OPHI
OXFORD POVERTY & HUMAN DEVELOPMENT INITIATIVE
www.ophi.org.uk

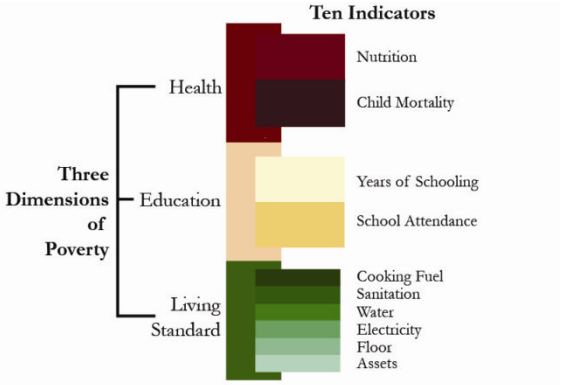
UNIVERSITY OF OXFORD

Is the Multidimensional Poverty Index robust to different weights?

Sabina Alkire, Maria Emma Santos, Suman Seth, Gaston Yalonetzky, November 2010

In 2010, the United Nations Development Programme Human Development Report and OPHI released an international Multidimensional Poverty Index (MPI) that reflects multiple deprivations in education, health and living standard that people face at the same time. How robust is this index to the different choices made in its design? This brief presents the results of tests that show that the MPI rankings are robust to a range of plausible changes in weights.

Figure 1: Dimensions and Indicators of the MPI



The MPI is constructed using three dimensions consisting of ten indicators described in Figure 1 (Alkire and Santos, 2010). In the MPI, each dimension is equally weighted at one third; each indicator within a dimension is also equally weighted. For example, the nutrition indicator is assigned 1/6 weight and the sanitation indicator receives 1/18 weight.

The decision to weight health, education, and standard of living dimensions equally arises in part from various discussions around, and robustness tests of, the Human Development Index, as well as both expert opinion and participatory analysis. Also, choosing dimensions such that they are roughly equal helps make the resulting measure easy to understand and use (Atkinson *et al.,* 2002).

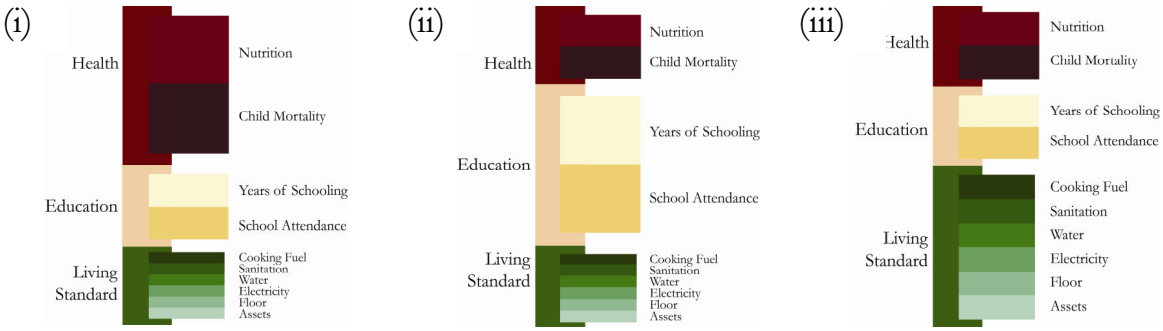
Amartya Sen, among others, sees the need to set weights in multidimensional measures as a strength, not an embarrassment: "There is indeed great merit… in having public discussions on the kind of weights that may be used" (1997). After all, any national budget implicitly sets weights on many dimensions of welfare, often with little debate. Yet given the legitimate diversity of human values, Sen also argues that it may not be necessary to agree on a precise set of weights: ideally, measures would be developed that are robust to a range of weights.

