



OPHI WORKING PAPER NO. 107

Measures of Human Development: Key Concepts and Properties

Sabina Alkire*

October 2016

Abstract

The measurement of human development is a rich field which has seen a veritable explosion of new and innovative international indices. With the advent of the Sustainable Development Goals' and their emphasis on interlinkages across deprived conditions, and with the promise of a 'data revolution' it seems that the proliferation of new indices that track multiple inter-related phenomena will continue. This paper sets out two basic set of criteria that, we believe, would be very helpful for structuring new policy metrics, and would provide human development experts, statisticians, and others with tremendously useful concepts to bear in mind when assessing and using different societal indices. The first criteria, ably articulated in Foster Seth Lokshin and Sajaia recently, is to clarify whether a human development indicator measures well-being, inequality, or poverty. These features of any population are distinct and each are of singular importance, but it is also useful not to confuse them. The second criterion is to ensure that the structure of the index is clearly explained. In particular, we discuss the importance of clarifying four methodological features: whether an index can be broken down by indicator; whether it can be disaggregated by population subgroup; whether it reflects the joint or overlapping conditions of a person, or evaluates dimensions one by one, and what kinds of weights or values are used to construct the composite index. While these criteria may seem, on the face of it, rather dry, a clear answer to each is essential (and also rather powerful) to understanding the policy relevance of each index.

Keywords: Human Development, Poverty, Inequality, Measurement, Social Progress

JEL classification: D63, I32, O15

* Oxford Poverty and Human Development Initiative (OPHI), University of Oxford. sabina.alkire@qeh.ox.ac.uk.

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

Acknowledgements

A first draft of this paper was prepared for the UNDP Latin American and Caribbean Bureau. I am very grateful to George Grey and UNDP colleagues, Milorad Kovacevic, Maria Emma Santos, Suman Seth, and Ana Vaz for insightful comments and corrections, to Claire Battye for research assistance and to Maarit Kivilo for formatting and finalization. All errors remain my own.

Funding: This paper was commissioned by UNDP Latin America Region.

Citation: Alkire, S. (2016). "Measures of Human Development: Key concepts and properties." *OPHI Working Paper 107*, University of Oxford.

The Oxford Poverty and Human Development Initiative (OPHI) is a research centre within the Oxford Department of International Development, Queen Elizabeth House, at the University of Oxford. Led by Sabina Alkire, OPHI aspires to build and advance a more systematic methodological and economic framework for reducing multidimensional poverty, grounded in people's experiences and values.

The copyright holder of this publication is Oxford Poverty and Human Development Initiative (OPHI). This publication will be published on OPHI website and will be archived in Oxford University Research Archive (ORA) as a Green Open Access publication. The author may submit this paper to other journals.

This publication is copyright, however it may be reproduced without fee for teaching or non-profit purposes, but not for resale. Formal permission is required for all such uses, and will normally be granted immediately. For copying in any other circumstances, or for re-use in other publications, or for translation or adaptation, prior written permission must be obtained from OPHI and may be subject to a fee.

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

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by OPHI or the University of Oxford, nor by the sponsors, of any of the views expressed.

1. Measures of Human Development: Expansion, Categories and Structure

Since UNDP's *Human Development Reports* were launched in 1990, the pioneering concepts of human development have moved dramatically from the margin to the mainstream. No longer is development success considered solely in terms of economic growth, although inclusive growth is rightly and vigorously pursued, particularly in the poorest countries.¹ No longer is the primary yardstick of human well-being GDP per capita; rather the limits of GDP as an indicator of well-being are recognized worldwide, including by the OECD and its member countries.² No longer is a money-metric definition of poverty considered to be sufficient; rather the Sustainable Development Goals urge nations to reduce poverty in “all its forms and dimensions”.³ And while economic inequality has justly acquired an unprecedented visibility, more and more actors are shining lights on inequalities in other spaces too.⁴ Alongside these topics, issues such as ecosystems, governance and democratic voice, gender parity, human rights, and freedom from violence have taken their rightful seat as necessary considerations in assessing societal progress.

Naturally, the tremendous shift in the concepts and objectives of development towards a human development approach is to be celebrated. Evidence of this shift is nowhere more visible than in the metrics and measures being proposed to assess human development. In 1990, when the Human Development Index was launched, it was original in content and in substance. Very few international indices of its kind existed, which could be used to compare most of the nations of the world, with a particular focus on developing countries.

Since that time there has been a veritable explosion of new and innovative international metrics of very different kinds. Drawn by the media interest in rankings and international comparisons, many institutions have been motivated to create international indices as a tool to attract attention from policy actors, to expose poor performance and to incentivize positive change. A small sample of these include:

¹ World Bank *Global Monitoring Report* 2015.

² http://ec.europa.eu/environment/beyond_gdp/index_en.html.

http://www.insee.fr/fr/publications-et-services/dossiers_web/stiglitz/doc-commission/RAPPORT_anglais.pdf.

³ This phrase is mentioned no less than 7 times in *Transforming our World* (2015), the document establishing the Sustainable Development Goals.

⁴ Dreze, J., and Sen, A., (2013) *An Uncertain Glory: India and Its Contradictions*, London, Allen Lane.

