparisons are robust and a significant part (23.44%) is not due to unanimity. Thus, there are possible

[15] Actually, the rank reversal of Equatorial Guinea is quite probably due to a data problem: the UN reports a (tenfold?) high GDP figure for this country and also uses it in its calculation of the HDI rank.

normalizations that corroborate the HDI ranking. This is mirrored by a serious decline in the number of non-decisive rankings (only 1.70%). However, one also observes somewhat more rank reversals (5.19%). To conclude, if we disregard the unanimity part —which are basically non-controversial rankings— the weak criterion seems to confirm the HDI ranking, while the strong criterion clearly disagrees with the HDI in the sense that it considers a significant part of countries as mutually incomparable.

We next want to obtain a more precise 'statistical' statement regarding the robustness of the HDI ranking, as well as of the UNDP's concomittant classification of the countries in human development groups. We proceed as follows. On the basis of the HDI ranking, the United Nations arranges countries into three groups, ranging from low over middle to high human development. We employ χ^2 goodness-of-fit tests to investigate the extent to which the distribution of the robust dominance relationship over the different classes complies with the expected distribution under the null hypothesis that the HDI ranking/classification is effectively robust.

Table 6 presents the results for the HDI country classification using the strong dominance criterion $\succ_*^{1/4}$. The second column (with label '#') contains the number of countries in each human development class. In the following columns, each cell ij (row i and column j) presents the number of times a country ranked by the UN in class i dominates a country ranked in class j; this number is expressed as a percentage of the maximum number of pairwise comparisons possible (for classes i and j). Moving from the lower left to the upper right corner, the percentages increase. This stands to reason: countries with a higher (lower) value for the original HDI ranking are more likely to dominate (to be dominated by) others.

j / i	#		1	2		3
1	55		51.11	84.78		100.00
2	86		0.00	48.04		90.44
3	36		0.00	0.00		36.67
with diagonals:	χ^2	=	1732.18	(p	=	0.000)
without diagonals:	χ^2	=	137.90	(p	=	0.000)

Table 6: Comparing $\succ_*^{1/4}$ and the United Nations' human development classification.

If the HDI *ranking* were a fully robust tool for ranking individual countries, then all cells on and above the diagonal in Table 6 would value 100 percent.

Hence, the χ^2 tests that include the diagonal values in the table can be conceived as a (weak) check for robustness of the HDI country ranking. Next, robustness of the HDI *classification* only requires 100 percent in the cells above the diagonal, whence the robustness of this classification can be checked by excluding the diagonal values in the χ^2 tests.

Table 6 reports values that are rather far below 100 percent, not only on the diagonal (even as low as 36.67%) but also above the diagonal (as low as 84.04%). The χ^2 tests confirm this impression. When including the diagonal entries, we strongly reject the hypothesis that the HDI is a robust *ranking* tool. When excluding the diagonal cells, which means that we consider robustness of the HDI *classification*, the test statistic decreases sharply, but the value remains far from reasonably probable under the null hypothesis of a strongly robust HDI classification.

This first assessment casts serious doubt on the robustness of the HDI's results. Yet, it is based on the rather stringent dominance criterion $\succ_*^{1/4}$, and may be conceived as too harsh. Table 7 uses the milder criterion $\succ^{1/4}$ as a base of comparison.

j / i	#		1	2		3
1	55		87.81	99.22		100.00
2	86		0.76	99.53		98.61
3	36		0.00	0.94		91.75
with diagonals:	χ^2	=	27.32	(p	=	0.000)
without diagonals:	χ^2	=	0.89	(p	=	0.642)

Table 7: Comparing $\succ^{1/4}$ and the United Nations' human development classification.

Notice that the percentages on and above the diagonal are now much closer to 100%. The results of the χ^2 tests are mixed: while they reject robustness of the HDI ranking (see the results that include the diagonal values), they support robustness of the HDI classification (see the results that exclude the diagonals).

