

Keeping things simple: why the Human Development Index should not diverge from its equal weights assumption

Lee M. Stapleton · Guy D. Garrod

Received: 23 October 2006 / Accepted: 6 December 2006 / Published online: 19 January 2007
© Springer Science+Business Media B.V. 2007

Abstract Using a range of statistical criteria rooted in Information Theory we show that there is little justification for relaxing the equal weights assumption underlying the United Nation’s Human Development Index (HDI) even if the true HDI diverges significantly from this assumption. Put differently, the additional model complexity that unequal weights add to the HDI more than counteracts the improvement in goodness-of-fit. This suggests that, in some cases, there may be limited validity in increasing the complexity of a range of other composite sustainability indices.

Keywords Complexity · Composite indices · Human Development Index · Information Theory · Sustainable development · Well-being

1 Introduction

Although there are many definitions of sustainable development perhaps the most famous and widely cited is that put forward by the World Commission on Environment and Development (WCED, 1987) which states that development is sustainable if it “meets the needs of the present without compromising the ability of future generations to meet their own needs”. Following this, Agenda 21 was presented at the United Nations Conference on Environment and Development (UNCED, 1992) in Rio de Janeiro as a blueprint for action, at various spatial scales from local to global, to facilitate a move towards sustainable development. The kind of sustainable development that Agenda 21 was aiming to facilitate is that articulated by the WCED (1987); interestingly, however, the actual text of Agenda 21 does not define sustainable development or sustainability despite using these terms a total

L. M. Stapleton (✉) · G. D. Garrod
Centre for Rural Economy, School of Agriculture, Food and Rural Development, Newcastle University, NE1 7RU Newcastle-upon-Tyne, UK
e-mail: L.M.Stapleton@ncl.ac.uk

of 373 times. In this context sustainability is only distinct from sustainable development in a linguistic sense; for an argument in favour of decoupling these two terms see Sneddon (2000).

Pannell and Glenn (2000) suggest that the stimulus for developing sustainability indicators was the fact that sustainability cannot be condensed to a single simple definition. Of equal importance was Chapter 40 of Agenda 21 which highlighted the need for a more systematic approach to the identification and utilisation of sustainable development indicators in order to ascertain whether or not development was becoming more, or less, sustainable over time. This call for action stimulated and continues to stimulate a myriad of research ventures, at different scales, within the broad area of indicators of sustainable development.

Indicators of sustainable development need to account for and address the fact that there are economic, environmental, social and institutional dimensions to sustainability. The number and array of indicators that have been proposed since the publication of Agenda 21 reflects these different dimensions of sustainable development with the exception of indicators covering the institutional dimension of sustainability which are currently hampered by methodological issues and a lack of data. Different criteria have been suggested to guide the selection of indicators of sustainable development (e.g. Hardi and Zdan, 1997) to try and ensure that a given indicator set is: adequate given the time-frame and spatial-scale under consideration; reflective of the different dimensions of sustainability; sensitive to intergenerational and intragenerational concerns. To this end, indicator frameworks have been employed to systematize the selection of indicators so that important elements are covered by the indicator set (e.g. Bossel, 1999).

Composite indices of sustainable development are, superficially at least, an attractive option relative to simple, non-aggregated, indicators of sustainable development; if the state of the world can be represented by a few key numbers why would this not be preferred to a longer list of numbers? In this paper we provide: a brief overview of composite indices of sustainable development (Sect. 2); a more specific discussion of those composite indices which have been formulated to measure well-being (Sect. 3); an examination of what an information theoretic approach to composite indices of sustainable development tells us in terms of the appropriateness of such indicators using a well-known indicator of well-being—the United Nation’s Human Development Index—as an illustrative example (Sect. 4); conclusions from this analysis (Sect. 5).