Human Development Index HDI⁵
 Multidimensional Poverty Index MPI⁶
 Inequality-Adjusted HDI⁷
 Ease of Doing Business Index⁸
 Legatum Prosperity Index⁹
 Global Peace Index GPI¹⁰
 Social Progress Index SPI¹¹
 Transparency International TI¹²
 Slavery Index¹³
 Physical Quality of Life Index PQLI¹⁴
 Aging Index¹⁵
 Your Better Life Index BLI¹⁶

Statisticians have long been suspicious of indices. Yet the recent scrutiny of indicators having pride-of-place in statistics, like GDP and \$1.90 poverty show these measures to exhibiting some of the same challenges of choice of components and weighting and aggregation mechanisms.¹⁷ So there is no avoiding indices; the task now is to scrutinize index methodologies and establish standards and guidelines.

This note sets out two basic set of criteria that, we believe, would help to order the chaos, and provide UNDP, statisticians, and others with tremendously useful features of different societal indices, to guide their evaluation and use. Note that this document focuses on indicators that primarily refer to the condition of human beings. Although these indices may incorporate variables related to other species, infrastructure, and institutional or environmental factors, this note focuses on indices related to human development.

⁵ HDRO at UNDP <http://hdr.undp.org/en>.

⁶ www.ophi.org.uk/ (all details) and national headlines on <http://hdr.undp.org/en/composite/MPI>.

⁷ HDRO at UNDP <http://hdr.undp.org/en/content/inequality-adjusted-hdi>.

⁸ <http://www.doingbusiness.org/>.

⁹ <http://www.li.com/activities/publications/2015-legatum-prosperity-index>.

¹⁰ <http://www.visionofhumanity.org/#/page/indexes/global-peace-index>.

¹¹ <http://www.socialprogressimperative.org/data/spi>.

¹² <https://www.transparency.org/>.

¹³ <http://www.globallslaveryindex.org/>.

¹⁴ Morris, 1979.

¹⁵ <http://www1.unece.org/stat/platform/display/AAI/Active+Ageing+Index+Home>.

¹⁶ <http://www.oecdbetterlifeindex.org/>.

¹⁷ Stiglitz Sen Fitoussi, 2009, Ferreira *et al.*, 2015.

The first set of criteria define: **What aspect of a given population does it measure: welfare, inequality, or poverty (size, spread, or base)?**¹⁸

Most indicators can be characterized as measuring one of the following: well-being, inequality, or poverty. This may seem to be a rather basic distinction. Yet it is misunderstood with surprising frequency. From a measurement perspective, measures of the three kinds are distinct, and the policies they can inform are also distinct. This note clarifies these distinctions.

The second set of criteria define: **What is the core structure of the index?** In particular, where does the index stand with respect to four methodological features: **dimensional analysis, decomposition** (or disaggregation), **linkages** across dimensions (joint distribution), and **weights**.

Let us elaborate each of these criteria in turn.

2. Kinds of Measures: Well-Being, Inequality and Poverty

Let us presume that we are designing a set of measures to provide information regarding a population in a given period. This could be the entire human population of a village, district, state, country, region, or continent. The study could also focus on some particular population sub-group such as children, older people, workers, youth, women, migrants, slaves, trafficked women, an ethnic minority, or urban dwellers.

Well-being One set of measures will show the well-being or *size* or overall level of *achievements* among this population. We might want to know the total wealth, the total income, total years of schooling, and so on which is enjoyed by every member of the population. In order to facilitate comparisons across groups, we usually take a population mean so that we can compare populations of different sizes. The most common average is the arithmetic mean or average. So we might measure the wealth per capita, income per capita, or the mean years of schooling of the population.

The current Human Development Index (HDI) is a classic measure of well-being, because it is based on GNI per capita, mean years of schooling, average school life expectancy, and average life expectancy. Each of these are averages of achievements across the entire national population. It is thus conceptually coherent as a well-being measure.

Inequality The mean level of achievements does not give us any information regarding the levels of inequality. Across three persons, the mean years of schooling could be six years, yet in one case, each of

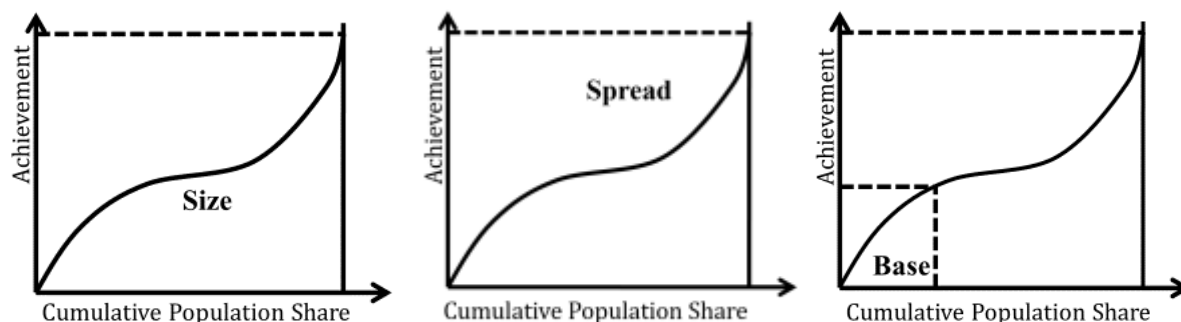
¹⁸ This distinction is established as the basic structure of measurement by Foster, J. E., Seth S., Lokshin, M., and Sajaia Z. (2013), and this paper draws upon their distinctions.

the three persons might have completed 6 years, whereas in the other case, two persons never attended school while the third completed 12 years of schooling plus a 4-year University degree plus a two year master’s degree, so the third person has all 18 years of schooling. The mean is the same; the distribution is very different.