So *is the MPI robust to a range of weights?* To test this, we estimated the MPI using three additional weighting structures (Figure 2): (i) giving 50% weight to health and 25% weight each to education and standard of living, (ii) giving 50% weight to education and 25% weight each to health and standard of living, and finally (iii) giving 50% weight to standard of living and 25% weight each to health and education.[1] Then we verify if the country rankings are stable using four approaches. First, we calculate the correlation coefficients between each pair of rankings using three different methods: Pearson's correlation coefficient, Spearman's rank correlation coefficient and Kendall's rank correlation coefficient (Tau-b).[2] Next, we estimate the concordance between all four rankings using three methods: Kendall and Dickinson-Gibbon

Citation: Alkire, Sabina, Santos, Maria Emma, Seth, Suman & Gaston Yalonetzky. 2010. Is the Multidimensional Poverty Index robust to different weights. Oxford Poverty & Human Development Initiative (OPHI) Research in Progress Series. Available at: www.ophi.org.uk/publications/ophi-research-in-progress.

(KDG), the multi-rank version of Spearman's coefficient (by Kendall, KS) and the multiple-rank concordance index of Joe (J), and perform a famous test of rank independence by Friedman. We also explore the percentage of pairwise country comparisons that are robust for all weighting structures, and explore 'large' changes in rankings among different countries.

Figure 2: Different MPI Weighting Structures Tested



Correlations: In Table 1, we report the three pairwise correlation coefficients between the rankings under the equal weight structure and each of the three alternative weighting structures for all 104 countries.

Changing the indicators' weights indeed affects the poverty estimates. However, the country rankings thus generated remain quite stable. In the third column of Table 1 we compare the rankings across all countries. There, one can see that the minimum of the three Pearson's correlation coefficient is 0.989, the minimum of the three Spearman's coefficient is 0.981 and the minimum of the three Kendall's Tau-b is 0.903.

Table 1: Correlations between MPI and adjusted MPIs having 50% weight on each dimension in turn and 25% on the remaining two dimensions

Pair of Rankings Compared	Correlation Coefficient	All Countries
MPI with Equal Weight and MPI with 50% weight on Education	Pearson	0.991
	Spearman	0.984
	Tau-b	0.903
MPI with Equal Weight and MPI with 50% weight on Health	Pearson	0.995
	Spearman	0.981
	Tau-b	0.909
MPI with Equal Weight and MPI with 50% weight on Living Stand.	Pearson	0.989
	Spearman	0.989
	Tau-b	0.916
Total Number of Observations		**104**

Interestingly, the correlation of the rankings obtained with the three alternative weighting systems is also high (see Table 2). While comparing all 104 countries, the lowest rank correlation between all measures with unequal weights is the Kendall Tau-b coefficient between the ranking obtained assigning 50% of relative weight to education and the one obtained assigning 50% of weight to health, and it is 0.856.

Table 2: Correlations among MPIs adjusted to have 50% on each dimension in turn and 25% on the remaining two dimensions

Pair of Rankings Compared	Correlation Coefficient	All Countries
MPI with 50% weight Education and MPI with 50% weight on Health	Pearson	0.985
	Spearman	0.957
	Tau-b	0.836
MPI with 50% weight Education and MPI with 50% weight on Living Stan.	Pearson	0.966
	Spearman	0.970
	Tau-b	0.854
MPI with 50% weight Health and MPI with 50% weight on Living Stan.	Pearson	0.978
	Spearman	0.968
	Tau-b	0.856
Total Number of Observations		**104**

Concordance: In addition to considering pairwise correlations we also look at measures of concordance across the four ranks: the original one generated by equal weighting and the three alternatives considered. As mentioned earlier, we use three indices of intra-group rank concordance: the one by Kendall and Dickinson-Gibbon (KDG), the multi-rank version of Spearman's coefficient (by Kendall, KS) and the multiple-rank concordance index of Joe (J). In the latter two cases we use the complete-tie-sensitive variations developed by Seth and Yalonetzky (2010).[3] These indices are equal to zero when the ranking criteria are independent from each other, and are equal to one if and only if the ranking criteria are perfectly consistent. We also performed a well-known test of rank independence by Friedman. The null hypothesis of rank independence, for all countries, is strongly rejected with 99% confidence (See p-value reported, along with the other concordance results, in Table 3).