Tables 6 and 7 partly confirm the rough picture sketched in table 5. Using the strong dominance criterion $\succ_*^{1/4}$ to test the robustness of both the HDI country ranking and the country classification leads to rejection. However, the weak dominance criterion $\succ^{1/4}$ is milder: although the complete

HDI ranking is not robust, at least as a classification device it can stand the robustness test. Therefore, as a main conclusion, we may state that our analysis suggests that the (incomplete) HDI classification is reasonably robust, while the opposite holds for the (complete) HDI country ranking.[16]

[16] Of course, the above findings are contingent on the specific disagreement level we have started from. Thus, it is certainly worth mentioning that our (weak) robustness test cannot reject robustness of the HDI ranking on the basis of the weak criterion $\succ^0$.

5 Conclusion

We have presented a robust approach to rank alternatives (e.g., different countries), when there are multiple relevant performance dimensions (e.g., the different dimensions of human development). First, we suggested a formal normalization procedure that seeks to express the usual controversies associated with this step in composite indicator bilding. Basically, this is achieved by incorporating a wide range of transformation possibilities, each associated with an alternative benchmark performance selection. Second, we proposed a generalized Lorenz dominance criterion for aggregating the different performance dimensions, which effectively obtains a ranking criterion that is consistent with the whole class of increasing, S-concave indices. Third, we proposed a weak as well as a strong dominance criterion, based on the previous normalization/aggregation step. Both versions can be implemented via linear programming, which is convenient from a practical point of view.

Our assessment of the Human Development Index illustrates the practical usefulness of the procedure. In particular, our application has focused on testing the robustness of (i) the United Nations' HDI itself as a tool for ranking individual countries and (ii) the corresponding country classification, which is based on the HDI, into low, middle or high development. We found that the HDI classification (i.e., essentially an incomplete country ranking) may be conceived as reasonably robust, while the opposite holds for the (complete) HDI ranking. Specifically, we have shown that there are many cases in which the HDI country ranking may be called 'disputable', which means that two countries turn out to be incomparable or are subject to a rank reversal when accounting for disagreement regarding the appropriate evaluation scheme. In fact, such disputable rankings are mainly situated 'within' the HDI classes rather than 'between' countries of different classes, which explains the more favorable testing results for the HDI classification. Of course these findings can be criticized, in that our normalization/aggregation methodology and the one actually used by UNDP are quite different. For example, the fact remains that we have used GDP per capita, whereas the HDI is (not uncontroversially) built on its logged value. Note however that this is at least partly neutralized by the very fact that we allow for a broad range of implicit trade-off values. The same holds for the aggregation procedure we proposed: it admittedly cannot produce HDI-like index numbers, but the rationale of this 'shortcoming' is precisely that we want to check several possible indices simultaneously. Finally, and more importantly, even if our results with regard to the HDI are deemed only partially substantive, our application still makes the general demonstration that the proposed methodology may be instrumental for a descriptive as well as a statistical robustness investigation of existing country rankings/classifications when multiple performance dimensions are relevant.

An interesting avenue for further research consists in endogenously defining a classification of countries on the basis of the (weakly or strongly) robust outranking criteria as they have been defined in this study. Specifically, this would boil down to conceiving a country classification such that each country in a higher class (resp. weakly or strongly) robustly dominates each country in a lower class, while the opposite never occurs. Such an endogenous classification exercise is similar in spirit to Noorbakhsh's (1988). Another interesting question is whether the gap between strong and weak dominance —between 'all' and 'at least one' allowable normalization(s) supporting a country— can be filled, e.g., by using a criterion that tells us what percentage of all possible normalizations favours one country over another.

Appendix

For the criteria $\succ_*^0$ (see "strong dominance") and $\succ^0$(see "weak dominance") we tabulate, for each individual country, the number of observations that is (robustly) dominated by that country and the number of observations that is dominating the country. We rank the countries for each criterion on the basis of a "net dominance" metric, which is calculated as the number of dominated minus the number of dominating observations.We stress the illustrative nature of these summarizing results; other ranking procedures are equally possible. Given the specific orientation of the current paper, our main focus is on the pairwise dominance relationships. More detailed (country-specific) pairwise dominance information is given in Tables 1-4; additional results are available from the authors upon simple request.

