Human Development Index and Multidimensional Poverty Index: Evidence on their Reliability and Validity

Beja, Edsel Jr.

Ateneo de Manila University

1 March 2021

Human Development Index and Multidimensional Poverty Index: Evidence on their Reliability and Validity

Edsel L. Beja Jr.

Department of Economics, Ateneo de Manila University
Quezon City, The Philippines 1108

Email: edsel.beja@gmail.com

Abstract

The paper tests the reliability and the validity of the Human Development Index (HDI) from the *United Nations Development Programme* and the Multidimensional Poverty Index (MPI) from the *Oxford Poverty & Human Development Initiative*. The results indicate that there is strong reliability within metric; and that there is strong validity between metrics. The findings imply that both HDI—more specifically, the inverse, (1-HDI)—and MPI embody equivalent descriptions about the level of human development or the lack of it.

Keywords: Human Development Index, Multidimensional Poverty Index, Reliability, Validity
JEL Codes: C43, I30

1. Introduction

The UNDP introduced the Human Development Index (HDI) in 1990 as a metric on the level of human development in a country. The HDI focuses on income, education, and health; and it uses four indicators to measure these essential dimensions. In particular, GDP per capita (PPP US$) is a proxy for income; life expectancy at birth is a proxy for health; and both mean years and expected years of schooling proxy for education. Since 2010, the calculation of the HDI relies on the geometric mean, thereby allowing the effect of a change in one dimension to be the same for a proportionate change in any of the other dimensions. With the HDI, a country knows how close it is in achieving human development.

In 2010, the UNDP introduced the Multidimensional Poverty Index (MPI) from the *Oxford Poverty & Human Development Initiative* (OPHI). The MPI shares with the HDI in terms of dimensions but goes further by incorporating more indicators. There are six indicators for income, namely electricity, sanitation, water, house flooring, fuel for cooking, and assets; two for education, namely years of schooling and child attendance at school; and two for health, namely child mortality and nutrition. There is an overall weight of 1/3 for each dimension; but, given the composition of the dimensions, there is in effect a relative weight of 1/6 for each item in education and in health and a relative weight of 1/18 for each item in income. The MPI is calculated using a threshold count approach—that is, an individual is deemed

poor if the number of her deprivations total to at least a third of the weighed items. In a way, the MPI sets a "minimum" condition as necessary to be able to pursue a life that is worth living. With the MPI, a country knows how far it is from achieving human development.

The measurement protocols for the HDI and MPI are already well known to analysts.[1] What is probably more important to point out is that the HDI reports values that range between 0 and 1 to represent the level of attainment in relation to the target of 1. The closer a country is to 1 the higher then is the level of human development. The MPI, while it also reports values between 0 and 1, is about the level of deprivation. In this case, the closer a country is to 1 the higher then is the lack of human development.

Juxtaposed, the HDI and MPI are thus potentially mirror-images of each other.[2] As such, it is sensible to assert not only their reliability given their respective conceptual bases but also their validity given their purported inverse relationship. Such an assertion—reliability (i.e., repeat measures of HDI or of MPI) and validity (i.e., the relation between HDI and MPI)—is empirically verifiable. Yet, there is no study to date that does so. This paper is therefore an attempt to fill the void by presenting some evidence.

The paper contains four parts. Following the introduction, Part 2 presents the methodology; then Part 3 gives the results. The last part concludes the paper.

2. Methodology

2.1 Concept

Anand and Sen (2003, p. 119) are clear in arguing that "an attainment perspective is more relevant in assessing how well a country is doing, whereas the shortfall perspective is more relevant in looking at the difficulty of the task still remaining." As such, one can assert that metrics that look into human development need to share the same "perspective". The same goes for metrics on human deprivation.

The HDI and MPI are alternative representations in the context of Anand and Sen: the HDI uses an attainment perspective; whereas the MPI uses a shortfall perspective. This paper in turn argues for a transformation of the HDI into a shortfall perspective; that is, more specifically, HDS = 1–HDI, where HDS stands for human development shortfall index. Doing so does not assert that the MPI is a better metric than the HDI. Rather, the transformation goes in line with the view of Alkire and Santos (2010; see also Sen 1979), who argue that micro-level data reflect the situation on the ground more accurately that macro-level data do. The main point is just for both HDS and MPI to share the same perspective.

---

[1] Excellent references on the HDI and MPI are Fukuda-Parr and Kumar (2003) and Alkire, Foster, et al. (2015), respectively. See also Ravallion (2001) for a discussion on indices of development. Beja (2014) presents a revision to the HDI. Beja (2016) is an attempt at combining objective and subjective indicators of ill-being.

