AN AUTOMATIC-DEMOCRATIC APPROACH TO WEIGHT SETTING FOR THE NEW HUMAN DEVELOPMENT INDEX

Dr. Chris Tofallis

University of Hertfordshire Business School Working Paper (2012)

The final version is accepted to appear in *Journal of Population Economics*: www.springerlink.com

AN AUTOMATIC-DEMOCRATIC APPROACH TO WEIGHT SETTING FOR THE NEW HUMAN DEVELOPMENT INDEX

Dr. Chris Tofallis

Statistical Services and Consultancy Unit
The Business School
University of Hertfordshire
College Lane
Hatfield
Herts
AL10 9AB
United Kingdom

Tel. +44 1707285486
Email: c.tofallis@herts.ac.uk

ABSTRACT

Perhaps the most difficult aspect of constructing a multi-dimensional index is that of choosing weights for the components. This problem is often bypassed by adopting the 'agnostic' option of equal weights, as in the Human Development Index. This is an annual ranking of countries produced by the United Nations Development Programme based on life expectancy, education, and per capita gross national income. These three dimensions are now aggregated multiplicatively. Whatever weights (exponents) are chosen for these dimensions, some nations will feel disadvantaged. To avoid the use of arbitrary weights we propose for consideration a two-step approach: (1) Find the most advantageous set of weights for each nation in turn. (2) Regress the associated optimal scores on the underlying indicators to find a single weight set. This approach has the properties of non-subjectivity, fairness, and convenience. The result is that the highest weight is placed on the life expectancy dimension.

Keywords: Human Development Index, index construction, social composite index, aggregation, ranking, data envelopment analysis.

1. INTRODUCTION

The United Nations Development Programme (UNDP) publishes its Human Development Report annually. This is ‘widely considered the most influential of the many regular reports by multilateral institutions. Unique among UN publications for their tradition of intellectual independence – though sponsored by UNDP, they do not represent its official views or policies’ (OPHI, 2010). Mahbub-ul-Haq (1934-1998), one time Chief Economist of Pakistan’s Planning Commission, led the team to produce the first UN Human Development Reports starting in1990. His aim was to “to shift the focus of development economics from national income accounting to people-centred policies’’, and asked his friend from student days, Amartya Sen, to assist in developing an HDI which was not purely focused on economic progress, but also on other advances in well-being. Thus one of the original aims was to depart from one-dimensional rankings based on Gross Domestic Product or income per capita, and to embark on using a more broadly based composite indicator. Sen was opposed to a single figure index because it could not capture the complexity of all the various aspects. However Haq succeeded in persuading him by arguing that only a single figure could catch the attention of policy-makers, the media, and politicians.

Chapter 1 of the first Human development report summarises the underlying philosophy thus:

> ‘The basic objective of development is to create an enabling environment for people to live long, healthy and creative lives. This may appear to be a simple truth. But it is often forgotten in the immediate concern with the accumulation of commodities and financial wealth…. Human development is a process of enlarging people’s choices. The most critical ones are to lead a long and healthy life, to be educated and to enjoy a decent standard of living.’ (UNDP, 1990)

Thus, ever since then, the reports have contained a ranking of nations based on a three component index known as the Human Development Index or HDI.

The HDI is important and influential for a number of reasons – so much so that a recent paper was entitled *The Tyranny of International Index Rankings* (Hoyland et al, 2012):

> ‘A particularly favorable or unfavorable position is likely to be widely noticed, and governments stand to lose by not commenting upon them. To attack an index is

> never appropriate for politicians. When the ranking is unfavorable, an attack would just make things worse; when the ranking is favorable, the praise is too tempting. In Norway, for instance, leading politicians regularly insist that the United Nations has chosen Norway as the best country to live in based on its position in the Human Development Index (at least prior to 2007)'.

One reason that the HDI matters is that it influences governments and non-governmental organisations when they are considering how to allocate foreign aid (O'Neill, 2005). The HDI is also used in price setting in different countries. For example, Merck Pharmaceuticals bases its local prices for AIDS drugs on a country's HDI category: It gives discounts of 90% to the lowest HDI group, and 75% discounts to the middle group (Bate and Boateng, 2007). HDI is also used in studies which try to find out what helps promote the development of nations. For example Reiter and Steensma (2010) looked at the effect of foreign direct investment (FDI) using regression analysis. They found that the benefits from FDI were greatest when a country's FDI policy restricted foreign investors from entering some economic sectors: 'When foreign investors are allowed free access, the effect it has on human development depends on whether or not foreign investors' objectives align with those that will promote human development, which would be merely fortuitous.' It was hypothesized that FDI was found to be most beneficial when it was directed towards those sectors where foreign expertise was needed, and that left to its own devices FDI could drive out local firms. They also found that the beneficial effect on human development of FDI was reduced in those countries where corruption was high. Also on the theme of FDI and human welfare, Gohou and Soumare (2011) looked at the differential effect of foreign direct investment across different African countries. They found that there was a statistically significant effect on the poorer countries of central and east Africa but that this was not the case for the less poor nations in northern and southern Africa.

Of course, the HDI is an average for a given population; there will be differences in welfare *within* a given country, and so there is much interest in adapting the human development index so that it includes adjustments for inequality (Hicks, 1997). Grimm et al 2008 studied the HDI for different quintiles in the income distribution to enable comparisons between poor and non-poor within and across countries. The 2010 UN Development Report has for the first time introduced an inequality-adjusted HDI, which is calculated for 139 nations. It takes into account inequalities in income, education, and health to reduce the indices for these dimensions, and thereby to give a reduced overall HDI figure. The results show that sub-Saharan African scores fall the most - due to extensive inequality across all three dimensions (UNDP, 2010, p.7). Global inequality has also been studied using the HDI by looking at inter-country differences in the index as a whole as well as the separate components of life expectancy, education, and standard of living (McGillivray and Markova, 2010). The results showed that inequalities in life expectancy have increased over the period studied, but that there have been reductions in inequality for the other variables.