strong dominance				
rank	country	number dominated by country	number dominating country	net dominance
1	Australia	162	0	162
2	Norway	161	0	161
3	Canada	159	0	159
4	Belgium	159	1	158
5	Sweden	157	0	157
6	Netherlands	155	0	155
7	Iceland	153	0	153
8	United Kingdom	151	2	149
8	Austria	151	2	149
10	Switzerland	148	0	148
10	France	151	3	148
10	New Zealand	151	3	148
10	Spain	150	2	148
14	Finland	149	2	147
15	Denmark	147	1	146
16	United States	145	1	144
17	Ireland	144	1	143
18	Japan	142	0	142
19	Germany	144	8	136
20	Luxembourg	131	0	131
20	Israel	137	6	131
22	Italy	138	8	130
23	Portugal	139	10	129
24	Greece	136	13	123
25	Slovenia	138	16	122
26	Korea, Rep. of	133	13	120
27	Barbados	132	13	119
28	Singapore	127	15	112
29	Argentina	124	19	105
30	Hong Kong, China (SAR)	100	1	99
30	Poland	119	20	99
32	Cyprus	116	18	98
33	Czech Republic	120	25	95
34	Estonia	101	9	92
35	Malta	106	15	91
36	Lithuania	108	21	87
36	Uruguay	112	25	87
38	Chile	109	27	82
39	Slovakia	109	29	80
40	Brunei Darussalam	101	23	78
40	Saint Kitts and Nevis	88	10	78
42	Croatia	107	30	77
43	Hungary	99	23	76
44	Seychelles	101	26	75
45	Costa Rica	93	22	71
46	Latvia	94	26	68
47	Cuba	82	24	58
48	Bahrain	87	30	57
49	Qatar	74	23	51
50	Bulgaria	84	34	50
51	Kuwait	76	27	49
52	Belarus	73	27	46
53	Libyan Arab Jamahiriya	83	40	43
53	Macedonia, TFYR	80	37	43
55	Mexico	78	39	39
55	Trinidad and Tobago	79	40	39
57	Panama	74	36	38
58	United Arab Emirates	58	21	37
59	Albania	73	37	36

60	Ukraine	68	33	35
62	Russian Federation	61	27	34
62	Saint Lucia	72	38	34
64	Malaysia	72	39	33
64	Suriname	74	41	33
66	Bahamas	59	28	31
67	Bosnia and Herzegovina	69	39	30
68	Tonga	62	34	28
69	Venezuela	68	41	27
70	Mauritius	66	40	26
70	Romania	68	42	26
70	Jamaica	60	34	26
73	Samoa (Western)	66	43	23
74	Armenia	58	36	22
75	Colombia	67	46	21
76	Saint Vincent and the Grenadines	60	41	19
77	Kazakhstan	54	36	18
78	Turkmenistan	54	37	17
79	Azerbaijan	57	41	16
80	Brazil	57	42	15
80	Philippines	62	47	15
80	Maldives	55	40	15
83	Lebanon	61	47	14
84	Oman	49	37	12
84	Thailand	61	49	12
86	Saudi Arabia	49	38	11
86	Dominica	58	47	11
86	Georgia	48	37	11
89	Jordan	60	53	7
90	Turkey	62	56	6
90	Tunisia	53	47	6
92	Fiji	59	54	5
92	Paraguay	61	56	5
94	Peru	60	56	4
95	Sri Lanka	56	53	3
96	Belize	56	55	1
96	Iran, Islamic Rep. of	56	55	1
96	Occupied Palestinian Territories	48	47	1
99	China	60	60	0
100	Grenada	52	53	-1
101	Guyana	48	50	-2
102	Ecuador	56	59	-3
103	Dominican Republic	54	58	-4
103	Uzbekistan	33	37	-4
105	Kyrgyzstan	30	38	-8
106	El Salvador	54	65	-11
107	Syrian Arab Republic	50	63	-13
108	Cape Verde	54	69	-15
109	South Africa	27	45	-18
110	Bolivia	47	67	-20
111	Algeria	47	69	-22
112	Mongolia	33	56	-23
113	Tajikistan	16	43	-27
114	Gabon	34	67	-33
115	Viet Nam	46	80	-34
116	Moldova, Rep. of	26	62	-36
117	Indonesia	49	86	-37
118	Botswana	12	56	-44
119	Honduras	45	91	-46
120	Nicaragua	44	91	-47