[2] HDI and MPI are arguably mirror-images of each other if and only if there is symmetry in the distribution of outcomes, which is unlikely.

2.2 Procedure

The paper uses the standard tests in checking for reliability and validity (see, e.g., Netemeyer et al. 2003 and Gregory 2015). Reliability is about the stability of a metric when measured repeatedly, and in this paper the test-retest approach is the appropriate procedure. In addition, the paper looks at the "steadiness" of correlations across the repeat measures as an indicator of the overall reliability of a metric. Steadiness, in turn, is verified through the evenness of the level of correlation across repeat validity tests.

Meanwhile, validity is the extent by which a metric measures what it is designed to measure. A way to check for validity is to use the divergent-convergent validity test, and in this paper a high degree of correlation between alternative metrics is evidence of the level of validity. Metrics that are mirror images of each other are expected to show a high level of negative correlation (i.e., divergent validity viz. HDI and MPI); whereas those that share the same perspective are expected to show a high level of positive correlation (i.e., convergent validity viz. HDS and MPI).

Lastly, regression analysis examines the relationship between the HDS and MPI. The standardized coefficients in turn show the effect of the HDS on MPI, or vice versa.

2.3 Dataset

The datasets used for the analysis are publicly available on the Internet. The HDI is from the UNDP website, and the MPI is from the OPHI website. Data on the HDI is available each year; but that on the MPI is not, given that the micro-level data are normally not available every year for most countries.

The OPHI dataset contains 121 countries (Alkire et al. 2020). A country may report one to six data points between 2010 and 2020. But for the analysis, the OPHI dataset is clustered into three periods with 2011, 2014, and 2018 as time markers. For example, a country with observation for 2010 but no observation for 2011 is counted for the period 2011. The procedure thereby excludes countries with relatively few MPI data points.

Then the HDI data—or, more precisely, HDS—are matched with the clustered MPI data. In the end, one gets a pseudo panel of 88 countries with three periods of HDS and MPI.

3. Results

The results of reliability test show a high degree of correlation for two-period pairings of HDS, with an average of 0.99 ($r_{1,2} = 0.992$ and $r_{2,3} = 0.996$; all $p < 0.001$). There are comparable results for the MPI, with an average of 0.96 ($r_{1,2} = 0.953$ and $r_{2,3} = 0.976$; both $p < 0.001$). The level of correlation for the period pairings appears to be stable. Further results on the steadiness of correlations are available in Table 1, which shows that the HDS-MPI correlation is at least 0.90. Overall, the results imply that both HDS and MPI are able to measure what they are designed to measure.

Figures 1, 2, and 3 present the data when sorted relative to the MPI. This presentation is just an attempt to remove the seemingly random character of the information when the values are

arranged alphabetically by country name. Notice, though, that there remain some "fluctuations" in the trends or there are points that show "opposite" directions between HDS and MPI. Nonetheless, one can still observe that the overall trend is positive, which present an initial evidence of convergent validity.