Table 3: Indicators of rank concordance for the four ranks of MPI generated using alternative weights.

Index	All Countries
Friedman statistic[4]	404.190
Degrees of freedom	103
P-value	<1%
KDG	0.981
KS	0.975
J	0.983

According to the results in Table 3, the degree of rank concordance is very high among the countries, at 0.975 or higher.

Pairwise Comparisons: As a third closely related exercise we compared the MPI estimates for all possible pairs of countries across all four different weighting structures. In 88% of the total possible pairs, one country has higher poverty than the other *regardless* of the weighting system.

Large Rank Changes: Finally, we looked at the countries that changed rank 10 places or more. To do so, we focus on the countries whose MPI scores range from 0.05 to 0.64. If we focus on these – the bottom 60 countries – we find that only 5 countries change ranks of 10 or more places. Three countries (Guyana, Zambia and Kenya) improve their rank position when 50% weight is applied to education (in 14, 11 and 10 places correspondingly). One country, Chad, improves its rank position in 12 places when health receives 50% of weight and Cote d' Ivoire

improves its rank position in 11 places when 50% weight is applied to the living standard dimension. Among the top 44 countries, there are 14 countries with rank changes at 10 or above; however the MPI values among these countries are very restricted – from 0 to 0.05 – and the MPI seems less able to discriminate among such small changes.[5]

Of course it is not only weights that are important. Alkire and Santos (2010) analyse the MPI robustness to changes in the indicators and dimensional cutoffs used, also finding high correlations between rankings. The MPI was also tested to changes in the value of the poverty cutoff. Recall that the MPI identifies a person as multidimensionally poor if they are deprived in 30% of the dimensions. To test the robustness of this decision, the percentage of dimensions in which a person must be deprived to be identified as multidimensionally poor was varied from 20% to 40%, instead of using only the value of 30%. When each country's estimate is compared with each other, it is found that in 95.5 % of all possible pairs one country is poorer than the other *regardless* of the poverty cutoff. These results suggest that the 30% poverty cutoff used for the MPI is not a critical choice that dramatically affects results. The rankings are quite stable and robust for a plausible range of values.

As detailed in the paper, data constraints call for cross-country comparisons to be undertaken with great caution. Data come from different years and from different surveys, 11 countries lack two or three indicators and a few countries have significant sample reductions due to missing information. In some cases (highlighted in the paper) estimates must be interpreted as lower or upper bounds of multidimensional poverty. Moreover, even if rankings are robust to a range of weights, the MPI values for each country and its subsequent analysis will reflect, of necessity, only one weighting structure and the analysis would be affected were another used.

However, what this note has tried to demonstrate is that, despite all the limitations, the ranking of countries is quite robust to the particular selection of weights. MPI ranks are robust for 88% of pair-wise comparisons, the correlations between the MPI and each of three alternative weighting structures is 0.90 or above, and the Friedman test of rank concordance rejects the null hypothesis of rank independence with 99% confidence. As analysed in the paper, rankings are also robust to the particular selection of indicators' cutoffs and poverty cutoffs.

References

Alkire, S. and Santos, M. E. (2010), "Acute Multidimensional Poverty: A New Index for Developing Countries", OPHI Working Paper No 38.

Atkinson, A. B., Cantillon, B., Marlier, T., Nolan, B. (2002). *Social indicators: The EU and social inclusion*. Oxford: Oxford University Press.

Seth, S. and Yalonetzky, G. (2010), "Assessing rank robustness with complete-tie-sensitive indices of rank concordance".

Sen, A. (1997). 'Quality of Life and Economic Evaluation, *Academia Economic Papers, 25*(3), 269-316.