Inequality measures make this difference visible. They reflect the **spread**, or range of achievements – of a distribution. Indices like the Gini index, the Atkinson index, the Theil 1 and 2, or the 90:10 ratio, all provide information regarding the levels of inequality in a society. For example, the Inequality-Adjusted HDI applies the Atkinson Inequality Measure to each of the dimensions of the HDI, and takes a general mean across them. All income inequality measures can be described as the ratio of some pair of income standards, where the income standards include mean, median, percentile incomes and Atkinson’s equally distributed equivalent income among others (Foster et al. 2013).

Poverty The third category of measure to describe the distribution of a population is a poverty measure. A poverty measure describes the **base** of the distribution – the people whose achievements are unacceptably low according to some standard(s). In a monetary poverty measure, for example, the distribution is censored at a poverty line, and the poverty measure aggregates information about the achievements of poor people into a society-wide measure.

It would be possible to have two distributions with the same level of welfare and inequality, yet very different levels of poverty. For a society it is also possible that both the levels of welfare and inequality improve over time, yet the situation of the poor deteriorates. For example, the level of welfare may improve over time due to increases in richer people’s incomes. And the level of inequality may improve due to a reduction in inequality between the rich and the middle-class. Yet the poor people may fail to enjoy any improvement or see a deterioration.



The figures above present, pictorially using quantile functions, the idea behind measuring well-being, which reflects the size of a distribution (how far right a distribution is?), measuring inequality, which reflects the spread of a distribution, and measuring poverty, which reflects the base of a distribution.

In the figure, we depict each person in the society. The height of the black curve shows their achievement in a variable such as income or years of education. People are ordered from the poorest persons, with the lowest achievements, on the left hand side, to the richest, on the right. The well-being measures will be a function of the area to the right of the curve. Inequality measures will reflect in some way the range of achievements (height) and distribution of achievements from poorest to richest. And poverty measures will reflect the situation of the people in the left hand portion whose achievements are below a cutoff, as shown.

Measures that summarize a distribution can be described with reference to the size, the spread, and the base of a distribution. Even measures that adjust for a second feature can still be classified. For example, Atkinson's Equally Distributed Equivalent Income (EDE, 1970), and following it, the IHDI, are inequality-adjusted well-being measures (Alkire and Foster 2010) – they measure size. The IHDI applies the Atkinson Inequality Measure to the original HDI, in order to obtain a very intuitive measure, which is the equivalent level of HDI that all persons actually enjoy if we apply some aversion to inequality, to recognise that the marginal contribution of an achievement unit to the wellbeing of the rich is lower than its contribution among the least well off. Another example is a relative poverty measure, which has a poverty cutoff and so reflects the base in some way. But because the poverty cutoff is relative to some feature of the distribution size, such as the median, it also reflects the spread or inequality of the distribution. A clear statement of the properties of the measures clarifies these features.

All Three Measure Types Needed: What are the implications of these distinctions? Sometimes we are asked, should we use the HDI or the Multidimensional Poverty Index (MPI) (Alkire and Santos 2010, HDR 2010)? This is a fascinating question, but it misunderstands these distinctions. The HDI is a well-being measure whereas the MPI is a poverty measure. The MPI complements the well-being indices. It is not necessary to choose between poverty and well-being measures. Just as countries normally measure income per capita, income inequality, and income poverty, so too the suite of human development measures have a measure of each type: a measure of well-being (such as the HDI), a measure of inequality (such as the IHDI), and a measure of Poverty (such as the MPI). All of them are used together to understand human development.

Increasingly, because human development also takes place on a shared planet, the variables that are included may also pertain to ecosystems and the natural environment. The literature underlying conceptually coherent combination of such measures is rapidly developing.

It may be worth observing that some composite indices cannot be categorized. The Social Progress Index is one of these. It mixes together indicators of poverty, well-being, inequality, and others. For example it includes deprivations like undernourishment, child mortality, lack of access to improved water and sanitation, and early marriage which are clearly ‘poverty’-related indicators in a measurement sense, because they reflect the base of their respective distributions. It also includes society-wide well-being indicators such as political rights, women’s average years in school, life expectancy, mobile phone subscriptions and internet users. It also includes inequality measures: inequality in educational attainments. The SPI is also not restricted to indicators of well-being, poverty or inequality of a given population, as it also includes data on people who have perished (from infectious diseases or air pollution or suicide or non-communicable diseases). The internal diversity of indicators creates significant challenges in interpretation.

3. Core Structural Features of Indices for Human Development

Having defined kinds of measures, we now turn to examine the **core structure of indices**. Given that the motivation of the Human development family of indices is to guide and structure policies to advance human development, it can be useful to understand how certain structural features of indices can advance this aim. For Human development in the context of the Sustainable Development Goals, four criteria are particularly important:

1. **dimensional analysis**;
2. **decomposition** (or disaggregation);
3. **linkages** across dimensions (joint distribution); and
4. **weights** and robustness

Dimensional Analysis. The 1990 Human Development Index results were electrifying because, using simple descriptive information on education and health in addition to income, they showed that national mean achievements in non-income indicators did not necessarily match income achievements. Rather, they showed that countries gained or lost rank in the HDI in comparison with the GDP per capita rankings. Furthermore, if a country had lost rank, policy actors could look across the row of the HDI table at its dimensional achievements to see which dimension was responsible for the low performance. At the time, this comparison of dimensional achievements in health, education, and income was tremendously new and useful.

Subsequently, the field of variables used to chart human development outcomes has expanded greatly, to include topics such as political rights, physical security, freedom of action, productivity, gender parity, community belonging, cultural liberty, sustainability, and human rights (Alkire 2010). Global and

National Human Development Reports have often examined the distribution of these achievements across countries, and found that ‘not all good things go together’ (*HDR 2010*), which again underscores the importance of considering multiple dimensions of human development.