121	Namibia	16	69	-53
122	Egypt	36	90	-54
123	Guatemala	39	97	-58
124	Sao Tome and Principe	24	83	-59
125	Morocco	28	91	-63
126	Equatorial Guinea	23	95	-72
127	Solomon Islands	24	100	-76
128	India	29	107	-78
129	Vanuatu	13	96	-83
130	Cambodia	28	112	-84
130	Ghana	29	113	-84
132	Swaziland	2	92	-90
133	Papua New Guinea	22	114	-92
134	Lao People's Dem. Rep.	22	116	-94
134	Zimbabwe	1	95	-94
136	Myanmar	16	113	-97
136	Lesotho	2	99	-97
138	Comoros	19	117	-98
139	Bhutan	16	116	-100
140	Sudan	17	120	-103
141	Cameroon	14	118	-104
142	Bangladesh	13	118	-105
143	Uganda	12	118	-106
144	Kenya	8	115	-107
145	Nepal	12	122	-110
145	Togo	12	122	-110
145	Congo	9	119	-110
148	Pakistan	7	118	-111
148	Mauritania	8	119	-111
150	Djibouti	10	123	-113
151	Haiti	12	126	-114
152	Yemen	5	124	-119
152	Guinea	3	122	-119
154	Nigeria	6	126	-120
154	Gambia	7	127	-120
156	Madagascar	4	126	-122
157	Timor-Leste	0	124	-124
158	Angola	0	125	-125
159	Benin	6	132	-126
160	Eritrea	5	132	-127
160	Rwanda	2	129	-127
162	Senegal	4	132	-128
162	Malawi	1	129	-128
164	Zambia	0	130	-130
165	Côte d'Ivoire	4	135	-131
166	Tanzania, U. Rep. of	1	134	-133
167	Central African Republic	1	142	-141
168	Congo, Dem. Rep. of the	2	144	-142
168	Mozambique	1	143	-142
170	Chad	1	144	-143
171	Mali	1	145	-144
171	Burkina Faso	0	144	-144
173	Ethiopia	2	150	-148
174	Burundi	1	151	-150
175	Niger	0	151	-151
176	Guinea-Bissau	0	155	-155
177	Sierra Leone	0	166	-166

weak dominance				
rank	country	number dominated by country	number dominating country	net dominance
1	Luxembourg	176	0	176
2	Norway	175	1	174
3	Ireland	173	2	171
4	United States	172	3	169
5	Denmark	172	4	168
6	Canada	171	5	166
7	Iceland	170	6	164
8	Switzerland	168	7	161
9	Netherlands	168	8	160
10	Australia	167	9	158
11	Austria	166	10	156
12	Belgium	165	11	154
13	Sweden	164	12	152
14	Japan	162	13	149
15	France	162	14	148
16	United Kingdom	161	15	146
17	Finland	160	16	144
18	Germany	159	17	142
19	Italy	157	18	139
20	Hong Kong, China (SAR)	157	19	138
21	Singapore	156	20	136
22	New Zealand	155	21	134
23	Spain	154	22	132
24	Israel	153	23	130
25	Greece	152	24	128
26	Slovenia	151	25	126
27	Portugal	150	26	124
28	Cyprus	149	27	122
29	Brunei Darussalam	148	28	120
30	United Arab Emirates	147	29	118
31	Korea, Rep. of	146	30	116
32	Malta	145	31	114
33	Qatar	144	32	112
34	Seychelles	143	33	110
35	Barbados	142	33	109
36	Czech Republic	140	35	105
37	Bahrain	140	36	104
38	Bahamas	139	37	102
39	Kuwait	138	38	100
40	Hungary	137	39	98
41	Estonia	136	40	96
42	Slovakia	135	41	94
43	Saint Kitts and Nevis	134	42	92
44	Argentina	132	43	89
44	Poland	132	43	89
46	Lithuania	129	45	84
46	Oman	130	46	84
48	Chile	128	47	81
48	Croatia	128	47	81
50	Antigua and Barbuda	127	49	78
51	Saudi Arabia	126	50	76
52	Latvia	125	51	74
53	Mauritius	124	52	72
54	Costa Rica	123	53	70
55	Trinidad and Tobago	122	53	69
56	Mexico	120	55	65
57	Uruguay	119	56	63
58	Malaysia	119	57	62
59	Russian Federation	118	58	60
60	Libyan Arab Jamahiriya	117	59	58