1.1 Composite indices of sustainable development: overview

A composite index is an aggregation of individual indicators which can be weighted to reflect the relative importance of each indicator (Nardo, Saisana, Saltelli, Tarantola, 2005a). According to Stevens (2005), the rationale for developing and using such composite indices to inform public policy is that they integrate a mass of information into easily understood formats for a general audience. Stevens (2005) also notes how “... their construction is not straightforward, they can provide misleading information...”. Similarly, Bossel (1999) notes how composite indices can hide serious deficits. To elaborate, a composite index could show positive increases over time suggesting that development is becoming more sustainable but this aggregate rise could mask declines in some components of the index which is

obviously converse to the notion of sustainability. Not surprisingly therefore, the development and use of composite indices of sustainable development has proponents and opponents.

1.2 Composite indices of sustainable development: well-being

There has been a significant research effort internationally to define and operationalise measures of well-being. This is at least partly due to the broad nature of the concept and thus the wide array of interpretations that can be placed upon it. In practice, measures of well-being tend to concentrate explicitly on economic and social dimensions; many also incorporate the environmental dimension. Here we focus on the most commonly used in the literature. Three essentially identical indicators involve the estimation of a range of economic, social and environmental benefits in monetary terms but with different names: the Genuine Progress Indicator (GPI; Cobb, Halstead, & Rowe, 1995); the Index of Sustainable Economic Welfare (ISEW; Daly & Cobb, 1989); and the Sustainable Net Benefit Index (SNBI; Lawn & Sanders, 1999). A fourth indicator, the United Nations Development Programme's Human Development Index (HDI; UNDP, 1990) aggregates life expectancy, adult literacy combined with years of education (hereafter education) and GDP per capita into one measure. In terms of the first three measures, three identical indices which go by different names are far from ideal and can cause confusion. Healy and Côté (2001), for example, describe the GPI and ISEW as distinct indices which they are not. Further research into these indices needs to converge upon an agreed nomenclature for what is being measured. Lawn (2006), who despite coining the SNBI nomenclature with a colleague in 1999 now opts for "the Genuine Progress Indicator as the best name so far devised". However, Herman Daly who co-coined the ISEW term in 1989 is less positive about the GPI name: "Index of Sustainable Economic Welfare has the advantage of being the most explicitly descriptive of what it is [...]; Sustainable Net Benefit Index is a close second; and Genuine Progress Indicator is a distant third, in my opinion. But for appeal to the general public maybe Genuine Progress Indicator is simpler" (H. Daly, Pers. Comm. 12 May 2006). We would argue that, to the academic, ISEW is the most appropriate nomenclature because welfare is a well-established concept in economics whereas the actual meaning of the *net benefits* in the SNBI and the *genuine progress* in the GPI are not immediately apparent. However, non-academics may have a semantic problem with the term welfare, associating it with financial aid from the government for example—particularly in an American context. Therefore, although ISEW is perhaps the most appropriate name from an academic viewpoint, all three names are problematic from the perspective of wider stakeholders.

There is a substantial literature on these three measures and a controversy surrounding their use. Atkinson (1995) and Neumayer (1999) for example offer methodological critiques which are answered by Lawn (2006) in terms of "illuminating a sound theoretical foundation" for these measures. However, because these three measures are expressed in monetary units their accuracy depends upon the quality of valuation methods used for this purpose which Lawn (2006) calls into question. Therefore, even if the various methodological critiques regarding the

measures themselves do not stand up to scrutiny, their dependence upon problematic valuation methodologies is not in doubt.

Since its proposition in 1990, the HDI has gained noteworthy prominence. HDIs are published annually in UN Human Development Reports (HDRs) which, according to Sagar and Najam (1999), represent “the flagship publication not only of the UNDP, but possibly of the entire UN system”. Not surprisingly, given the prominence of this measure to the UN, it has received significant attention in the literature. The nomenclature—Human Development Index—does not appear to raise obvious semantic problems within the context of sustainable development (however outside of this context, human development could be taken to mean the evolution of *Homo sapiens*) and because it does not depend on the valuation of non-market goods or services it avoids the problems associated with the three measures of well-being discussed above. This aside, critiques of the methodology abound questioning the quality of the data employed (e.g. Murray, 1991), the failure of the indicator to acknowledge within country differences (Sagar and Najam, 1998) and the method used to aggregate the indicator’s three components (Srinivasan, 1994; Booyesen, 2002). However, these issues are not unique to the HDI and could equally be levied against an array of composite indices. The HDI is also problematic because the methodology used for its calculation has evolved since its inception in 1990 making comparisons over time difficult (Morse, 2003). The UN is explicit about the shortcomings of the HDI and, therefore, implicit in its continued use is the statement that the index is at least as good as other related composite indices of sustainable development.