The Human Development Index, pioneering in its day, now appears quite limited in terms of its dimensional coverage because it only has three dimensions. The Stiglitz Sen Fitoussi Commission (2009), for example, identified not three but eight dimensions of Quality of Life that, authors argued, should be considered in its measurement. Further, many of the indices mentioned above, like the Global Peace Index, OECD’s Better Life Index, and the Social Progress Index use an expanded range of dimensions in their well-being indices. However, it is important to note that only some of these measures offer direct dimensional breakdown, so that countries can readily ascertain what indicators are lagging and which are satisfactory, *and* trust that improvements in the component indicators will improve the overall index. The Global MPI for example satisfies the principle of dimensional breakdown according to its component indicators, and if any MPI deprivation of any poor person in any indicator is removed, poverty is reduced.¹⁹ OECD’s Better life index has such a feature, but the Social Progress Index does not.

Disaggregation. In 1990, the Human Development Index was ground-breaking because it provided an alternative national headline to GDP per capita. However, the Sustainable Development Goals have called for disaggregation of many measures by relevant variables, which might include gender, age group, ethnicity, rural-urban area, geographic region, and disability. Disaggregation enables policymakers to see intra-country variations. It enables state or local policy makers to have relevant information to shape their responses. And to fulfil the SDG goal of ‘Leaving No One Behind’ it is necessary to be able to define who the poorest of the poor are, and how their situation is changing. Compare the maps below. Figure 1 shows the level of Poverty for 102 countries, whereas Figure 2 breaks down most countries into 962 subnational regions (Alkire Jindra Robles and Vaz 2016) and where one can see that several subregions in each country are poorer as well as less poor than the average. It is this kind of disaggregation that tremendously aids policy design to expand human development.

Because the Global MPI satisfies the property of subgroup decomposability, it can be rigorously disaggregated by any subgroup for which the data are representative. At present decompositions for 1468 subnational regions have been done; for example all Low Income Countries have disaggregated MPI poverty data as do 41 countries in Africa.

¹⁹ Alkire and Foster (2016) articulate the importance of dimensional breakdown. But a number of indicators do not satisfy dimensional breakdown. For example, the Social Progress Index uses factor analysis to aggregate outcome indicators, but because the weights generated by factor analysis are endogenous, there is no necessary linkage between the improvements in any one indicator and the overall SPI, and it does not satisfy dimensional breakdown.

The Human Development Index in theory can be computed for any group or population of any size who have the required variables. However in practice, life expectancy is not widely available for subnational groups. So where local or regional HDIs are produced, they tend to use other variables to indicate health. The Global Peace Index has been computed subnationally for three countries to date. Indices that combine many component indicators from different data sources face a larger challenge in disaggregation, because not all indicators are available for all regions.

Figure 1. Multidimensional Poverty by Country²⁰

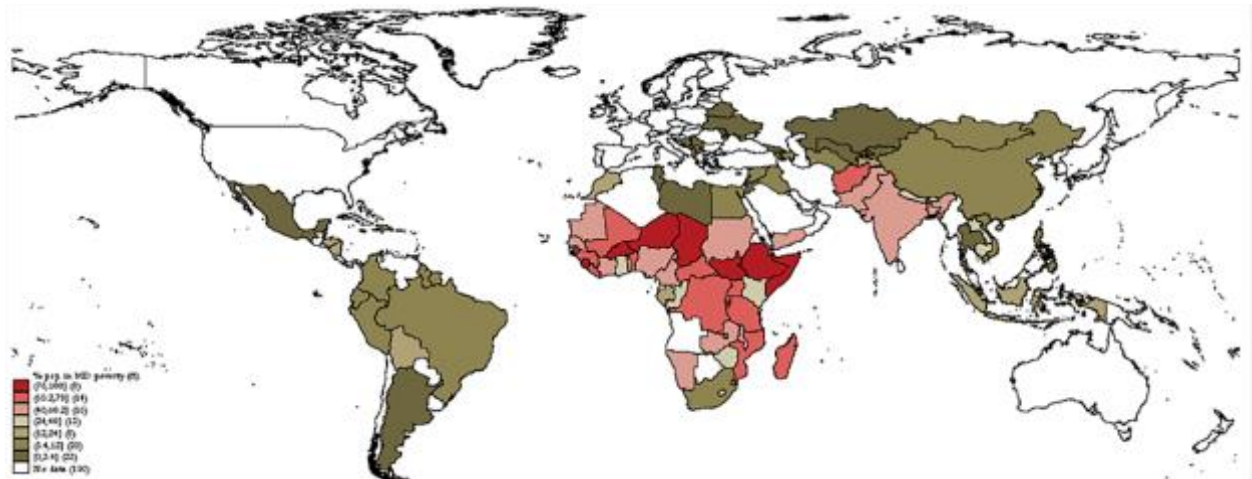
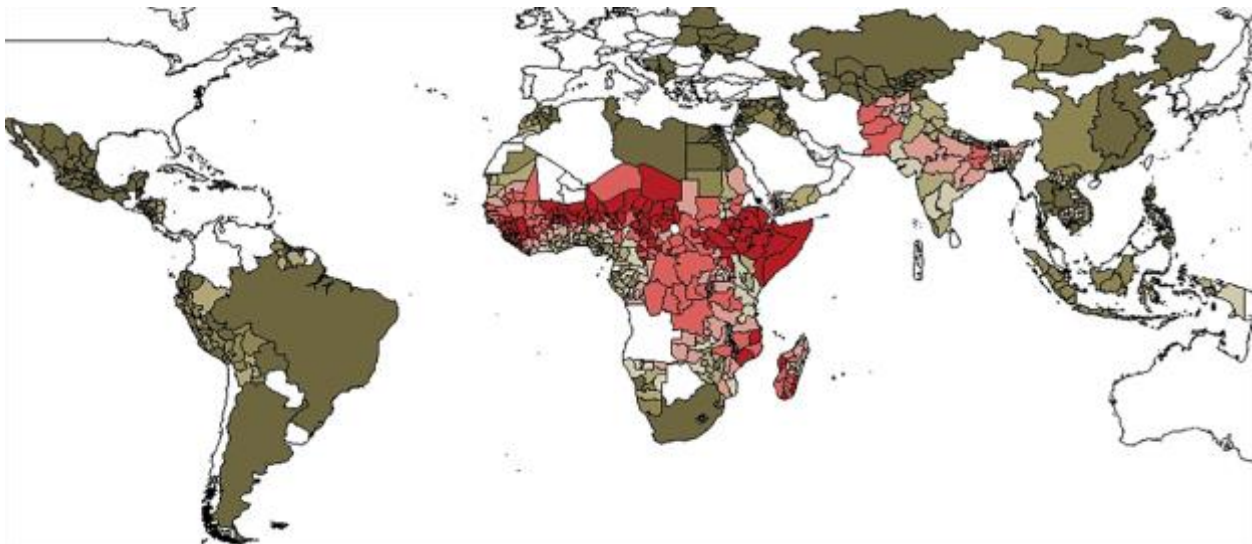