Human development indices are also used at a sub-national level to allocate support to regions where it is most needed. For example, since 1997 Brazil has been publishing its *Atlas of Human Development* which displays development levels by color coding in thousands of municipalities. This is used by the Alvorada Program to allocate funding for poverty reduction. There is an interesting parallel here with the introduction of the HDI as a shift away from purely economic measures, because the Alvorada Program was

> ‘developed in response to criticism of the government for focusing too intensely on economic conditions and not devoting sufficient attention to social issues. It provides various kinds of support… educational programs (adult literacy programs, remedial courses for students, scholarships for poor families); providing safe water and sanitation for 16,000 schools and

> more than 1.3 million families; establishing 6,000 new health care teams to serve an estimated 31 million people; and energy services programs' (Henninger and Snel, 2002, p.18)

Brazil has indeed embraced the HDI and has made efforts to raise awareness in society by incorporating it into the secondary school curriculum and in university entrance exams (Henninger and Snel, 2002, p.29)

The UNDP has resisted the temptation to stick to the original methodology and has in fact been open to new ideas regarding the way the HDI is calculated. As a result there have been numerous adjustments over the years. Indeed Morse (2003) used these methodological variations as a way of demonstrating the volatility of country rank positions. This was done using a set of 114 countries which was common to the 12 year analysis, and so avoided the issue of rank shifts caused by nations entering or being excluded over this period. The results were presented as a set of deviations from what the ranks would have been if a different UNDP methodology had been used. Morse found that 'the volatility that can result from such recalculation is shown to be substantial (10-15 ranks), yet reports in the popular press are frequently sensitive to movements of only a few ranks. Such movement can easily be accounted for by changes in the HDI methodology rather than genuine progress in human development'. The paper also quotes a number of lively newspaper pieces which speak of rank changes over time, such as 'We're not No. 1! Canada drops in UN rankings. Seven-year reign at top is coming to an end' - The National Post (Canada; July 3, 2001). Another was: 'Kudos for India in Human Development Report. . . India has 'moved up four notches' giving enough reason for satisfaction. - The Hindu (India; 30 June 2000). It is not always clear to what extent such rank changes over time are due to alterations in the methodology, and it is for this reason that the recent UNDP reports also provide historical trend information using their current method of calculation across a number of years.

The most significant change in methodology occurred in November 2010 and we turn to that in the next section. This is followed by a discussion of the weights used in the HDI. Our objective in this paper is to present a way of deriving weights from the data such that each country influences the final result. Section 4 explains how the technique of data envelopment analysis can be used to provide weights which are specific for each country. The resulting scores are then used in a regression to generate a single set of weights to be applied across all nations. We emphasise that this is merely a way of proposing weights for subsequent consideration since one cannot guarantee a priori that the results will be acceptable. The acceptability has to de decided by the user.

2. FROM ARITHMETIC MEAN TO GEOMETRIC MEAN

From 1990 to 2009 the HDI was calculated using an equally weighted sum of three indicators or sub-indices which lie in the range 0 to 1. The three sub-indices dealt with life expectancy, education, and per capita Gross Domestic Product adjusted for purchasing power parity. Desai (1991) criticized the use of this additive approach for the HDI, saying ‘additivity over the three variables implies perfect substitution which can hardly be appropriate’. This was echoed by Herrero et al (2010) “no matter how bad the health state could be, it can always be compensated with further education or income at a constant rate”, and in the critical review of the HDI by Sagar and Najan (1998), who argue that allowing trade-offs in this way ‘suggests that you can make up in one dimension the deficiency in another. Such a reductionist view of human development is completely contrary to the UNDP’s own definition’. They recommended a multiplicative scheme because this would mean that ‘the more severe the deprivation on any dimension, the more difficult it is to have a high HDI. This better addresses UNDP’s concerns about focusing on the state of the more vulnerable segments of society in determining the level of human development in any country’.

Sagar and Najan (1998) pointed out another valuable aspect of the multiplicative approach. Consider an increase of 0.1 in one of the components. This would be a much more significant change for a country which has shifted from 0.1 to 0.2 on this component, than for a more developed country that improved from 0.8 to 0.9. Under the additive scheme both countries would achieve the same increase in the overall score, whereas under the multiplicative scheme the overall score would rise by a greater amount for the less developed country – it would reflect the fact that it had experienced a greater percentage change.

In 2010 the UNDP adopted the multiplicative approach for aggregating the three dimensions. The inputs to the HDI were: life expectancy, education (measured as the geometric mean of years of schooling and expected years of schooling), and standard of living (measured as the natural log of gross national income per capita adjusted for purchasing power parity). These were each normalized into the zero to one range to provide sub-indices which we denote as L, E, and Y respectively. In a multiplicative setting the weights are applied by raising each variable to a power. Equal weights were adopted.

Thus the formula used by the UNDP in 2010 was: $HDI = L^{1/3}\ E^{1/3}\ Y^{1/3}$

3. EQUAL WEIGHTING CRITICIZED

The lead author of the 2009 and 2010 Human Development Reports (Klugman, 2011), which contain the HDI results, has observed that “the choice of equal weights has been criticized. This point has been recognized by Anand and Sen - one of the architects of the HDI, 1997), who wrote that ‘any choice of weights should be open to questioning and debating in public discussions’”. One of the first to question the use of equal weights in the HDI was Kelley (1991). Chowdhury and Squire (2006) described the use of equal weights as “obviously convenient but also universally considered to be wrong", and embarked on a major exercise of asking researchers in the field of development from around the world to offer their choices. Lind’s (2010) argument against equal weights is that ‘an index of human development should reflect humans' revealed opinion of their own needs and interests'. He then proceeds to use this approach to develop a 'calibrated HDI' which has a multiplicative functional form: 'All young person's will at some time decide to end their education or training and to start working, doing so whenever their expected lifetime earnings and cultural benefits would not improve by more time spent as a student. The choice reflects their relative valuation of greater earnings versus more education'.

Using this type of reasoning the weights are deduced using revealed preferences from data on 26 nations. Since the study focused on developed nations these preferences would not be universally applicable. The largest weight so derived went on life expectancy and the lowest to education.

Nguefack-Tsague et al (2011) also felt that using equal weights was arbitrary; they applied principal component analysis based on the correlations between the three components. The objective of this technique is to find linear combinations of the variables with maximum variance. They found that the first such principal component had roughly equal weights. Chowdhury (2005) used a modified form of principal component analysis which involved using data which had not been standardised by dividing by the standard deviation. Under this scheme those variables with higher coefficient of variation (standard deviation divided by the mean) receive a higher weight. This resulted in GDP and life expectancy having very similar weight, but education receiving twice the weight of these two dimensions. Chowdhury observes that 'the idea that equal weighting should be the norm and the burden of proof should fall on differential weighting is not well founded'. Decancq and Lugo (2008) speak of 'researchers that would like to avoid the hazardous question of how to set the weights, and therefore choose for equal weighting'; they describe this as an 'agnostic viewpoint'. They also point out that principal component analysis is actually a way of summarising the data rather than a procedure for weight setting.

A normative approach to choosing a set of weights is to ask 'experts'. Chowdhury and Squire (2006) adopted this method when they surveyed experts in the field of development research. Two hundred questionnaires were sent out and efforts were made to stratify for income and across seven geographic regions. Respondents were asked to choose a weight between 0 and 10 for each dimension, and these were then averaged. Analysing the 105 responses from 60 countries, they found that they “received more responses from researchers from the low-income and lower-middle income countries, proportionately about the right number of responses from the upper-middle and high-income non-OECD countries, and fewer responses from the high-income OECD countries.”