1.3 The Human Development Index: an information theoretic analysis

Occam’s Razor is a principle that is often invoked in scientific research to justify economy, simplicity and parsimony in the development of new theories. Expressed as the Law of Succinctness (*lex parsimoniae*) Occam’s Razor states that entities should not be multiplied beyond necessity. While such a principle seems to be in tune with the rationale for the formulation of composite indices where the main aim is to simplify the state of the world into a few key numbers (as opposed to having to refer to an array of simple non-aggregated indicators) the reality may be somewhat different. In fact, composite indices may be further removed from reality compared to simple indicators because of: (1) assumptions about the functional form used to combine different indicators; (2) assumptions about any weights used to prioritise different indicators within a given functional form (Fig. 1). This would appear to suggest that composite indices could introduce additional complexity, which is not associated with simple indicators.

Information Theory is a branch of applied mathematics that has been applied to problems in many fields over the last 50 years such as electrical engineering, physics and psychology. More recently it has been applied within the environmental sciences to determine, for example, whether more complex models of terrestrial nutrient flux are justified compared to simpler, nested alternatives (Stapleton et al., 2006). Specifically, instead of comparing models on the basis of their goodness-of-fit to a dataset, statistics rooted in Information Theory can be used which include a goodness-of-fit component whilst penalising complexity. The Root Mean Squared Deviation (RMSD), the Akaike Information Criterion (AIC; Akaike, 1974) and the

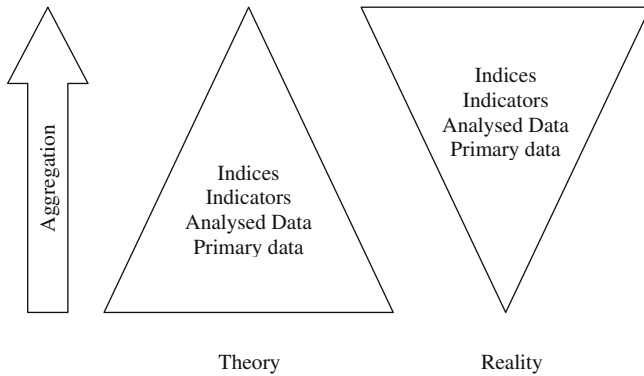


Fig. 1 The sustainable development information pyramid (based on Segnestam, 2002)

Schwarz Information Criterion (SIC; Schwarz, 1978) take into account the number of adjustable (free) parameters in a given model. The Minimum Description Length (MDL; Rissanen, 1987) and the Information-Theoretic Measure of Complexity (ICOMP; Bozdogan, 1990) take into account the number of adjustable model parameters in a model as well as the complexity of its functional form. Therefore, a model which is selected using one or more of these statistics (Eq. 1) could be described as parsimonious (striking a better balance between goodness-of-fit and complexity) relative to alternative models.

$$\text{RMSD} = \sqrt{\text{SSE}/(n - p)}$$

$$\text{AIC} = -2 \log(\text{ML}) + 2p$$

$$\text{BIC} = -2 \log(\text{ML}) + p \log(n)$$

$$\text{MDL} = -\log(\text{ML}) + 0.5\log|H(\theta)|$$

$$\text{ICOMP} = -\log(\text{ML}) + 0.5p \log \left[\frac{\text{trace}(\Omega(\theta))}{p} \right] - 0.5 \log |\Omega(\theta)| \quad (1)$$

where ML = the maximised likelihood function, p = number of adjustable parameters, n = number of data points, SSE = model sum of squares error, $H(\theta)$ = Hessian matrix of the likelihood, $\Omega(\theta)$ = Covariance matrix of parameter estimates.