Appendix: Robustness and Surveys

Note that the MPI was estimated using three internationally comparable surveys: the Demographic and Health Survey (DHS) for 48 countries, the Multiple Indicators Cluster Survey (MICS) for 35 countries and the World Health Survey (WHS) for 19 countries. In undertaking the robustness checks it is evident that countries whose data come from the World Health Survey behaved differently than the others, although the rankings still show concordance. In this section we report the same results as above. However we have additional reported results for different subgroups of countries, in order to have a better understanding of the source of non-robustness, if any. In particular, we look at the bottom 75 countries, as well as those whose data

are drawn from DHS, MICS, and WHS surveys. Two countries – Mexico and Argentina – drew the data from national surveys and these are included in the 'non-WHS' column.

Table 4, below, repeats the results for Table 1 for the bottom 75 countries and the different survey sources. While all results remain strong, it is evident that the country rankings among the WHS countries are least robust. If we consider the non-WHS countries, the minimum of Kendall's Tau-b is 0.910, whereas in WHS it is 0.635.

Table 4: Correlations between MPI and adjusted MPIs by Survey type

Pair of Rankings Compared	Correlation Coefficient	All Countries	Bottom 75 Countries	DHS Only	MICS Only	DHS & MICS	Non WHS	WHS Only
MPI with Equal Weight and MPI with 50% weight on Education	Pearson	0.991	0.987	0.986	0.997	0.991	0.991	0.981
	Spearman	0.984	0.982	0.984	0.969	0.986	0.987	0.936
	Tau-b	0.903	0.893	0.906	0.889	0.908	0.910	0.830
MPI with Equal Weight and MPI with 50% weight on Health	Pearson	0.995	0.993	0.994	0.997	0.996	0.996	0.987
	Spearman	0.981	0.991	0.991	0.991	0.996	0.995	0.774
	Tau-b	0.909	0.930	0.929	0.940	0.950	0.945	0.635
MPI with Equal Weight and MPI with 50% weight on Living Stand.	Pearson	0.989	0.983	0.981	0.996	0.988	0.988	0.991
	Spearman	0.989	0.983	0.983	0.986	0.989	0.989	0.955
	Tau-b	0.916	0.896	0.903	0.916	0.915	0.915	0.864
Total Number of Observations		**104**	**75**	**48**	**35**	**83**	**85**	**19**

Correlations among all 'Adjusted' MPIs (having 50% weight on some dimension) are also high for all survey groups except among WHS countries (see Table 5). While comparing the non-WHS countries, the minimum of the Kendall's Tau-b always exceeds 0.8. However, when we compare the nineteen WHS countries only, the Kendall's Tau-b is only 0.472 between the ranking obtained assigning 50% of relative weight to education and the ranking obtained assigning 50% of weight to health.

Table 5: Correlations among MPIs adjusted by Survey

Pair of Rankings Compared	Correlation Coefficient	All Countries	Bottom 75 Countries	DHS Only	MICS Only	DHS & MICS	Non WHS	WHS Only
MPI with 50% weight Education and MPI with 50% weight on Health	Pearson	0.985	0.978	0.979	0.994	0.985	0.986	0.962
	Spearman	0.957	0.970	0.980	0.949	0.981	0.979	0.616
	Tau-b	0.836	0.854	0.885	0.842	0.883	0.880	0.472
MPI with 50% weight Education and MPI with 50% weight on Living Stan.	Pearson	0.966	0.949	0.940	0.989	0.962	0.963	0.950
	Spearman	0.970	0.951	0.946	0.971	0.970	0.972	0.859
	Tau-b	0.854	0.805	0.819	0.872	0.856	0.860	0.715
MPI with 50% weight Health and MPI with 50% weight on Living Stan.	Pearson	0.978	0.966	0.965	0.990	0.978	0.979	0.984
	Spearman	0.968	0.968	0.965	0.975	0.980	0.978	0.763
	Tau-b	0.856	0.854	0.849	0.882	0.881	0.876	0.639
Total Number of Observations		**104**	**75**	**48**	**35**	**83**	**85**	**19**

As one might expect, the rank concordance indices are also reduced to between 0.81 and 0.86 in the case of the subsample of countries for which the WHS was used, whereas they remain high for the other country groupings. But when we perform the three tests of rank concordance on all of the sub-groups of countries, we find that all – including the WHS countries – reject the null hypothesis of rank independence with 99% confidence.