It is apparent that surveys are difficult for a number of reasons, including the decision of whose opinions should be sought. Indeed the authors state:

“Our first conclusion is that...future research along these lines should either survey a very different sample, say, policy-makers for example, or employ very different methods such as carefully structured interviews.”

Surveys also suffer from the well known issues of non-response bias, sampling bias, expense, and the fact that they are time consuming to carry out.

One major difficulty with asking people to provide weights is that even if they have an ordering of criteria, they will still be uncomfortable in deciding actual values because there is an infinity of possibilities – they will feel as though they are ‘picking numbers out of the air’ – a sense of arbitrariness. Using mathematical procedures such as optimisation can help avoid this difficulty; values can be generated and then presented for approval or rejection.

4. HDI AND DATA ENVELOPMENT ANALYSIS (DEA)

We turn to an approach which is very different from surveying people from each country although it allows for each country to influence the result; it also does not suffer from any of the problems associated with surveys: data envelopment analysis (DEA). Researchers have been drawn to data envelopment analysis (DEA) for weight setting because it offers the possibility of generating weights automatically from the data, provided that the variables have already been selected. It therefore removes a large element of subjectivity and so provides a move toward objectivity.

The idea behind DEA is that a separate set of weights is computed for each country: The weights for a country are chosen so as to maximise its aggregate score subject to the condition that these weights do not lead to any country exceeding a score of 100%. It can thus be viewed as an optimal set of weights for that country. DEA uses a frontier as a benchmark. This frontier is delineated by the best-practice countries according to a principle of dominance known as convex-dominance. This is best understood using a diagram and the reader who is not familiar with DEA is referred to Appendix 1.

DEA has been applied to the original additive form of the HDI by several authors e.g. Bougnol et al (2010). Perhaps the earliest such application was by Mahlberg and Obersteiner (2001) who were attracted to DEA because of its 'scientific' approach.

DEA has already been applied to the new formulation of the HDI by Blancard and Hoarau (2011). This followed the initial application to the multiplicative HDI form by Zhou et al (2010) using 27 nations in the Asia and Pacific region; this was remarkably prescient as at that time the UNDP had not decided to switch to a weighted product formula. Neither of these papers presented a single set of weights to be applied to all countries – that was not their aim. Rather their objective was to make use of the fact that DEA provides weights for *each* country in a non-subjective way.

In the DEA literature the first multiplicative model was proposed by Charnes et al (1982) and this was the one used by Zhou et al (2010) as well as by Blancard and Hoarau (2011). Zhou et al (2010) reported that their results suffered from not being scale-invariant; specifically, when they multiplied each of the three sub-indicators by 10 they found that some of the results were different, but were unable to account for this effect. The originators of the multiplicative DEA model had actually produced a follow-up paper (Charnes et al, 1983) which succeeded in resolving this issue in a simple way. The corrected method simply involves an extra factor (k) in the multiplicative performance measure which is optimized ($k\ L^a\ \ S^b\ Y^c$). This variable (k) becomes an additive term when logarithms are taken, and so absorbs the effect of multiplying any indicator by a constant; it had originally been omitted and this was the likely cause for the lack of scale-invariance. We shall be adopting the improved units-invariant form in this paper.

Also worth mentioning here is that mathematical optimization has been used to derive HDI weights but without going through the DEA route. Hatefi and Torabi (2010) used a single stage optimization to find a common set of weights for all countries. They considered the deviations of the scores from 100% as one-sided residuals and computed the weights which minimised the largest of these deviations (i.e. minimax deviation from 100%). This is equivalent to finding the weights which squeeze all the scores within the narrowest range (with 100% being the upper end of this range). It is also equivalent to choosing the weights which maximize the lowest score. Such a minimax approach to weight setting has been shown to possess a number of drawbacks by Tofallis (2010). Most importantly, these include the fact that it is the data from the worst performing countries that ends up controlling the final weights – in an effort to make their score as high as possible.

5. AN 'AUTOMATIC-DEMOCRATIC' APPROACH TO SETTING WEIGHTS

Our proposed approach is to use DEA as a first stage to find the specific weights for each country which maximises its score. The advantage of this is that countries cannot then argue that the weights have been biased against them. The second stage is to use the resulting scores as the dependent variable in a regression on the underlying measures. The regression coefficients will provide the common set of weights we are seeking. Appendix 2 explains the details of the DEA model.

Our work is in some ways similar to that of Despotis (2005a, 2005b) in that after the DEA computations a second stage is employed to estimate a common set of weights. However, Despotis used a goal programming model for the second stage which included a user-controlled parameter (t) for adjusting the mixture of the minimax norm ($t = 0$) and the L1 norm ($t = 1$):

"Varying the parameter *t* between these two extreme values, we provide the model with the flexibility to 'compromise' between the two norms and to explore different sets of common weights and consequently different global efficiency patterns" (Despotis, 2005a).

For most of the range, $t = 0$ to 0.99, Despotis found that the highest weight went on the education index (0.613), followed by the life expectancy (0.433), and 0.032 for per capita gross domestic product. For $t = 0.991$ to 0.995 the weights were 0.815 on life expectancy, 0.267 on education, and 0.043 for GDP. Finally, for $t = 0.996$ to 1.0 they were 0.834 for life expectancy, 0.25 for education, and 0.002 for GDP. Note that these weights were derived using the additive formula for HDI and so are not comparable with those that we shall derive as we shall be working with the new multiplicative form.

For the reason stated earlier we choose to avoid the minimax approach, and furthermore we prefer to avoid introducing the *t* parameter to maintain simplicity.

6. RESULTS BASED ON NORMALIZED DATA

Appendix 3 displays the optimal (DEA) scores for each country based on the three sub-indices that the UNDP published in 2010, i.e. using data that has been normalized using their scheme. There were six countries which attained the upper limit score of 1.0. These were Australia, Hong Kong, Japan, Liechtenstein, New Zealand, and Norway. The lowest score was 0.39, obtained for Afghanistan.

These optimal scores were used as the dependent variable in a non-linear least squares regression with the three sub-indices as explanatory variables using the following model:

$$\text{DEA score} = K\, L^{A}\, E^{B}\, Y^{C} + \text{residual}$$

The Levenberg-Marquardt solution method was used within the SPSS statistical program to identify the values of the parameters K, A, B, and C. We did not impose any constraints on the parameters - such as forcing them to sum to unity - in order to achieve a better fit to the data. For a discussion on the use of regression on DEA results see McDonald (2009), who recommends ordinary least squares regression because it gives consistent estimators, and argues that tobit regression is inappropriate.