Composite indices of sustainable development could be regarded as models but models which fit the data perfectly; it's not possible to go out into the field and measure these theoretical constructs directly in order to determine whether such models are an accurate representation of reality. The existing HDI attaches equal weights to its three components and although a survey of 1,547 researchers recently concluded that this was optimal relative to differentiating these weights, this equal weighting approach has been heavily criticised (Chowdhury & Squire, 2006). Instead

of an equal weighting approach let us assume that the true HDI lies about unequal coefficients. A HDI based on equal coefficients will necessarily have a lower goodness-of-fit to this true case but differentiating such coefficients adds complexity to the HDI model so that this improvement in goodness-of-fit might not justify this additional complexity according to model selection statistics. Assume therefore:

$$\text{Model 1 [current HDI]} = (\alpha \cdot \text{Life Expectancy Index}) + (\alpha \cdot \text{Education Index}) \\ + (\alpha \cdot \text{GDP Index})$$

where

$$\alpha = 0.333$$

$$\text{Index} = (\text{Value for Country} - \text{Minimum across all Countries}) / \text{Range} \quad (2)$$

$$\text{Model 2 [more complex HDI]} = (\alpha_1 \cdot \text{Life Expectancy Index}) \\ + (\alpha_2 \cdot \text{Education Index}) + (\alpha_3 \cdot \text{GDP Index})$$

where

$$\alpha_1, \alpha_2, \alpha_3 \text{ are all adjustable under model fitting subject to: } \alpha_1 + \alpha_2 + \alpha_3 = 1;$$

$$\text{HDI} < = 1 \quad (3)$$

Assume also that there are two HDI datasets calculated as follows from the life expectancy, education and GDP indices which constituted the HDIs for 177 countries in 2003:

$$\text{HDI Hypothetical Datasets} = (\alpha_1 \cdot \text{Life Expectancy Index}) \\ + (\alpha_2 \cdot \text{Education Index}) + (\alpha_3 \cdot \text{GDP Index})$$

where

$$\left. \begin{array}{l} \alpha_1 = 0.5 \pm \text{rand } 10\% \\ \alpha_2 = 0.3 \pm \text{rand } 10\% \\ \alpha_3 = 0.2 \pm \text{rand } 10\% \end{array} \right\} \text{Dataset 1 (somewhat different coefficients)}$$

$$\left. \begin{array}{l} \alpha_1 = 0.05 \pm \text{rand } 10\% \\ \alpha_2 = 0.85 \pm \text{rand } 10\% \\ \alpha_3 = 0.1 \pm \text{rand } 10\% \end{array} \right\} \text{Dataset 2 (highly different coefficients)} \quad (4)$$

For Dataset 1, Model 2 provides a closer fit compared to Model 1 in terms of R^2 and RSS where the values of α_1 , α_2 and α_3 are optimised using the Marquardt method (Press, Teukolsky, Vetterling, & Flannery, 2002) in order to minimise the deviation between prediction and observation (Table 1, Fig. 2) and because 177 (number of data points) $< \infty$, optimised coefficients in Eq. (3) converge around, but not absolutely to those specified in Eq. (4). Only the RMSD and ICOMP favour Model 2 over Model 1; the AIC, SIC and ICOMP favour Model 1 i.e., the extra complexity of Model 2 does not justify the increased goodness-of-fit of this model. More interestingly, the same pattern of results occurs for the more divergent dataset, Dataset 2 (Table 1, Fig. 3): Model 2 outperforms Model 1 in terms of the goodness-of-fit statistics R^2 and RSS but the same three out of five model selection statistics (AIC,

Table 1 Goodness of fit and model selection statistics where hypothetical data were generated around the 2003 Human Development Index (HDI) and then tested against the existing HDI (Model 1) and a more complex alternative with adjustable coefficients (Model 2)