Table 6: Indicators of rank concordance for all Survey Sources

Index	All Countries	Bottom 75 Countries	DHS Only	MICS Only	DHS & MICS	Non WHS	WHS Only
Friedman statistic[6]	404.190	290.28	184.46	133.3	323.97	331.76	61.902
Degrees of freedom	103	74	47	34	82	84	18
P-value	<1%	<1%	<1%	<1%	<1%	<1%	<1%
KDG	0.981	0.981	0.981	0.980	0.987	0.987	0.863
KS	0.975	0.974	0.975	0.973	0.983	0.983	0.817
J	0.983	0.974	0.968	0.984	0.982	0.983	0.862

World Health Survey: Clearly from this analysis, the World Health Survey values have contributed disproportionally to the lack of robustness to weights. Ten of the WHS countries show a rank change of 10 or more places. In addition, for eight of them the MPI values are, in any case, only a lower or upper bound. Three of the eight WHS countries which MPI estimate should be interpreted as a lower bound are also among the ten that show a rank change of 10 or more places. These are Latvia, Sri Lanka and Tunisia.[7] Thus, for 15 of the 19 WHS countries (Sri Lanka is in both categories), the MPI values appear to be less accurate. While the concordance tests are still valid, clearly this dataset has been problematic. There are two plausible reasons for this. The first is data quality: while the DHS and MICS are implemented regularly, the WHS was only implemented once, and the quality of the data collected is known to vary considerably among countries. The second is MPI values. Most of the large rank changes occur in the very top of the MPI spectrum, among countries that have relatively low (and similar) MPI values. For example the MPI range among the top 44 countries is 0 to 0.05; the remaining 60 countries cover a range from 0.05 to 0.64. Sixteen of the top forty-four countries use WHS data, and the comparisons among these are arguably much more sensitive than among the bottom 60 countries.

[1] In such way, in one of the alternative weightings each of the educational indicators weight 25%, each of the health indicators 12.5% and each of the living standard indicators 4.16%. In the other, each of the health indicators weights 25%, each of the two education indicators weights 12.5% and the living standard indicators the same as above. Finally, in the weighting structure that gives higher weight to living standard, each of these indicators weights 8.33 % and each of the health and education indicators weights 12.5%.

[2] Pearson's correlation coefficient measures the linear relationship between a pair of rankings. The Spearman's rank correlation coefficient is based on the changes in country ranks between a pair of rankings. The Kendall's Tau-b coefficient is calculated by comparing each pair of countries in a pair of rankings.

[3] The index by Kendall and Dickinson-Gibbon is a complete-tie-sensitive version of an earlier index by Kendall.

[4] Under the null hypothesis of independence of ranks the statistic is asymptotically distributed as Chi-square with degrees of freedom equal to the number of rank observations (e.g. countries) minus one.

[5] These 14 countries are: Argentina, Brazil, Czech Republic, Ecuador, Estonia, Hungary, Jordan, Latvia, Occupied Palestinian Territories, Sri Lanka, Trinidad and Tobago, Tunisia, Ukraine and Uruguay.

[6] Under the null hypothesis of independence of ranks the statistic is asymptotically distributed as Chi-square with degrees of freedom equal to the number of rank observations (e.g. countries) minus one.

[7] Latvia worsens its relative position (ie. becomes poorer) 14 places when health receives 50% of weight. Sri Lanka improves its relative position (ie. becomes less poor) 18 places when education receives 50% of weight and worsens its relative position 11 places when living standard receives 50% of weight. Tunisia improves its relative position 16 places when education receives 50% of weight. Hungary, Czech Republic and Estonia are three other WHS countries which experience big changes in their rank positions. When health receives 50% of weight, Hungary and Czech Republic worsen their relative position by 22 and 27 places correspondingly, and Estonia improves its relative position in 27 places.