Figure 1: HDS and MPI, period 1

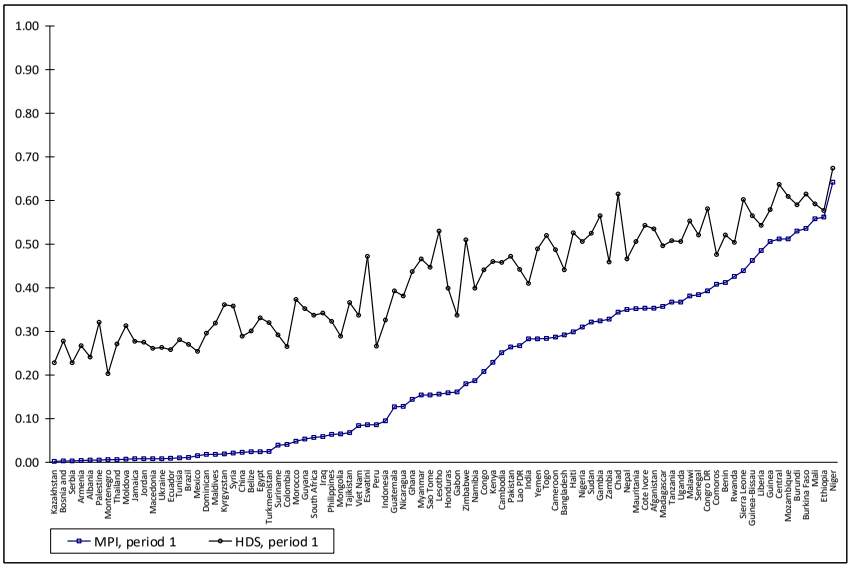

Figure 2: HDS and MPI, period 2

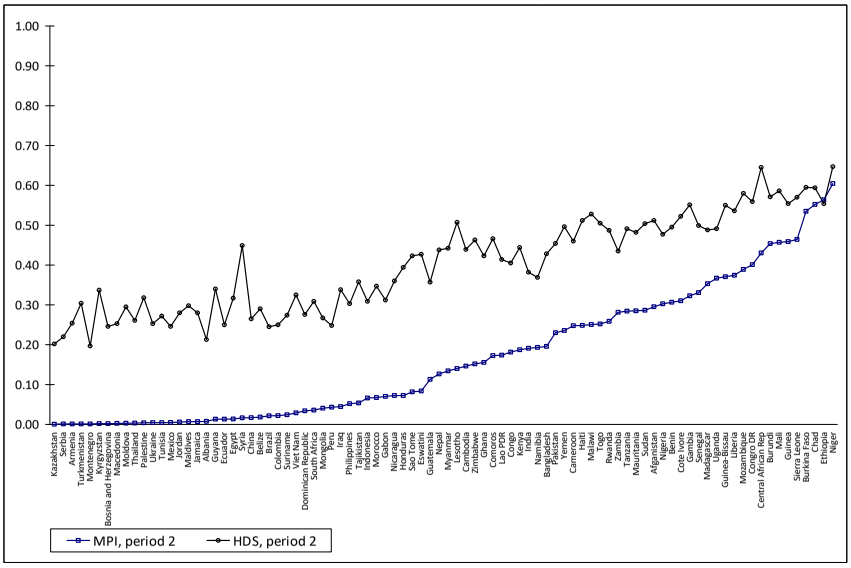

Figure 3: HDS and MPI, period 3

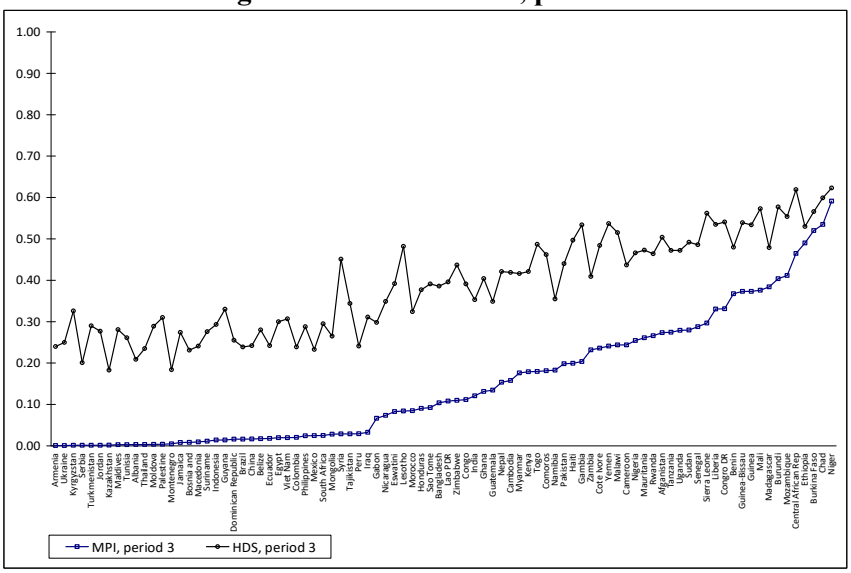

Table 1, in turn, presents a robust proof of convergent validity between the HDS and MPI. The results show rather high degree of correlations for the full data as well as for each period of the paired metrics. In addition, Table 1 presents results when the dataset is split into a "downward" or an "upward" segment, using the HDS as a reference for the identification given a more pronounced pattern relative to the MPI. For instance, a series like 0.30, 0.26, 0.20, 0.35, 0.22, 0.38, 0.40, etc. for the HDS is split as 1, 0, 0, 1, 0, 1, 1, etc., where 0 means a "downward" segment and 1 means an "upward" segment. This procedure is an alternative to the usual split data approach of the validity test to see if there is any significant difference in the degree of correlation between the split data. The results in Columns 3 and 4 suggest that the correlations are consistent with the results in Column 1—that is, Column 3 shows an average correlation of 0.94 ($p < 0.001$ for all period pairings); and Column 4, an average correlation of 0.93 ($p < 0.001$ for all period pairings). By inspection, these numbers are not significantly different from each other. On the whole, Table 1 presents strong evidence on the convergent validity between the HDS and MPI.