Model	Data	Parameters	SEs(±)	Data points	Degrees freedom	RSS	R ²	RMSD	AIC	SIC	MDL	ICOMP
1	1	$\alpha = 0.333$	0.001	177	176	0.206	0.965	0.035	2.215	5.391	3.477	0.107
2	1	$\alpha_1 = 0.510$ $\alpha_2 = 0.302$ $\alpha_3 = 0.186$	0.016 0.015 0.018	177	174	0.115	0.98	0.026	6.115	15.64	3.415	2.026
1	2	$\alpha = 0.333$	0.003	177	176	1.491	0.751	0.088	3.37	6.546	4.054	0.685
2	2	$\alpha_1 = 0.057$ $\alpha_2 = 0.853$ $\alpha_3 = 0.087$	0.024 0.023 0.028	177	174	0.278	0.956	0.039	6.278	15.8	3.496	2.107

Standard Error (SE); Residual Sum of Squares (RSS); Root Mean Squared Deviation (RMSD); Akaike Information Criterion (AIC); Schwarz Information Criterion (SIC); Minimum Description Length (MDL); Information-Theoretic Measure of Complexity (ICOMP). The shaded line is used to distinguish models applied to different datasets: goodness of fit and model selection statistics cannot be applied between datasets. All statistics are unitless. The model which maximises R² and minimises RSS, RMSD, SIC, MDL and ICOMP should be chosen

SIC and ICOMP) favour Model 1. Furthermore, if we acknowledge that RMSD is an informal measure with no statistical justification ever posited for its use as a model selection determinant (Myung, 2000) then the preference for Model 1 over Model 2 in terms of model selection criteria becomes even more pronounced. It may seem

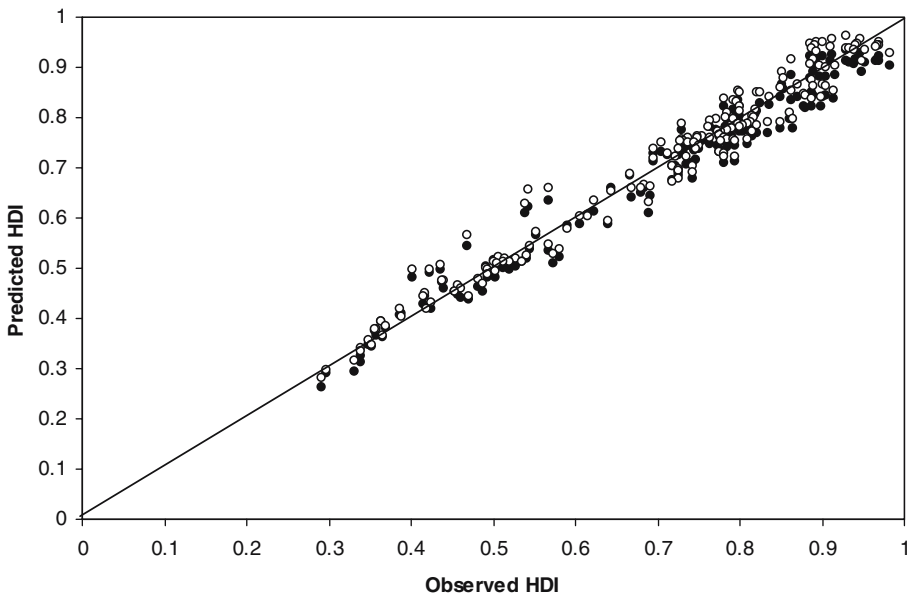


Fig. 2 Prediction versus observation where a hypothetical dataset (Dataset 1) is generated for testing the 2003 HDI. The solid diagonal is a 1–1 line where prediction and observation are equal. ○ shows how the HDI (Model 1) compares to this dataset ● shows how an optimised HDI with adjustable coefficients (Model 2) compares to this dataset