Figure 2. Multidimensional Poverty by Subnational Region (Data Permitting)



The Human Development Index in theory can be computed for any group or population of any size who have the required variables. However in practice, life expectancy is not widely available for

²⁰ Both maps depict the MPI according to Alkire, Jindra, Robles and Vaz (2016).

subnational groups. So where local or regional HDIs are produced, they tend to use other variables to indicate health. The Global Peace Index has been computed subnationally for three countries to date. Indices that combine many component indicators from different data sources face a larger challenge in disaggregation, because not all indicators are available for all regions.

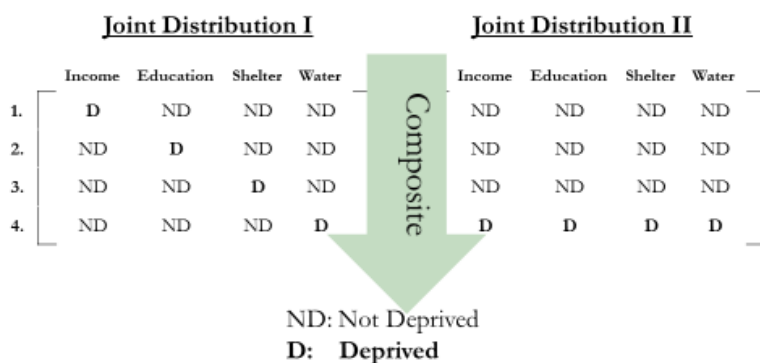
Linkages across Dimensions. A further feature of Human Development which has been emphasized vigorously is the need for measures and policies to look across different dimensions. The dimensions of poverty and of well-being are interlinked, and their realization requires an integrated policy response. This is stated in the Preamble of the SDG document *Transforming our World* (2015), which says, “The interlinkages and integrated nature of the Sustainable Development Goals are of crucial importance in ensuring that the purpose of the new Agenda is realised.” In his report to the UN Economic and Social Council in July 2015, the Secretary General of the United Nations hoped that the SDGs would “inject new impetus for embracing integrated approaches to development.” Beyond the evident ethical appeal of reducing multiple deprivations, another reason for this is cost effectiveness. When integration is lacking, synergies are not realized – as the prominent 50-country UNDP report famously proved (2010). When policy coherence is evident, progress accelerates.

Multidimensional indices that bring together indicators from different dimensions of development support more of a joined up view. However, there are two very different types of indices, and the information they provide is fundamentally distinct. Composite indices offer only an indirect understanding of linkages whereas counting-based multidimensional measures transparently map linkages with a very high resolution.

In composite indices, such as the HDI, GPI, SPI, and OECD’s BLI, different indicators (which may come from different surveys) are selected, that are pertinent to the concept at hand. These are aggregated first across the population into some overall score such as the mean years of schooling, child

mortality rates. Each indicator contains some information regarding a given population (a nation, or region). The indicators may be treated or normalized, and then are aggregated into a composite index. As we discussed above in the case of the HDI, the aggregate index will combine achievements in different indicators to provide an overall statistic which may

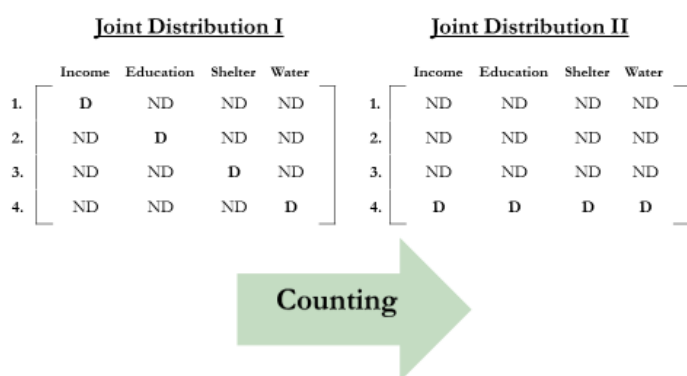
Composite Indices: Order of Aggregation



rank countries or regions quite differently than its component indicators. So these indicators provide some idea of whether ‘all good things go together’. But they do not investigate at the individual or household level, precisely which achievements are enjoyed by which persons. For example, in each of the two societies depicted above, 25% of people are deprived in income, education, shelter and water. But in society I, 25% of the citizens experience exactly one of these deprivations at the same time. In Society II, all four deprivations are experienced by the *same* 25% of the population. Composite indicators cannot distinguish between these two societies because they only use the column totals, which are 25% in each case.