The estimated weights (and standard errors) were found to be as follows:

K = 1.026 ± 0.005

A = 0.732 ± 0.028

B = 0.056 ± 0.015

C = 0.074 ± 0.013

Thus the resulting model is:

$$\text{Score} = 1.026\, L^{0.732}\, E^{0.056}\, Y^{0.074}$$

Thus we see that the largest weight is placed on the Life Expectancy index, with a much lower weight on Gross National Income, and education receiving the lowest. Interestingly this ordering of the weights is the same as that obtained by Lind (2010) using revealed preferences. Weights in a multiplicative formula are interpreted in percentage terms, thus in the 2010 UNDP formula where all the weights equal 1/3, the interpretation is that a marginal increase of 1% in any one of the sub-indices leads to a 0.33% increase in the HDI score; it follows that a 1% increase in one of the indices is equivalent to a 1% increase in any one of the others. In the proposed scheme however this does not apply. We have that a 1% rise in the life expectancy index leads to a 0.732% rise in the overall score. By contrast a 1% rise in the education index leads to only a 0.056% rise in the overall score. Putting these two figures together it follows that a 13% rise (0.732/0.056) in the education index equates to a 1% rise in the life expectancy index. Similarly, we can see that a 10% rise in the income index roughly equates to a 1% change in the life expectancy index.

A high goodness of fit was achieved when the above model was fitted: $R^2 = 0.964$. The correlations (r) between parameter estimates were: r(A,B) = -0.449, r(A,C) = -0.470, r(B,C) = -0.369.

The value of the factor K does not of course affect the rankings. One can for example reduce it to ensure that the resulting scores do not exceed the DEA scores – which were after all the most optimistic estimates. Changing K to 0.9046 achieves this effect and leads to a score range of 0.338 (Zimbabwe) to 0.883 (Japan). By comparison the UNDP HDI range in 2010 was from 0.140 (Zimbabwe) to 0.938 (Norway).

The scores based on the above model are displayed in Appendix 3, together with the official HDI results produced by the United Nations Development Programme.

For comparison purposes we shall consider the differences in rank between the UNDP's HDI score and the one proposed here. The average change is not small, in absolute terms it is a shift of 10.4 positions. At the top now is Japan, followed by Australia (maintaining its HDI position of second), then Hong Kong, Switzerland, and Norway (which came first under HDI). The biggest fallers are Kazakhstan, from 66 to 109, and the Russian Federation, from 65 to 103. These large changes are explained by the greater weight that is now being placed on life expectancy: Kazakhastan's life expectancy is ranked 119th out of the 169 countries in the table, and Russia's is ranked 111th. The biggest climbers under the new rating are Syria, up 34 places from 111 to 77, and Vietnam, up 31 places from 113 to 82. Once again it is the weight on life expectancy that is the main reason for these shifts: life expectancy in Syria is 74.6 years and is ranked 55th, and in Vietnam it is 74.9 years and is ranked 53rd.

Often, precise scores are not used by policy and decision makers but rather their general category. The UN Development Programme classifies countries as having 'very high development' if they are in the top quartile for the HDI, the second quartile countries are said to have 'high development', the third quartile is termed 'medium' and the lowest quartile nations are described as having 'low' development. Since 169 nations were listed in the 2010 report, the UNDP placed 43 countries in the top category and 42 in each of the others. Using this same scheme we can see which nations have shifted from one category to another under the proposed methodology. We find that five countries would move from the top quartile to the second quartile and the same number move in the opposite direction. Six countries rise from the third quartile to the second; interestingly these include China, whereas the

Russian Federation would fall from the second to the third quartile. When it comes to foreign aid we naturally expect most attention to be focused on the lowest group. The changes here would be that five countries would rise out of the lowest quartile, including Bangladesh, Nepal, and Yemen; whilst those moving to the lowest group would be Botswana, Congo, Equatorial Guinea, Swaziland, and South Africa (where life expectancy is at rank 151).

7. RESULTS BASED ON DATA WITHOUT NORMALIZATION

From 1990 to 2009 the three dimensions of the HDI were aggregated using an additive approach. Since each component dimension is measured in different units, conversion to an index in the range 0 to 1 was used to enable them to be added together. This was done using range normalization as follows:

Component index value = (actual value − lower limit) / (upper limit − lower limit)

From 2010 onwards the three dimensions are no longer added together, rather they are aggregated using multiplication, and so it is not essential to use such a normalization. Note that changing the units of measurement of any dimension (e.g. measuring life expectancy in months rather than years) merely multiplies all observations by a constant factor (12) and so the final rankings remain unchanged even after the multiplicative aggregation of all three dimensions. This is a very useful property. Avoiding the above normalization also means that proportionality is retained, apart from GNI where the logarithm is used to model diminishing marginal returns.

Another problem with the above normalization is the choice of upper and lower limits. This is illustrated by the fact that the UNDP has itself altered these on a number of occasions over the years. Most recently, in 2010 it decided to change the lower limit on life expectancy from 25 years to 20 years. Whilst the Gross National Income per capita lower limit was set at $163 in 2010 and was reduced to $100 in 2011. There is no agreement on what these thresholds should be and consequently they are controversial.

We therefore repeated our analysis using the un-normalised data used in the 2011 Human Development Report. Specifically, we use life expectancy in years, education measured as the geometric mean of expected years of schooling and mean years of schooling (this is equivalent to the index used by the UNDP because they use a zero lower limit for schooling and hence normalization has no effect), and ln(GNI) per capita.

The estimated weights (and standard errors) were found to be as follows:

$K = 0.038 \pm 0.003$

$A = 0.590 \pm 0.022$

$B = 0.025 \pm 0.009$

$C = 0.289 \pm 0.022$

The goodness of fit to the DEA scores as measured by R^2 is 0.969.

Thus the resulting formula is:

$$\text{Score} = 1.039\, L^{0.59}\, E^{0.025}\, Y^{0.289}$$

This is perhaps more satisfactory than the formula based on normalized data because the dominance of life expectancy over the other two dimensions is less overwhelming than before, with the weight falling from 0.732 to 0.59. The weight on GNI has risen from 0.074 to 0.289, whilst education still retains the lowest weight. These changes also illustrate the hidden effect of normalization. This new formula is simpler and more direct; it is longer influenced by the choice of lower and upper thresholds.

Comparing the rankings using the formula and the official HDI ranks one finds that the average shift is the same as before: 10.4 positions. Hong Kong now takes the lead, followed by Japan, Switzerland, and Norway.

One can simplify the formula further by removing the log transformation on Gross National Income. Repeating our analysis then gives:

$$\text{Score} = 0.051\, L^{0.602}\, E^{0.029}\, Y^{0.031}$$

The weights on education and GNI are now almost the same. Fortunately this simplification has very little effect on the rankings from the analysis which has the log transformation for GNI: the mean absolute rank change is only 0.6 and the largest rank shift is just three positions. Thus the diminishing returns effect of the logarithm is now effectively reproduced by a reduced exponent. This simpler formula is also preferable because it makes interpretation of weights and substitution rate calculations easier.