Table 1: Correlation analysis, full and split dataset

HDS with MPI	Full	Up	Down
Full	0.917	0.924	0.948
Period 1	0.928	0.964	0.931
Period 2	0.917	0.917	0.956
Period 3	0.905	0.934	0.915

Note: All results are statistically significant at $p < 0.001$. "Full" is the full dataset. For the split dataset, "Up" refers to the "upward segment" y; and "Down" refers to the "downward segment".

Lastly, Table 2 presents the results of regression analysis. The beta coefficient indicates that a

unit increase in the HDS (i.e., a decrease in HDI) means a 0.89 increase in the MPI, or vice versa. Though not enough can be said from Table 2 with respect to the causality between HDS and MPI (see, e.g., Beja 2019 on how causality might test for validity), Table 2 still presents enough evidence to confirm the initial findings on validity. Perhaps, from such estimate one can make sample predictions of HDS or MPI for, say, countries that were not included in the panel because of lack of observations. More interestingly, Table 2 points to equivalence in the information embodied in the HDS and MPI. That is, both metrics point out the degree of shortfall in human development and how large the remaining problem is for a country. Put another way, these metrics lead to the same interpretation about the level of deprivation in a country.

Table 2: Regression analysis

Variable	**Estimate**	**Beta**	**t-test**	**p-value**
Constant	-0.318		-13.886	< 0.001
HDS	1.247	0.890	23.959	< 0.001
Dummy = period 2	-0.014		-2.118	0.035
Dummy = period 3	-0.013		-2.325	0.021
Variance components				
Variable	**Estimate**		**t-test**	**p-value**
Residual	0.001		9.356	< 0.001
Intercept = country	0.003		5.659	< 0.001

Note: Dependent variable = MPI. R^2 = 83.8.

4. Conclusion

This paper tested the reliability and validity of the Human Development Index (HDI) from the UNDP and the Multidimensional Poverty Index (MPI) from the OPHI. For the analysis, the paper used the inverse of the HDI, namely HDS = 1-HDI, in order to focus the examination on metrics that shared the same perspective.

The results on the test-retest approach of reliability test showed not only a high degree of correlation for two-period pairings of the respective metrics but also a high level of steadiness in the level of their correlations. Such finding indicated that both HDS and MPI are able to measure what they were designed to measure.

Trends of both HDS and MPI were presented as initial evidence of their convergent validity. The results of the paired correlation analysis indeed confirmed a high degree convergent validity. Moreover, the results of the split data analysis validated the initial findings. Regression analysis in turn confirmed a positive relationship between the HDS and MPI, with an estimated coefficient of 0.89 for the countries covered in this study.

Lastly, the results from reliability and validity tests suggested that both indicators drew from the same information source. Overall, the findings indicated that not only do they embody similar information but also gave equivalent descriptions about the level of human development or the lack of it.

References

Alkire, A. and Santos, M. (2010). "Acute Multidimensional Poverty: A New Index for Developing Countries." OPHI Working Paper 38, Oxford: University of Oxford

Alkire, S., Kanagaratnam, U. and Suppa, N. (2020). "The Global Multidimensional Poverty Index (MPI) 2020," OPHI Methodological Notes 49, Oxford: University of Oxford

Alkire, S., Foster, J., Seth, S., Santos, E., Roche, J., and Ballon, P. (2015). *Multidimensional Poverty Measurement and Analysis*, Oxford: Oxford University Press

Anand, S. and Sen, A. (2003). "Human Development Index: Methodology and measurement," in S. Fukuda-Parr and A. Kumar (ed.), *Readings in Human Development* (pp. 114-127), Oxford: Oxford University Press

Beja, E. (2014). "Yet, Two More Revisions to the Human Development Index," *Forum for Social Economics*, 43(1), pp. 27-39

Beja, E. (2016). "Measuring Economic Ill-Being using Objective and Subjective Indicators: Evidence for the Philippines," *International Review of Applied Economics*, 30(2), pp. 151-166

Beja, E., (2019). "Consumer Expectations Survey and Quarterly Social Weather Survey: Evidence of Convergent Validity and Causality," *Philippine Statistician*, 8(2), pp. 131-147

Fukuda-Parr, S. and Kumar, A. (2003). *Readings in Human Development*, Oxford: Oxford University Press

Gregory, R. (2015). *Psychological Testing*, 7th ed., Essex: Pearson Education Ltd.

Netemeyer, R. Bearden, W., and Sharma, S. (2003). *Scaling Procedures*, London: Sage Publications

Ravallion, M. (2001). "Mushup Indices of Development," *World Bank Research Observer*, 27(1), pp. 1-32

Sen, A. (1979). "Poverty: An Ordinal Approach to Measurement," *Econometrica*, 44(2), pp. 219-231