In contrast, a counting-based index like Bhutan’s Gross National Happiness (GNH) Index (Ura et al 2015) or the Multidimensional Poverty Index (MPI) begins at the level of the individual or household, and builds a profile of their attainments or deprivations across indicators for each person. From this profile, each person or household is identified as poor (or as happy). The MPI, which is an overall aggregate multidimensional measure, can be broken back down to show which people experience which

Counting Measures (MPI) Order of Aggregation



deprivations at the same time. Combining this strong and transparent mapping of linkages between dimensions, with the disaggregation property mentioned above, means that if the sample design permits, one can analyse the MPI poverty profiles by region, by ethnicity, by household size, by gender, by age profile, by education or gender of the household head, and so on. So for example, in the panel to the left, we see that counting horizontally, in society I, each person is deprived in 25% of

the indicators, whereas in society II, person 4 is deprived in 100% of the indicators, and persons 1-3 in zero. Counting measures make use of this information as their basic information source.

Weights and Robustness. Both composite indices and counting-based indices aggregate across dimensions using some weighting vector. These weights are often perceived to be ‘a difficulty’ in composite measures, because they seem likely to be contentious and difficult to defend (Alkire Foster Santos 2011). Thus a policy-oriented measure should be robust to a range of plausible weights and the actual weights used should be made explicit, transparent and easy to understand.

There are a number of important observations that have been made in the academic literature. Two are of fundamental importance to policy users of the measures as well. The second of these two points is slightly technical, but of tremendous importance.

The first observation is that it is not necessary for people to agree on one precise set of weights, in order to use multidimensional measures. What is required is that multidimensional measures must be built such that their key policy recommendations are robust to a range of weights. Why? People are diverse, and values may differ across people and for the same person over time. Hence the relative valuations people place on different indicators of poverty or well-being normally differ somewhat. While this is often perceived to be a fundamental problem, it has a technical response: we must explore the robustness of indices to a range of plausible weights. Amartya Sen writes, ‘It can, of course, be the case that the agreement that emerges on the weights to be used may be far from total’, but continues, ‘we shall then have some good reason to use ranges of weights on which we may find some agreement. This need not fatally disrupt evaluation of injustice or the making of public policy.... A broad range of not fully congruent weights could yield rather similar principal guidelines’ (2009, p 243). It is not guaranteed that all sets of weights will be broadly congruent, and the robustness has to be confirmed empirically in each case. The Global MPI robustness tests in 2015, for example, show that if the weight on each dimension varies freely from 25% to 50%, 84% of pairwise comparisons of countries are unchanged, considering the standard errors of the MPI values and requiring comparisons to be strict (which is the most demanding situation).

The second observation is that the nature of the weights in the composite vs counting-based multidimensional measures are fundamentally different. This may seem a technical point, but it really matters, so is worth grasping.²¹

In measures like the MPI, ‘weights’ across component indicators are simply turned on if a deprivation is present, and turned off if it is not. This is because a deprivation is defined (unsafe water, inadequate sanitation, housing), often using ordinal or ordered categorical data, and the original data are dichotomised. If the deprivation is present, weights show the relative value of that deprivation – how much it increases the person’s deprivation score by. But the key thing is that there are only two situations: on, or off – deprived, or non-deprived.

²¹ Klugman (2011) and Ravallion (2012) debate the weights of the HDI, and in doing so clarify many of the assumptions about the precision weights and their particular implications in the HDI’s current functional form. Alkire Foster Seth Santos Roche and Ballon (2015) make the distinction that is being drawn here, in simpler language, between ‘precision’ weights and deprivation values and clarify the nature of weights (deprivation values) for the MPI.

In most composite indices such as the HDI, SPI, BLI and so on, the weights are much more complex. We call them ‘precision’ weights (in Alkire et al 2015). They fix the marginal rates of substitution across dimensions, and also calibrate the impact of changes at different levels of achievement in the same dimension on the overall measure.

To understand this further, we will compare how weights function in the least complex composite indicator which is OECD’s Better Life Index, and in the MPI.

The Better Life Index has 11 domains, and 24 indicators. To make the index, each indicator is normalized to lie between zero and 1. Then weights are selected for each of the 11 domains. There are one to four indicators within each domain, and each indicator receives an equal weight. For simplicity let us assume that the eleven domains are roughly equal in value, so apply equal weights.

Domain	Housing			Income		Jobs				Comm- unity	Education			Environ- ment		Civic Engage- ment		Health		Life Satis- faction		Safety		Work Life Balance
Domain weight	$\frac{1}{11}$			$\frac{1}{11}$		$\frac{1}{11}$				$\frac{1}{11}$	$\frac{1}{11}$			$\frac{1}{11}$		$\frac{1}{11}$		$\frac{1}{11}$		$\frac{1}{11}$		$\frac{1}{11}$		
Indicator weight	$\frac{1}{33}$	$\frac{1}{33}$	$\frac{1}{33}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{44}$	$\frac{1}{44}$	$\frac{1}{44}$	$\frac{1}{44}$	$\frac{1}{11}$	$\frac{1}{33}$	$\frac{1}{33}$	$\frac{1}{33}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{11}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$	$\frac{1}{22}$
Score (Example)	.92	.88	.75	.90	.43	.37	.76	.99	.32	.86	.96	.87	.78	.64	.57	.88	.95	.89	.73	.72	.84	.48	.67	.73

Imagine a country *Aadvardia* which has the normalized scores shown in the third row (OECD does not publish scores).²² The BLI (as in the pre-2010 HDI) is a weighted sum of its normalized scores, so the weights are multiplied by the 24 normalized scores, then summed to obtain the BLI value. Notice that here the weights are applied not to value of 0 or 1 (as they are in the MPI, we shall see), but to normalized scores that take any numeric value in that range.