8. CONCLUSION

In the field of multi-dimensional social indicators the issue of weight selection is possibly the one that researchers find most difficult. It is inevitable that there will be a difference of views. One respoinse to this is to try and aggregate the views of the chosen participants. But as Decancq and Lugo (2008) point out:

> ‘The main source of concern with participatory methods relates to the selection of participants, a concern that holds true for any sets of groups (experts, representative individuals, and policy-makers). Selection of participants can be biased - some groups being under-represented, or simply uninformed’.

To avoid controversy the path of least resistance is often adopted, namely to attach equal weight to all dimensions. Some might argue that this is merely a way of avoiding having to think about the problem. This is, in a way, understandable if not excusable, because even if one were certain about the relative ordering of criteria there is still an unlimited number of combinations of weight values that will satisfy that ordering. One way out of this perplexing puzzle is to use mathematical procedures such as optimisation to generate weights which can then be considered.

Of the data-driven methods for weight-setting in the HDI, principal component analysis (PCA) seems to have been the one most often suggested, until the recent interest in using DEA. The PCA idea is to find a weighted combination which explains as much of the variation in the data as possible. Unfortunately the issues associated with PCA are not often appreciated:

> “Principal components are generally changed by scaling and are therefore not a unique characteristic of the data. If one variable has a much larger variance than the other variables, then this variable will dominate….whereas if the variables are all scaled to have unit variance, then the first principal component will be quite different in kind”.

> “The conventional way of getting round the scaling problem is to analyse the correlation matrix rather than the covariance matrix…This ensures that all variables are scaled to have unit variance. This scaling procedure is still arbitrary to some extent. If the variables are not thought to be of equal importance, then the analysis of the correlation matrix is not recommended” (Chatfield and Collins, 1992, p. 70-71)

DEA seems to be more relevant to the task of measuring performance than PCA because the underlying scores are based on a best observed practice frontier, as illustrated in Figure 1 of Appendix 1. They are sensible in that higher scores are assigned to countries which are closer to the frontier. The DEA approach has the advantage that no particular shape for the frontier is imposed *a priori*, (this is why it is referred to as a non-parametric method). The slope of any facet on the frontier implies a particular set of weights. A possible drawback of DEA is that by allowing great weight flexibility one may obtain zero weights, which is against the spirit of a multidimensional index. Hence analyses and rankings based on DEA scores alone are hard to justify since it is difficult to justify a comparison of two countries when one country has its score based on three dimensions and the other has only used two. There is a strand of the DEA literature which attempts to deal with this issue by imposing restrictions on the weights, and indeed Zhou et al (2010) suggest asking experts to agree on a particular range for the contribution of each dimension, and also give results based on restricted ranges. This shifts the problem of deciding on weights to one of deciding on weight restrictions, and since we have adopted a non-subjective philosophy for the former it seems odd to give up this philosophy at a later stage. We have chosen not to go down that route because the act of choosing such restrictions seems ad hoc or arbitrary;

Another difficulty with DEA is that for countries lying on the frontier the same DEA score can be attained using alternative optimal weights. For that reason we have not attempted to use those weights subsequently – only the scores are used since these are unique. Given these scores and the underlying data we then needed to arrive at a common set of weights. We used regression to do this - this is an attempt to stay as close to the DEA scores as possible.

A possible disadvantage of our approach is that the weights can change from one year to the next as the underlying data changes. We expect such changes to be small because they arise from small shifts in the location of around 180 points/countries and the effect on the associated regression line. It is worth observing that the existing methodology for the HDI is also affected by data changes - those associated with the upper limit observations; though this is less noticeable or transparent because the changes affect the component scores via the normalization i.e. implicitly, whilst the explicit weights remain nominally equal and unchanged.

In summary, the approach we are presenting involves two steps. The first step tries to deal with the potential concern that assigned weights are disadvantageous to particular nations by deliberately finding the weights which are most advantageous to each nation in turn. We then used these optimal nation-specific weights to calculate an overall index score for that country. We did not attempt to average these weights across nations because they are not guaranteed to be unique i.e. a given nation can achieve the same optimal score with different weight combinations (alternative optima).

In the second step, we sought a common set of weights to apply to all nations. These were chosen so that the final scores would be as close as possible to the optimal scores found in step 1 by minimizing the sum of squares of these deviations from the optimal scores (i.e. least squares regression). We have thus shown how to produce a simple transparent formula for the final scores and rankings – something which DEA alone does not provide.

The general approach may be described as having an 'automatic-democratic' quality because it automatically generates weights from the data, whilst at the same time ensuring that each country influences the results in an equal and fair way without outside interference.

APPENDIX 1: Explanation of DEA using a diagram

a) Explaining DEA using a 2-dimensional scatter-graph.

For the purposes of illustration we shall restrict attention to just two dimensions; the principles apply equally well in three dimensions. Suppose we plot the life expectancy and years of schooling for a number of countries as points on a graph as in Figure 1.

DEA identifies the observed best–practice without having to pre-specify weights of importance. Instead it uses a dominance argument: in the diagram below country P is dominated by country B because the latter has higher life expectancy and a higher level of schooling. Thus *whatever* weights are attached to the two criteria would cause B to have a higher score than P.

Countries A, B, and C are not dominated by any other country and so form the best-practice frontier, or envelope. These are then assigned the maximum score of unity. Countries behind the frontier have their score calculated by the proportion of their distance to the frontier, thus for P the score would be given by the ratio OP/OP' [where P' is the point where the ray from the origin through P meets the frontier]. So if P is three quarters of the way to the frontier then its score would be ¾ = 0.75.