61	Brazil	116	60	56
62	Bulgaria	115	61	54
63	Tonga	114	62	52
64	Thailand	113	62	51
65	Macedonia, TFYR	111	64	47
66	Suriname	111	65	46
67	South Africa	110	66	44
68	Romania	109	67	42
69	Grenada	107	68	39
70	Panama	107	69	38
71	Bosnia and Herzegovina	105	71	34
71	Colombia	104	70	34
73	Belarus	103	72	31
74	Cuba	103	73	30
74	Dominican Republic	102	72	30
76	Tunisia	101	75	26
77	Kazakhstan	100	76	24
78	Turkey	99	77	22
79	Iran, Islamic Rep. of	98	78	20
80	Samoa (Western)	97	79	18
81	Venezuela	96	80	16
82	Saint Lucia	95	81	14
83	Fiji	94	82	12
84	Belize	93	83	10
85	Ukraine	91	84	7
86	Albania	91	85	6
87	Saint Vincent and the Grenadines	89	87	2
87	Dominica	88	86	2
89	Peru	88	88	0
90	Maldives	87	89	-2
91	Algeria	86	90	-4
92	Paraguay	85	91	-6
93	China	83	91	-8
94	Gabon	83	93	-10
95	Lebanon	82	94	-12
96	Cape Verde	80	93	-13
97	Turkmenistan	81	95	-14
98	Philippines	78	96	-18
98	El Salvador	79	97	-18
100	Jordan	77	99	-22
101	Botswana	76	100	-24
102	Jamaica	75	101	-26
103	Guyana	74	102	-28
104	Namibia	73	103	-30
105	Sri Lanka	72	104	-32
106	Armenia	70	104	-34
106	Ecuador	71	105	-34
108	Azerbaijan	69	107	-38
109	Syrian Arab Republic	68	108	-40
110	Guatemala	67	109	-42
111	Indonesia	65	110	-45
112	Egypt	65	111	-46
113	Georgia	64	112	-48
114	Occupied Palestinian Territories	63	113	-50
115	Morocco	62	114	-52
116	Bolivia	61	115	-54
117	Honduras	60	116	-56
118	Viet Nam	59	117	-58
119	Nicaragua	57	118	-61
120	Swaziland	57	119	-62

121	Equatorial Guinea	56	120	-64
122	Uzbekistan	54	121	-67
122	Kyrgyzstan	54	121	-67
124	India	53	122	-69
125	Mongolia	52	124	-72
126	Moldova, Rep. of	51	125	-74
127	Vanuatu	50	126	-76
128	Ghana	49	127	-78
129	Cambodia	48	128	-80
130	Solomon Islands	47	129	-82
131	Papua New Guinea	46	130	-84
132	Sao Tome and Principe	45	131	-86
133	Lesotho	44	132	-88
134	Zimbabwe	43	133	-90
135	Cameroon	41	135	-94
136	Tajikistan	39	134	-95
137	Lao People's Dem. Rep.	40	136	-96
138	Bhutan	39	136	-97
139	Comoros	38	137	-99
140	Sudan	37	138	-101
141	Pakistan	35	139	-104
141	Mauritania	36	140	-104
143	Bangladesh	34	142	-108
144	Djibouti	33	143	-110
145	Togo	32	144	-112
146	Uganda	31	145	-114
147	Myanmar	30	146	-116
148	Haiti	29	147	-118
149	Nepal	28	148	-120
150	Guinea	27	148	-121
151	Gambia	26	150	-124
152	Kenya	25	151	-126
153	Congo	24	152	-128
154	Angola	22	152	-130
155	Senegal	22	154	-132
156	Rwanda	21	155	-134
157	Côte d'Ivoire	20	156	-136
158	Nigeria	19	157	-138
159	Yemen	18	158	-140
160	Benin	17	159	-142
161	Madagascar	16	160	-144
162	Eritrea	15	161	-146
163	Central African Republic	14	162	-148
164	Chad	13	163	-150
165	Zambia	12	163	-151
166	Mozambique	11	165	-154
167	Tanzania, U. Rep. of	10	166	-156
168	Timor-Leste	9	167	-158
169	Malawi	8	168	-160
170	Ethiopia	7	169	-162
171	Congo, Dem. Rep. of the	6	170	-164
172	Burundi	4	169	-165
173	Guinea-Bissau	5	171	-166
174	Mali	3	173	-170
175	Burkina Faso	2	174	-172
176	Sierra Leone	1	175	-174
177	Niger	0	176	-176