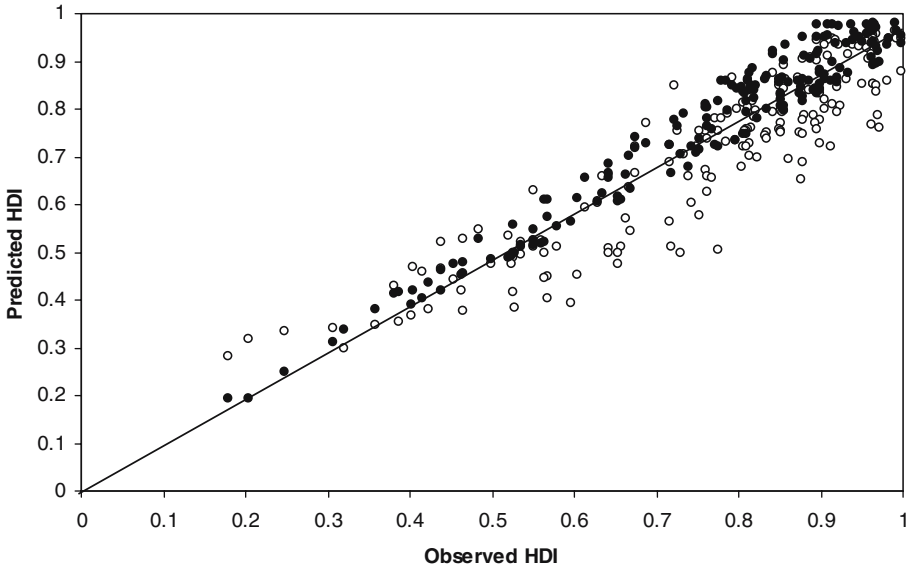


Fig. 3 Prediction versus observation where a hypothetical dataset (Dataset 2) is generated for testing the 2003 Human Development Index (HDI). The solid diagonal is a 1–1 line where predication and observation are equal. ○ shows how the HDI (Model 1) compares to this dataset ● shows how an optimised HDI with adjustable coefficients (Model 2) compares to this dataset

Table 2 Pearson’s correlation matrix for the life expectancy, education and GDP indicators used to calculate the Human Development Index for 177 countries in 2003

	Life Expectancy	Education	GDP
Life Expectancy	1		
Education	0.729	1	
GDP	0.767	0.760	1

counter-intuitive that Model 2 isn’t generally selected over Model 1 in the case of Dataset 2 given that this dataset diverges very significantly from the equal weights assumption; this can be explained in terms of the fact that the three variables (indicators) which constitute the HDI are highly collinear (Table 2). If this were not the case then model selection criteria would be less likely to favour the simpler equal weights model (Model 1).

2 Conclusion

Although it is not possible to go out into the field and directly measure the HDI and conduct the kind of statistical analysis outlined above, the nature of the results have important practical implications for the HDI. The nature of the data generated in this work suggests that alternative HDIs with different coefficients provided a better goodness-of-fit to these datasets when compared to the current HDI which attaches equal weights to life expectancy, education and GDP. However, attaching such equal weights implies only one parameter compared to three parameters if the weights

associated with life expectancy, education and GDP are differentiated. Statistics rooted in Information Theory suggest that, even if the true weights are significantly different from each other, there is a lack of justification for acknowledging this in the functional form of the HDI, not least because the three variables (indicators) which constitute the HDI are collinear. Put differently, although the sensitivity of composite indicator outputs to different changes associated with the construction of such indicators has been examined previously (e.g. Morse, 2003; Nardo et al., 2005a; Nardo, Saisana, Saltelli, Tarantola, Hoffman, & Giovannini, 2005b) the work presented here goes further by illustrating that if such changes increase the complexity of the indicator under consideration then the additional assumptions this brings forth may not be parsimonious relative to simpler alternatives.