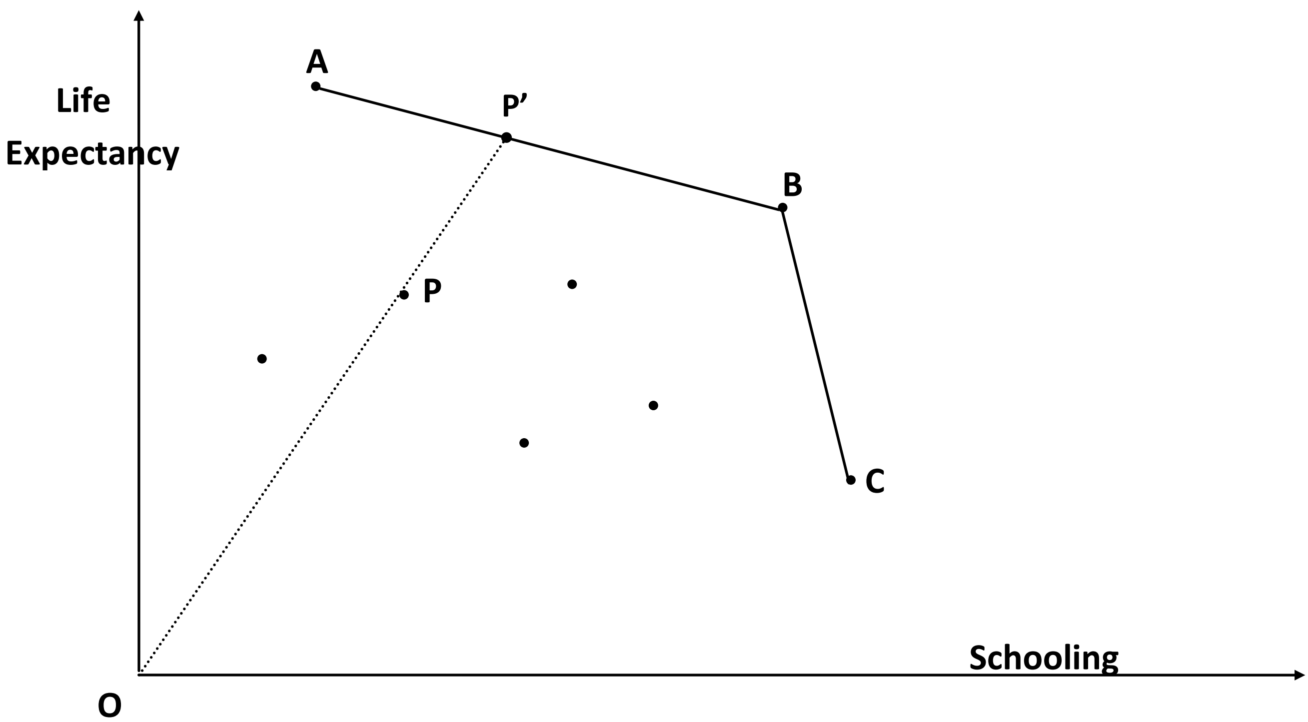

Figure1. The frontier according to DEA

b) Going beyond DEA to obtain a single set of weights:

The slope of line AB implies one set of weights – these would favour countries A and B. The slope of line BC represents another set of weights - these would favour countries B and C. The method in this paper uses regression to identify an intermediate set of weights that does not favour any one country, but rather allows every country to influence the outcome whilst keeping as close as possible to the scores found in the DEA analysis.

APPENDIX 2: THE DEA MODEL

We solve a separate optimization problem for each country. Let H_o denote the score of the particular country being assessed.

We wish to find an individual set of weights (K_o, A_o, B_o, C_o) which maximizes the score for that country:

maximize $H_o = K_o L^{Ao} E^{Bo} Y^{Co}$

subject to the constraints:

$A_o + B_o + C_o = 1$

$K_o, A_o, B_o, C_o \geq 0$ (all weights non-negative),

and

$H_i \leq 1$ for i = 1 to n (n is the number of countries)

i.e. none of the country scores exceed 100% using these weights.

The above problem is converted to a linear programming problem by taking logarithms: Use K\` to denote ln(K), L\` to denote ln(L) etc.

The problem then becomes:

Maximize $H`_o = K`_o + A_o L`_o + B_o E`_o + C_o Y`_o$

subject to the constraints

$A_o + B_o + C_o = 1$

$H`_i \leq 0$ (because ln(1) = 0), for i = 1 to n

i.e. $K_o` + A_o L`_i + B_o E`_i + C_o Y`_i \leq 0$, for i = 1 to n

and $A_o, B_o, C_o \geq 0$

Once the optimal value is found in this way it can be converted back using anti-logs:

$H = \exp H`$

Note that the optimal value of H arising from the above linear programming problem is unique, although it may be possible to achieve the same score with different weights – alternative optima. We do not use the weights in the regression stage, only the unique H values.

The scores can also be obtained using a different formulation which normalizes the score of the country being assessed to unity (which means the log is zero), and then minimizes the maximum score across all countries. The score of the assessed unit is then the reciprocal of the largest score:

Minimize $h`_o$

subject to the constraints

$$A_o + B_o + C_o = 1$$

$$K`_o + A_oL`_o + B_oE`_o + C_oY`_o = 0$$

$$K_o` + A_oL`_i + B_oE`_i + C_oY`_i \leq h`_o, \quad \text{for } i = 1 \text{ to } n$$

$$\text{and } A_o, B_o, C_o \geq 0$$

We applied both formulations and found identical scores for all countries.

Appendix 3. Table of results.

Nation	HDI rank (UNDP)	Optimal weights score (DEA)	DEA rank	New score	New rank	Rank difference (New vs. HDI)
Afghanistan	155	0.39	169	0.39	167	-12
Albania	64	0.90	42	0.79	48	16
Algeria	84	0.84	78	0.74	75	9
Andorra	30	0.97	18	0.85	23	7
Angola	146	0.55	149	0.44	157	-11
Argentina	46	0.88	50	0.79	47	-1
Armenia	76	0.86	66	0.76	67	9
Australia	2	1.00	1	0.88	2	0
Austria	25	0.97	23	0.85	22	3
Azerbaijan	67	0.81	96	0.73	86	-19
Bahamas	43	0.88	51	0.78	51	-8
Bahrain	39	0.90	41	0.80	41	-2
Bangladesh	129	0.74	117	0.64	116	13
Barbados	42	0.91	40	0.81	38	4
Belarus	61	0.80	99	0.72	92	-31
Belgium	18	0.97	24	0.85	18	0
Belize	78	0.90	43	0.78	50	28
Benin	134	0.67	130	0.59	129	5
Bolivia	95	0.75	115	0.67	111	-16
Bosnia and Herzegovina	68	0.88	52	0.78	57	11
Botswana	98	0.74	118	0.56	133	-35
Brazil	73	0.84	77	0.75	72	1
Brunei Darussalam	37	0.95	32	0.82	35	2
Bulgaria	58	0.85	68	0.77	60	-2
Burkina Faso	161	0.53	153	0.47	152	9
Burundi	166	0.50	157	0.44	158	8
Cambodia	124	0.67	131	0.60	126	-2
Cameroon	131	0.51	156	0.49	151	-20
Canada	8	0.98	13	0.86	9	-1
Cape Verde	118	0.82	91	0.70	101	17
Central African Republic	159	0.44	168	0.42	166	-7
Chad	163	0.46	161	0.43	161	2
Chile	45	0.93	38	0.82	33	12
China	89	0.85	73	0.75	73	16
Colombia	79	0.85	74	0.75	71	8
Comoros	140	0.73	120	0.62	123	17
Congo	126	0.54	151	0.52	144	-18
Congo	168	0.44	166	0.39	168	0
Costa Rica	62	0.94	35	0.81	39	23
Croatia	51	0.90	45	0.80	42	9
Cyprus	35	0.95	34	0.84	29	6
Czech Republic	28	0.92	39	0.82	36	-8
Denmark	19	0.95	33	0.84	30	-11
Djibouti	147	0.57	146	0.52	142	5
Dominican Republic	88	0.84	79	0.74	78	10