References

[1] Anand, S., Sen, A. (2000) The income component of the Human Development Index, *Journal of Human Development* 1, 83-106.

[2] Booysen, F. (2002) An overview and evaluation of composite indices of development, *Social Indicators Research* 59, 115-151.

[3] Chakravarty, S.R. (2003) A generalized Human Development Index, *Review of Development Economics* 7, 99-114.

[4] Chatterjee, S.K. (2005) Measurement of human development: an alternative approach, *Journal of Human Development* 6, 31-53.

[5] Cherchye, L., Moesen W., Van Puyenbroeck T. (2004) Legitimately diverse, yet comparable: on synthesising social inclusion performance in the EU, *Journal of Common Market Studies* 42, 919-955.

[6] Dar, H. A. (2004) On making human development more humane, *International Journal of Social Economics* 31, 1071-1088.

[7] Despotis, D.K. (2004) A reassessment of the Human Development Index via data envelopment analysis, *Journal of the Operational Research Society*, 1-12.

[8] Dowrick, S., Dunlop, Y., Quiggins, J. (2003) Social indicators and comparisons of living standards, *Journal of Development Economics* 70, 501-529.

[9] Ebert, U., Welsch, H. (2004), Meaningful environmental indices: a social choice approach, *Journal of Environmental Economics and Management* 47, 270-283.

[10] Fleurbaey, M., Hagneré, C., Trannoy, A. (2003) Welfare comparisons with bounded equivalence scales, *Journal of Economic Theory* 110, 309-336.

[11] Fukada-Parr, S. (2001) Indicators of human development and human rights- overlaps, differences and what about the Human Development Index?, *Statistical Journal of the United Nations* 18, 239-248.

[12] Fukada-Parr, S., Shiva Kumar, A.K. (2003) *Readings in Human Development*, Oxford University Press: New Delhi.

[13] Klasen, S. and Bardhan, K. (1999), UNDP's Gender-Related Indices: A Critical Review, *World Development* 27, 985-1010.

[14] Lind, N. (2004) Values reflected in the Human Development Index, *Social Indicators Research* 66, 283-293.

[15] Luchters, G., Menkhoff, L. (2000) Chaotic signals from HDI measurement, *Applied Economics Letters* 7, 267-270.

[16] Maasoumi, E., Nickelsburg, G. (1988) Multivariate measures of well-being and an analysis of inequality in the Michigan data, *Journal of Business and Economic Statistics* 6, 327-334.

[17] Mahlberg, B., Obersteiner, M. (2001), Remeasuring the HDI by data envelopment analysis, International Institute for Applied Systems Analysis Interim Report 01-069.

[18] Majumder, A., Chakravarty, S. (1996) Achievement and improvement in living standards, *Journal of Development Economics* 50, 189-195.

[19] Mazumdar, K. (2003) A new approach to Human Development Index, *Review of Social Economy* 61, 535-549.

[20] Nardo M. et al. (2005) *Handbook on Constructing Composite Indicators: Methodology and User's Guide*, OECD Statistics Working Paper (JT00188147).

[21] Noorbakhsh, F. (1988) A modified Human Development Index, *World Development* 25, 517-528.