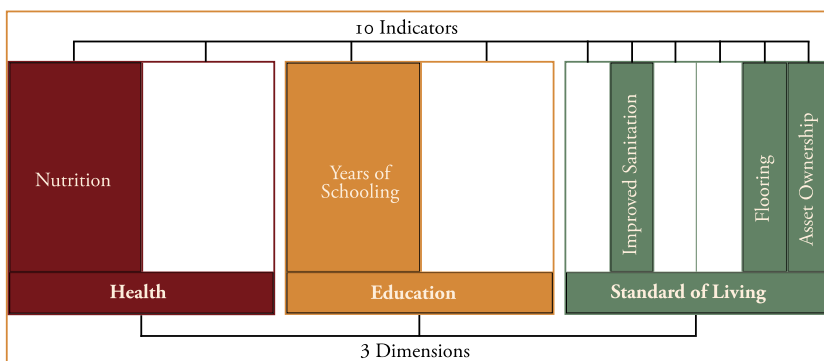
Let’s say that the first indicator, 0.92 in housing, changes to 0.96. The BLI will increase. However, we could increase the BLI by that same amount by increasing housing expenditure instead. But how much? We would have to work back through the normalization procedure of each indicator to find this out. So we can obtain marginal rates of substitution for each possible pair of indicators. And we must scrutinize these tradeoffs to ensure they are justifiable, and not ‘troubling’ as Martin Ravallion titled a paper on this topic (2012). However in practice the implications that weights have for policy incentives in the form of tradeoffs across indicators are often not articulated, nor justified, nor tested for robustness.

²² Interestingly the HDI tables from 1990 to 1998 did not publish HDI’s dimensional index scores. From 1999-2008 they did. From 2009 they only publish the dimensional indicators, which is more intuitive and direct, but obscures the tradeoffs introduced through the normalization process, so de-emphasises the composite HDI, perhaps correctly.

For these kinds of composite indices and their marginal rates of substitution to create meaningful measures of well-being, the normalized scores involved in these kinds of indices should be ratio scale (Alkire and Foster 2010, Alkire Foster Seth Santos Roche and Ballon 2015, Ch. 2). A ratio scale variable (like Kelvin temperature) has a ‘natural zero’ which is the same for all variables in the space of well-being. Also, a one-point difference on the scale has the same well-being effect no matter the level of achievement. Ideally the normalization step creates this. In practice, normalized scores are rarely justified as ratio scale variables, which makes the interpretation of the final indices yet more challenging.

Let’s contrast the role of weights in composite indices with their role in counting-based indices using dichotomised data, such as the MPI. The MPI begins by asking whether each person is deprived or not according to some agreed standards. As the global MPI was launched during the Millennium Development Goals, most deprivation cutoffs reflect those goals. For example: are they malnourished by WHO standards? Has a child died? Do they have clean water by MDG definitions? Adequate sanitation? Clean Cooking Fuel? And so on.

By answering these questions, a deprivation profile is developed for each person. For example, Olga is deprived in the shaded indicators below. We could think of her as having a row of tick marks for her deprivations, and circles in indicators in which she’s not deprived:



Tick=Deprived
0 = not Depr.

1 0 1 0 0 1 0 0 1 1 **Deprivations**

$\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{18}$ $\frac{1}{18}$ $\frac{1}{18}$ $\frac{1}{18}$ $\frac{1}{18}$ $\frac{1}{18}$ **Weights**

$\frac{1}{6}$ 0 $\frac{1}{6}$ 0 0 $\frac{1}{18}$ 0 0 $\frac{1}{18}$ $\frac{1}{18}$ **Profile (50%)**

We write the ticks – signifying deprivations – as ones, and the non-deprivations as zeros, as in the row labelled ‘Deprivations’. The indicator weights are one-sixth for two dimensions and one-eighteenth in the third dimension. You can see the weights visually by looking at the length of the cells in the deprivation profile. You multiply the deprivations by the weights to obtain Olga’s deprivation profile (last line), which shows that she is deprived in 50% of the weighted indicators. Note that the indicators of sanitation and flooring both have the same weight. In the MPI, this means the MPI will change by precisely the same amount whether Olga gets clean sanitation or flooring – the policy incentives are transparent. If you disagree with those weights, you can change them – the point is, what you see is what you get. Robustness tests are routinely applied to explore how the policy relevant analyses change when applying a plausible range of weights (Alkire and Santos 2014).

This last section has explained one last consideration in assessing indices. It is that measures that should be used for policy need to make their weights explicit, and should be robust to a plausible range of weights. We observed that weights have different implications a) if they are applied to a 0-1 vector of deprivations, vs b) when they are applied to a variable that takes any numbers within some range, such as a normalized index. In the first (a), weights (or deprivation values) are transparent, easy to understand and think about. The second case (b), involving what might be called ‘precision’ weights, requires analysts to interpret and justify the tradeoffs between an indicator itself at different levels of attainment, and tradeoffs between every indicator in the index – (which should be ratio scale). Precision weights are more complicated to set, interpret and justify, but this task must be taken on squarely when composite indices are released, in order to make the policy incentives implicit in composite indices transparent, and to assess the sensitivity of the final results to changes in this parameter.