Ecuador	77	0.88	54	0.77	58	19
Egypt	101	0.80	98	0.71	96	5
El Salvador	90	0.82	90	0.73	85	5
Equatorial Guinea	117	0.79	102	0.50	149	-32
Estonia	34	0.88	53	0.78	53	-19
Ethiopia	157	0.57	147	0.50	148	9
Fiji	86	0.79	101	0.71	99	-13
Finland	16	0.96	25	0.85	21	-5
France	14	0.98	11	0.87	7	7
Gabon	93	0.72	123	0.63	121	-28
Gambia	151	0.58	144	0.52	143	8
Georgia	74	0.84	80	0.74	83	-9
Germany	10	0.97	22	0.86	16	-6
Ghana	130	0.59	140	0.54	134	-4
Greece	22	0.95	30	0.84	27	-5
Guatemala	116	0.80	97	0.70	102	14
Guinea	156	0.62	137	0.53	138	18
Guinea-Bissau	164	0.45	163	0.42	164	0
Guyana	104	0.76	110	0.68	108	-4
Haiti	145	0.66	133	0.57	132	13
Honduras	106	0.83	82	0.73	91	15
Hong Kong	21	1.00	1	0.87	3	18
Hungary	36	0.87	57	0.78	52	-16
Iceland	17	0.99	9	0.87	6	11
India	119	0.70	125	0.63	120	-1
Indonesia	108	0.82	95	0.71	94	14
Iran	70	0.82	92	0.74	79	-9
Ireland	5	0.97	21	0.86	14	-9
Israel	15	0.98	15	0.86	11	4
Italy	23	0.97	17	0.86	12	11
Ivory Coast	149	0.61	139	0.54	136	13
Jamaica	80	0.83	84	0.74	80	0
Japan	11	1.00	1	0.88	1	10
Jordan	82	0.84	76	0.75	74	8
Kazakhstan	66	0.80	100	0.68	109	-43
Kenya	128	0.57	148	0.53	139	-11
Korea (Republic of)	12	0.96	29	0.85	24	-12
Kuwait	47	0.96	26	0.82	37	10
Kyrgyzstan	109	0.77	108	0.68	107	2
Lao	122	0.73	122	0.64	117	5
Latvia	48	0.85	71	0.76	64	-16
Lesotho	141	0.49	159	0.42	163	-22
Liberia	162	0.62	136	0.51	146	16
Libya	53	0.86	61	0.78	55	-2
Liechtenstein	6	1.00	1	0.85	20	-14
Lithuania	44	0.84	75	0.76	69	-25
Luxembourg	24	0.98	14	0.85	25	-1
Macedonia TFYR	71	0.86	62	0.77	62	9
Madagascar	135	0.65	135	0.58	131	4
Malawi	153	0.55	150	0.50	150	3
Malaysia	57	0.87	58	0.78	56	1
Maldives	107	0.83	85	0.72	93	14
Mali	160	0.46	160	0.43	159	1

Malta	33	0.95	31	0.84	28	5
Mauritania	136	0.59	141	0.54	137	-1
Mauritius	72	0.82	89	0.74	76	-4
Mexico	56	0.90	44	0.80	45	11
Micronesia	103	0.78	106	0.69	105	-2
Moldova	99	0.78	105	0.69	104	-5
Mongolia	100	0.75	112	0.68	110	-10
Montenegro	49	0.87	55	0.78	54	-5
Morocco	114	0.82	93	0.71	95	19
Mozambique	165	0.45	164	0.42	165	0
Myanmar	132	0.68	128	0.60	127	5
Namibia	105	0.67	129	0.63	122	-17
Nepal	138	0.75	113	0.63	119	19
Netherlands	7	0.97	16	0.86	15	-8
New Zealand	3	1.00	1	0.86	10	-7
Nicaragua	115	0.85	72	0.73	90	25
Niger	167	0.51	155	0.45	155	12
Nigeria	142	0.46	162	0.45	154	-12
Norway	1	1.00	1	0.87	5	-4
Pakistan	125	0.75	114	0.65	114	11
Panama	54	0.89	49	0.79	49	5
Papua New Guinea	137	0.66	134	0.58	130	7
Paraguay	96	0.83	86	0.73	87	9
Peru	63	0.85	70	0.76	66	-3
Philippines	97	0.83	83	0.73	88	9
Poland	41	0.89	47	0.80	44	-3
Portugal	40	0.94	36	0.83	32	8
Qatar	38	1.00	7	0.80	40	-2
Romania	50	0.85	69	0.77	63	-13
Russian Federation	65	0.77	107	0.70	103	-38
Rwanda	152	0.49	158	0.47	153	-1
Sao Tome and Principe	127	0.73	121	0.64	118	9
Saudi Arabia	55	0.87	59	0.77	61	-6
Senegal	144	0.57	145	0.53	141	3
Serbia	60	0.86	64	0.77	59	1
Sierra Leone	158	0.45	165	0.42	162	-4
Singapore	27	0.99	10	0.85	17	10
Slovakia	31	0.89	48	0.79	46	-15
Slovenia	29	0.93	37	0.83	31	-2
Solomon Islands	123	0.74	116	0.65	115	8
South Africa	110	0.70	124	0.52	145	-35
Spain	20	0.97	20	0.86	13	7
Sri Lanka	91	0.86	63	0.75	70	21
Sudan	154	0.62	138	0.54	135	19
Suriname	94	0.78	103	0.70	100	-6
Swaziland	121	0.58	143	0.45	156	-35
Sweden	9	0.98	12	0.87	8	1
Switzerland	13	0.99	8	0.87	4	9
Syrian	111	0.86	60	0.74	77	34
Tajikistan	112	0.75	111	0.67	113	-1
Tanzania	148	0.58	142	0.53	140	8
Thailand	92	0.78	104	0.71	98	-6
Timor-Leste	120	0.67	132	0.61	125	-5

Togo	139	0.69	127	0.59	128	11
Tonga	85	0.83	81	0.73	84	1
Trinidad and Tobago	59	0.83	87	0.73	89	-30
Tunisia	81	0.86	65	0.76	68	13
Turkey	83	0.83	88	0.74	81	2
Turkmenistan	87	0.73	119	0.67	112	-25
Uganda	143	0.54	152	0.51	147	-4
Ukraine	69	0.82	94	0.71	97	-28
United Arab Emirates	32	0.96	27	0.82	34	-2
United Kingdom	26	0.96	28	0.85	26	0
United States	4	0.97	19	0.85	19	-15
Uruguay	52	0.90	46	0.80	43	9
Uzbekistan	102	0.77	109	0.68	106	-4
Venezuela	75	0.86	67	0.76	65	10
Viet Nam	113	0.87	56	0.74	82	31
Yemen	133	0.69	126	0.61	124	9
Zambia	150	0.44	167	0.43	160	-10
Zimbabwe	169	0.52	154	0.34	169	0

REFERENCES

Anand, S., and Sen, A. (1993). *Human Development Index: Methodology and Measurement*. Human Development Report Office Occasional Paper 19. UN Development Programme, New York.