[22] Panigrahi R., Sivramkrishna, S. (2002) An adjusted Human Development Index: robust country rankings with respect to the choice of fixed maximum and minimum indicator values, *Journal of Human Development* 3(2), 301-311.

[23] Sagar, A., Najam, A. (1998) The Human Development Index: a critical review, *Ecological Economics* 25, 249-264.

[24] Saisana, M., Saltelli, A., Tarantola, S. (2005) Uncertainty and sensitivity analysis as tools for the quality assessment of composite indicators, *Journal of the Royal Statistical Society Series A* 168, 1-17.

[25] United Nations Development Programme (1990-2004) *Human Development Reports*. http://www.undp.org.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	norway		0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	1	0	1	0	1	1	1	1
2	Sweden	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	1
3	Australia	0	0		0	0	1	0	0	0	0	0	1	1	0	0	1	0	1	1	0	1	1	1	1
4	Canada	0	0	0		0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1	1	1	1
5	Netherlands	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1
6	Belgium	0	0	0	0	0		0	0	0	0	0	1	1	0	0	0	0	1	1	0	1	1	1	1
7	Iceland	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0
8	United States	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
9	Japan	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
10	Ireland	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	Switzerland	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	1	0	1	1	0	0
12	United Kingdom	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	1	1
13	Finland	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	1	1
14	Austria	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	0	0	1	0	0
15	Luxembourg	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0
16	France	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	1	0	0
17	Denmark	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	1	1
18	New Zealand	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	1	1	1
19	Germany	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	0
20	Spain	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	0	1
21	Italy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0
22	Greece	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0
23	Portugal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1
24	Korea, Rep. of	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Table 1: strong dominance relations for OECD subsample; interquartile range as benchmark range (0 = row element does not dominate column element; 1 = row element dominates column element)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	norway		0	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
2	Sweden	0		0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	1
3	Australia	0	0		0	0	1	0	0	1	0	0	1	1	0	0	1	0	1	1	1	1	1	1	1
4	Canada	0	0	0		0	0	0	0	1	0	0	0	1	1	0	1	0	0	1	1	1	1	1	1
5	Netherlands	0	0	0	0		0	0	0	1	0	0	0	1	1	0	1	0	0	1	1	1	1	1	1
6	Belgium	0	0	0	0	0		0	0	1	0	0	1	1	0	0	1	0	1	1	1	1	1	1	1
7	Iceland	0	0	0	0	0	0		0	0	0	0	0	0	1	0	1	0	0	1	0	1	1	1	1
8	United States	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
9	Japan	0	0	0	0	0	0	0	0		0	0	0	0	0	0	1	0	0	1	1	1	1	1	1
10	Ireland	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	1	0	1	1
11	Switzerland	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	1	0	1	1	1	1
12	United Kingdom	0	0	0	0	0	0	0	0	0	0	0		1	0	0	0	0	0	0	1	1	1	1	1
13	Finland	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	1	1	1	1	1
14	Austria	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	0	1	1	1	1
15	Luxembourg	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1
16	France	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	1	1	1	1
17	Denmark	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	0	1	1	1	1
18	New Zealand	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	0	1	1	1
19	Germany	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	1	1	1
20	Spain	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	1	1
21	Italy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1
22	Greece	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1	1
23	Portugal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1
24	Korea, Rep. of	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Table 2: weak dominance relations for OECD subsample; interquartile range as benchmark range (0 = row element does not dominate column element; 1 = row element dominates column element)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
1	Seychelles		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	Mauritius	0		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	Cape Verde	0	0		0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	Equatorial Guinea	0	0	0		0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	South Africa	0	0	0	0		0	0	1	1	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	Gabon	0	0	0	0	0		1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	São Tomé and Principe	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	0	1	0	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1
8	Namibia	0	0	0	0	0	0	0		0	0	0	1	0	0	0	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	Botswana	0	0	0	0	0	0	0	0		0	0	1	0	0	0	1	1	0	1	1	1	1	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	Ghana	0	0	0	0	0	0	0	0	0		0	0	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	Comoros	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
12	Swaziland	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	1	0	1	1	0	0	1	0	1	0	1	1	1	1	1	1
13	Sudan	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
14	Cameroon	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	1	0	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
15	Togo	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	1	1	0	0	1	0	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1
16	Congo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	1	1	1	0	1	1
17	Lesotho	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	1	1	1	1	1	1
18	Uganda	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1
19	Zimbabwe	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1	1	1	1	1	1
20	Kenya	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1	0	1	0	1	1	1	0	1	1
21	Madagascar	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	0	0	1