Naturally, while the four considerations mentioned in this section seem to have a particular relevance in linking human development indicators to policy, the additional properties those indicators satisfy should also be clarified, and the applications of those indicators should also follow standard good practices such as reporting standard errors, and assessing statistical significance of results.

4. Conclusion

The tremendous shift in the concepts and objectives of development towards a human development approach is to be celebrated. Evidence of this shift is nowhere more visible than in the explosion of metrics and measures being proposed to assess human development.

This note sets out two basic set of criteria to order the chaos, and provide UNDP, statisticians, and others with tremendously useful features of different societal indices, to guide their evaluation and use. The first set of criteria define: What aspect of a given population does it measure: welfare, inequality, or poverty (size, spread, or base)?

Most indicators can be characterized as measuring one of the following: well-being, inequality, or poverty. This may seem to be a rather basic distinction. Yet it is misunderstood with surprising frequency. From a measurement perspective, measures of the three kinds are and should be distinct, and the policies they can inform are also distinct.

This note also clarifies that all three measure types are needed. Just as countries normally measure income per capita, income inequality, and income poverty, so too the suite of human development measures have a measure of each type: a measure of well-being (such as the HDI), a measure of inequality (such as the IHDI), and a measure of Poverty (such as the MPI). All of them bring distinctive insights, and so should be used together to understand human development. Some composite indices, we noted, blend different measures, which creates challenges in interpretation.

The second set of criteria define the core structure of the index. In particular, it is important to clarify precisely where each index stands with respect to four methodological features: dimensional analysis, decomposition (or disaggregation), linkages across dimensions (joint distribution), and robustness to a plausible range of weights. The motivation of the Human development family of indices is to guide and structure policies to advance human development, and each of these four features contributes to the kinds of descriptive policies that are in use.

References

- Alkire, S. (2010). "Human development: Definitions, critiques, and related concepts." *OPHI Working Paper 36*, University of Oxford.
- Alkire, S. and Foster, J. (2010). "Designing the Inequality Adjusted Human Development Index (IHDI)". *OPHI Working Paper 37*, University of Oxford.
- Alkire, S. and Foster, J. (2011a). "Counting and multidimensional poverty measurement". *Journal of Public Economics*, 95(7–8): 476–487.
- Alkire, S. and Foster, J. (2011b). "Understandings and misunderstandings of multidimensional poverty measurement". *Journal of Economic Inequality*, 9: 289–314.
- Alkire, S., Foster, J., and Santos, M. E. (2011). "Where did identification go?" *Journal of Economic Inequality*, 9(3): 501–505.
- Alkire, S., Foster, J., Seth, S., Santos, M. E., Roche, J. M., Ballon, P. (2015). *Multidimensional Poverty Measurement and Analysis*. Oxford University Press.
- Alkire, S., Jindra, C., Robles, G., Vaz, A. (2016). "Multidimensional Poverty Index 2016: Brief methodological note and results." *OPHI Briefing 42*, University of Oxford.
- Alkire, S. and Santos, M. E. (2014). "Measuring acute poverty in the developing world: Robustness and scope of the Multidimensional Poverty Index". *World Development*, 59: 251–274.
- Alkire, S., Santos, M. E., Seth, S., and Yalonetzky, G. (2010). "Is the Multidimensional Poverty Index robust to different weights?" *OPHI Research in Progress 22a*, University of Oxford.
- Atkinson, A. B. (1970). "On the measurement of inequality". *Journal of Economic Theory*, 2(3): 244–263.
- Ferreira, F. H. G., Chen, S., Dabalen, A., Dikhanov, Y. M., Hamadeh, N., Jolliffe, D. M., Narayan, A., Prydz, E. B., Revenga, A. L., Sangraula, P., and Serajuddin, U. (2015). "A global count of the extreme poor in 2012: Data issues, methodology and initial results." *Policy Research Working Paper Series 7432*, The World Bank.
- Foster, J. E., Seth S., Lokshin, M., and Sajaia Z. (2013). *A Unified Approach to Measuring Poverty and Inequality: Theory and Practice*. The World Bank.
- Klugman, J., Rodriguez, F., and Choi, H. (2011). "The HDI 2010: New controversies, old critiques". *Journal of Economic Inequality*, 9:249–288
- Morris, D. (1979). *Measuring the Condition of the World's Poor: The Physical Quality of Life Index*. New York, Pergamon.
- Ravallion, M. (2012). "Troubling tradeoffs in the Human Development Index". *Journal of Development Economics*, 99(2): 201–209.
- Santos, M. E., Villatoro, P., Mancero, X. and Gerstenfeld, P. (2015). "A Multidimensional Poverty Index for Latin America." *OPHI Working Paper 79*, University of Oxford.
- Sen, A. K. (2009). *The Idea of Justice*. Penguin.
- Stiglitz, J. E., Sen, A., and Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress*. www.stiglitz-sen-fitoussi.fr.
- UNDP. (2010). *What Will it Take to Achieve the Millennium Development Goals? An International Assessment*. United Nations Development Programme (UNDP), United Nations.
- UNDP. *Human Development Reports* (various years).
- UN. (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. A/RES/70/1.
- World Bank. (2015) *Global Monitoring Report*. The World Bank, Washington DC.