Anand, S., and Sen, A. (1997). Concepts of Human Development and Poverty: A Multidimensional Perspective, in *Human Development Papers 1997: Poverty and Human Development*: 1 – 19. UNDP-HDRO, New York.

Bate, R and Boateng, K. (2007). Drug Pricing and Its Discontents, At Home and Abroad. *Health Policy Outlook*, vol. 9, August 2007.

Blancard, S and Hoarau, J-F. (2011). Optimizing the formulation of the United Nations' human development index: An empirical view from data envelopment analysis. *Economics Bulletin*, 31(1), 989-1003.

Bougnol, M-L; Dula, JH; Estellita Lins, MP; and Moreira da Silva, AC. (2010). Enhancing standard performance practices with DEA. *Omega*, 38, 33‑45.

Charnes, A., Cooper, WW and Seiford, L. (1982). A multiplicative model for efficiency analysis. *Socio-Economic Planning Sciences*, 16(5), 223-224.

Charnes, A., Cooper, WW., Seiford, L and Stutz, J. (1983). Invariant multiplicative efficiency and piecewise Cobb-Douglas envelopments. *Operations Research Letters*, 2(3), 101-103.

Chatfield, C and Collins, AJ. (1992). *Introduction to multivariate analysis*. Chapman and Hall, London.

Chowdhury, S. (2005). The Human Development Index: An Exercise in Objectivity. Downloadable from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=787586

Chowdhury, S. and Squire, L. (2006). Setting weights for aggregate indices: An application to the commitment to development index and human development index, *Journal of Development Studies* 42(5), 761–771.

Decancq, K and Lugo, MA. (2008). *Setting weights in multidimensional indices of well-being and deprivation.* Oxford Poverty & Human Development Initiative, Working Paper 18.

Desai, M. Human development: concepts and measurement. (1991). *European Economic Review*, 35 (2/3), 350-357.

Despotis, D.K. (2005a). A reassessment of human development index via data envelopment Analysis. *Journal of the Operational Research Society* 56(8), 960-980.

Despotis, D.K. (2005b). Measuring human development via data envelopment analysis: the case of Asia and the Pacific. *Omega* 33(5), 385-390.

Gohou, G and Soumare, I. (2011). Does Foreign Direct Investment Reduce Poverty in Africa and are there Regional Differences? *World Development*, in press. Available online: http://www.sciencedirect.com/science/article/pii/S0305750X11001446

Grimm, M, Harttgen, K, Klasen, S, and Misselhorn, M. (2008). A Human Development Index by Income Groups. *World Development*, 36(12), 2527–2546.

Hatefi, SM, and Torabi, SA. (2010). A common weight MCDA–DEA approach to construct composite indicators. *Ecological Economics*, 70, 114–120.

Henninger, N. and Snel, M. (2002). *Where are the poor? Experiences with the development and use of poverty maps.* World Resources Institute, Washington, DC and UNEP/GRID-Arendal, Arendal, Norway.

Herrero, C, Martinez, R and Villar, A. (2010). Multidimensional social evaluation: an application to the measurement of human development. *Review of Income and Wealth*, 56(3), 483-497.

Hicks, DA. (1997). The inequality-adjusted human development index: A constructive proposal. *World Development*, 25(8), 1283-1298.

Hoyland, B, Moene, K, and Willumsen, F. (2012). The Tyranny of International Index Rankings. *J of Development Economics*, 97(1), 1-14.

Kelley, A. C. (1991). The human development index: "handle with care". *Population and Development Review*, 17(2), 315-324.

Klugman, J, Rodriguez, F and Choi, H-J. (2011). *The HDI 2010: New Controversies, Old Critiques*. Human Development Research Paper 2011/01. UNDP, New York.

Lind, N. (2010). A Calibrated Index of Human Development. *Social Indicators Research*, 98, 301–319

Mahlberg, B. and Obersteiner, M. (2001). Remeasuring the HDI by data envelopment analysis. International Institute for Applied Systems Analysis (IIASA), Interim Report IR-01-069, Austria.

McDonald, J. (2009). Using least squares and tobit in second stage DEA efficiency analyses. *European Journal of Operational Research* 197, 792–798.

McGillivray, M and Markova, N. (2010). Global Inequality in Well-being Dimensions. *Journal of Development Studies*, 46(2), 371–378.

Morse, S. (2003). For better or for worse, till the human development index do us part? *Ecological Economics*, 45, 281-296.

Nguefack-Tsague, G, Klasen, S, and Zucchini, W. (2011). On Weighting the Components of the Human Development Index: A Statistical Justification. *Journal of Human Development and Capabilities*, 12(2), 183-202.

O'Neill, H. (2005). Ireland's Foreign Aid in 2004. *Irish Studies in International Affairs*, 16, 279–316.

OPHI (Oxford Poverty and Human Development Initiative). (2010). Oxford University and UNDP join forces to launch a better way to measure global poverty. http://www.ophi.org.uk/news/press-releases/

Reiter, SL and Steensma, HK. (2010). Human development and foreign direct investment in developing countries: the influence of FDI policy and corruption. *World Development*, 38(12), 1678-1691.

Sagar, AD and Najam, A. (1998). The human development index: a critical review. *Ecological Economics*, 25, 249-264.

Tofallis, C. (2010). Multicriteria Ranking Using Weights Which Minimize the Score Range. In *New developments in multiple objective and goal programming*, Eds. Jones, D, Tamiz, M, and Ries, J. Springer.

UNDP (United Nations Development Programme). (1990). *Human Development Report 1990*. New York: Oxford University Press

UNDP (United Nations Development Programme). (2010). *Human Development Report 2010: The Real Wealth of Nations.* UN Development Programme, New York: Palgrave Macmillan.

UNDP (United Nations Development Programme). (2011). *Human Development Report 2011: Sustainability and Equity: A Better Future for All*. UN Development Programme, New York: Palgrave Macmillan.

Zhou, P, Ang, BW, and Zhou, DQ. (2010). Weighting and aggregation in composite indicator construction: a multiplicative optimization approach. *Social Indicators Research*, 96(1), 169-181.