22	Nigeria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	1	0	1	1	0	0	1	1
23	Mauritania	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	1	0	1	0	0	0	0	0	1	0	1	0	1	0	1	1	1	1	1	1
24	Gambia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1
25	Eritrea	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	1	1
26	Senegal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	1	1
27	Rwanda	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	1	0	0	1	0	0	1	0	1	1	1	1	1
28	Guinea	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1
29	Benin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	1	0	0	1	0	1	1	1	0	1	1
30	Tanzania, U. Rep. of	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
31	Côte d'Ivoire	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	1	1	1	1	1	1	1	1	1
32	Zambia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	1	0
33	Malawi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	1
34	Angola	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	1	0	1	1	1	1	1	1
35	Chad	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	1	0	1	0	1	1
36	Congo, Dem. Rep. of the	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	0	0	0	1
37	Central African Republic	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	1	1	1
38	Ethiopia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	0	0	0	0	1
39	Mozambique	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	1	1
40	Guinea-Bissau	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1
41	Burundi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	1
42	Mali	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	0
43	Burkina Faso	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0
44	Niger	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
45	Sierra Leone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Table 3: strong dominance relations for Subsaharan African subsample; interquartile range as benchmark range (0 = row element does not dominate column element; 1 = row element dominates column element)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
1	Seychelles		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	Mauritius	0		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	Cape Verde	0	0		1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	Equatorial Guinea	0	0	0		0	0	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	South Africa	0	0	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	Gabon	0	0	1	1	0		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	São Tomé and Principe	0	0	0	0	0	0		0	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	Namibia	0	0	0	1	0	0	1		0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	Botswana	0	0	0	1	0	0	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	Ghana	0	0	0	0	0	0	1	0	0		1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	Comoros	0	0	0	0	0	0	0	0	0	0		0	1	0	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	Swaziland	0	0	0	1	0	0	1	0	0	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	Sudan	0	0	0	0	0	0	0	0	0	0	0	0		0	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	Cameroon	0	0	0	0	0	0	0	0	0	0	1	0	1		1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	Togo	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	Congo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	1	0	0	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	Lesotho	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	Uganda	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0		0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	Zimbabwe	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	0	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	Kenya	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1	1	0	0	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	Madagascar	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	1	1	1	1	1	1
22	Nigeria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1		0	0	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	1	1	1	1	1	1
23	Mauritania	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	Gambia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0		1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	Eritrea	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	0	1	1	0	0	1	0	1	0	1	1	1	1	1	1
26	Senegal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1		1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
27	Rwanda	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0		0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
28	Guinea	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	Benin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0		1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1
30	Tanzania, U. Rep. of	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	0	0	0	0	0	0	0	1	1	1	1
31	Côte d'Ivoire	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1	1		1	1	0	1	1	1	1	1	1	1	1	1	1	1
32	Zambia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0		1	0	0	1	0	1	0	1	1	1	1	1	1
33	Malawi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1	1	1	1
34	Angola	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	0	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1
35	Chad	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0		1	0	1	0	1	1	1	1	1	1
36	Congo, Dem. Rep. of the	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	1	1	1	1
37	Central African Republic	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1		1	1	1	1	1	1	1	1
38	Ethiopia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	1	1	1	1	1
39	Mozambique	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1		1	1	1	1	1	1
40	Guinea-Bissau	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1	1	1	1	1
41	Burundi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1	1	1	1
42	Mali	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1	1	1
43	Burkina Faso	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1	1
44	Niger	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
45	Sierra Leone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	

Table 4: weak dominance relations for Subsaharan African subsample; interquartile range as benchmark range (0 = row element does not dominate column element; 1 = row element dominates column element)