Acknowledgements The authors are working on a European Union Sixth Framework Project (System for Environmental and Agricultural Modelling: Linking European Science and Society; SEAMLESS) which helped provide inspiration for this work.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, pp. 716–723.
- Atkinson, G. (1995). Measuring Sustainable Economic Welfare: A Critique of the UK ISEW. Centre for Social and Economic Research on the Global Environment Working Paper, Norwich and London.
- Booyens, F. (2002). An overview and evaluation of composite indices of development. *Social Indicators Research*, 59, pp. 115–151.
- Bossel, H. (1999). *Indicators for sustainable development: Theory, method, applications*. Manitoba: International Institute for Sustainable Development.
- Bozdogan, H. (1990). On the information-based measure of covariance complexity and its application to the evaluation of multivariate linear models. *Communications in Statistics, Theory and Methods*, 19, pp. 221–278.
- 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, 761–771.
- Cobb, C. Halstead, T., & Rowe, J. (1995). *The genuine progress indicator: Summary of data and methodology*. San Francisco: Redefining Progress.
- Daly, H., & Cobb, J. (1989). *For the common good*. Boston: Beacon Press.
- Hardi, P., & Zdan, T. (1997). *Assessing sustainable development: Principles in practice*. Winnipeg: International Institute for Sustainable Development.
- Healy, T., & Côté, S. (2001). *The well-being of nations: The role of human and social capital*. Paris: Centre for Educational Research and Innovation, Organisation for Economic Co-operation and Development.
- Lawn, P. (2006). An assessment of alternative measures of sustainable economic welfare. In P. Lawn (Ed.), *Sustainable development indicators in ecological economics*, (pp. 139–165). Cheltenham: Edward Elgar.
- Lawn, P., & Sanders, R. (1999). Has Australia surpassed its optimal macroeconomic scale: finding out with the aid of “benefit” and “cost” accounts and a sustainable net benefit index. *Ecological Economics*, 28, 213–229.
- Morse, S. (2003). For better or for worse, till the human development index do us part? *Ecological Economics*, 45, 281–296.
- Murray, C. J. L. (1991). Development Data Constraints and the Human Development Index, Discussion Paper, United Nations Research Institute for Social Development, Geneva.
- Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology*, 44, 190–204.

- Nardo, M. Saisana, M. Saltelli, A., & Tarantola, S. (2005a). *Tools for composite indicator building*. Ispra: Institute for the Protection and Security of the Citizen, European Commission Directorate-General Research Centre.
- Nardo, M. Saisana, M. Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini E. (2005b) *Handbook on constructing composite indicators: Methodology and user guide*. Paris: Organisation for Economic Co-operation and Development.
- Neumayer, E. (1999). The ISEW—not an index of sustainable economic welfare. *Social Indicators Research*, 48, 77–101.
- Pannell, D. J., & Glenn, N. A. (2000). A framework for the economic evaluation and selection of sustainability indicators in agriculture. *Ecological Economics*, 33, 135–149.
- Press, W. H. Teukolsky, S. A. Vetterling, W. T., & Flannery, B. P. (2002). *Numerical recipes in C++*. Cambridge: Cambridge University Press.
- Rissanen, J. (1987). Stochastic complexity and the MDL principle. *Econometric Reviews*, 8, 85–102.
- Sagar, A. D., & Najam, A. (1998). The human development index: a critical review. *Ecological Economics*, 25, 249–264.
- Sagar, A. D., & Najam, A. (1999). Shaping human development: which way next? *Third World Quarterly*, 20, 743–751.
- Schwarz, G. (1978). Investigating the dimension of a model. *The Annals of Statistics*, 6, 461–464.
- Segnestam, L. (2002). Indicators of Environment and Sustainable Development: Theories and Practical Experience. Paper No. 89, Environmental Economic Series, World Bank, Washington DC.
- Sneddon, C. S. (2000). ‘Sustainability’ in ecological economics, ecology and livelihoods: a review. *Progress in Human Geography*, 24, 521–549.
- Srinivasan, T. N. (1994). Human development: a new paradigm or reinvention of the wheel? AEA Papers and Proceedings 84, pp. 238–43.
- Stapleton, L. M., Laybourn-Parry, J., Poulton, P. R., Tye, A. M., West, H. M., Young, S. D., & Crout, N. M. J. (2006). Parsimonious modelling of nutrient fluxes for a terrestrial ecosystem on Svalbard. *Biogeochemistry*, 80, 57–69.
- Stevens, C. (2005). *Measuring sustainable development*. Paris: Organisation for Economic Co-operation and Development.
- United Nations Conference on Environment and Development. (1992). *Agenda 21*. New York: United Nations Department of Economic and Social Affairs.
- United Nations Development Programme. (1990). *Human development report*. New York: Oxford University Press.
- World Commission on Environment and Development. (1987). *Our Common Future*. Oxford: Oxford University Press.