

Böhringer, Christoph; Jochem, Patrick

Working Paper

Measuring the immeasurable: a survey of sustainability indices

ZEW Discussion Papers, No. 06-073

Provided in Cooperation with:

ZEW - Leibniz Centre for European Economic Research

Suggested Citation: Böhringer, Christoph; Jochem, Patrick (2006) : Measuring the immeasurable: a survey of sustainability indices, ZEW Discussion Papers, No. 06-073, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim

This Version is available at:

<https://hdl.handle.net/10419/24527>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Discussion Paper No. 06-073

**Measuring the Immeasurable:
A Survey of Sustainability Indices**

Christoph Böhringer and Patrick Jochem

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 06-073

Measuring the Immeasurable: A Survey of Sustainability Indices

Christoph Böhringer and Patrick Jochem

Download this ZEW Discussion Paper from our ftp server:

<ftp://ftp.zew.de/pub/zew-docs/dp/dp06073.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.

Nontechnical Summary

National sustainability indices provide a one-dimensional metric to evaluate country-specific information on the three holistic dimensions of sustainable development: economic, environmental, and social conditions. At the policy level, they suggest an unambiguous yardstick against which a country's development can be measured and even a cross-country comparison can be performed. However it remains questionable whether we can meaningfully speak from unambiguity in the holistic subject of sustainability development.

This paper reviews the consistency and meaningfulness of eleven sustainability indices that are widely used in policy practice: the Living Planet Index (LPI), Ecological Footprint (EF), City Development Index (CDI), Human Development Index (HDI), Environmental Sustainability Index (ESI), Environmental Performance Index (EPI), Environmental Vulnerability Index (EVI), Index of Sustainable Economic Welfare/Genuine Progress Index (ISEW/GPI), Well-Being Index (WI), Genuine Savings Index (GS), and Environmental Adjusted Domestic Product (EDP).

We find that – although the sustainability indices are imputed to be concise and transparent – they fail to meet fundamental scientific requirements with respect to the three central steps of indices formation: normalization, weighting, and aggregation. Normalization and weighting of indicators – which in general are associated with subjective judgments – reveal a high degree of arbitrariness without mentioning or systematically assessing critical assumptions. As to aggregation, there are scientific rules which could guarantee consistency and meaningfulness of composite indices. Yet, these rules are often not taken into account. As a consequence, sustainability indices currently employed in policy practice are doomed to be useless if not misleading with respect to concrete policy advice.

Measuring the Immeasurable – A Survey of Sustainability Indices

Christoph Böhringer^{a,b} and Patrick E.P. Jochem^c

^a Centre for European Economic Research, P.O. Box 103443, 68034 Mannheim, Germany.

^b University of Heidelberg, Department of Economics, Heidelberg, Germany.

^c Scholarship student of the DBU at Institute for Economic Policy at University of Karlsruhe,
P.O. Box 6980, 76128 Karlsruhe, Germany.

Abstract. Sustainability indices for countries provide a one-dimensional metric to evaluate country-specific information on the three dimensions of sustainable development: economic, environmental, and social conditions. At the policy level, they suggest an unambiguous yardstick against which a country's development can be measured and even a cross-country comparison can be performed. This paper reviews the explanatory power of various sustainability indices applied in policy practice. We show that these indices fail to fulfill fundamental scientific requirements making them rather useless if not misleading with respect to policy advice.

Key Words: Sustainability Indices, Composite Indicators, Sustainability, Indices

JEL classification: C43, Q01

1 Introduction

Sustainable development has become one of the most popular catchwords on the world's policy agenda. Nearly all governments have committed themselves to sustainable development by integrating economic welfare, environmental quality and social coherence. As a consequence, there is a strong political desire for the comprehensive assessment of changes in economic, environmental, and social (including institutional) conditions: An issue that can not be clearly measured will be difficult to improve.

Monitoring progress towards sustainable development (hereafter: SD) requires in first place the identification of operational indicators that provide manageable units of information on economic, environmental, and social conditions. The central role of SD indicators has already been emphasized by the United Nations Conference on Environment and Development (UNCED), held in Rio de Janeiro in 1992, that calls on individual countries as well as international governmental and non-governmental organizations to “develop and identify indicators of SD in order to improve the information basis for decision-making at all levels” (UNCED, 1992, Agenda 21, Chapter 40). Since the early 90ies a multitude of indicator lists has been developed. The Compendium of Sustainable Development Indicator Initiatives mentions more than 500 sustainable indicator efforts (Parris and Kates, 2003).

For policy practice, the variety of SD indicators poses a huge problem, especially since policy makers demand an aggregate index¹ that can be unambiguously interpreted and easily communicated to the general public (Dalal-Clayton and Bass, 2002, p. 135, and Hammond et al., 1995, p. 2). Reflecting this policy demand, the construction of aggregate SD indices have a long history going back to pioneering work on national wellbeing indices by Nordhaus and Tobin (1971), Zolotas (1981), or Osberg (1985).

¹ In the following we use the term “index” in the sense of Alberti and Parker (1991) for both index and composite indicator.

In this paper, we scrutinize eleven widely applied SD indices as to their consistency and meaningfulness: the Living Planet Index (LPI), Ecological Footprint (EF), City Development Index (CDI), Human Development Index (HDI), Environmental Sustainability Index (ESI), Environmental Performance Index (EPI), Environmental Vulnerability Index (EVI), Index of Sustainable Economic Welfare/Genuine Progress Index (ISEW/GPI), Well-Being Index (WI), Genuine Savings Index (GS), and Environmental Adjusted Domestic Product (EDP).

We find that – although the bulk of SD indices are imputed to be concise and transparent – they fail to meet fundamental scientific requirements. There are three central issues to be addressed. Firstly, in selecting input variables one should be conscious that themes determine the thematic aggregation method and units determine the technical aggregation method. Secondly, as there are no general rules for normalization of these variables and their weighting these procedures should be treated in a transparent way with great reserve and be subject to comprehensive sensitivity analysis. Thirdly, commensurability of input variables (Ebert and Welsch, 2004) should be assured. The sustainability indices reviewed in this paper in general do not appropriately handle these issues. Thus we conclude that politically desired sustainability indices are inherently inconsistent and therefore useless if not misleading with respect to concrete policy advice.

The remainder of this paper is organized as follows. In section 2 we discuss central requirements for SD indices to prove scientifically viable and meaningful. We highlight in particular the basic requirements for aggregation as discussed by Ebert and Welsch (2004) and Welsch (2005). It will turn out that fundamental formal conditions for meaningful aggregation are widely neglected even though they could easily be met for some of the indices. In section 3 we describe eleven SD indices used in policy practice. In section 4, we evaluate these indices in view of the requirements for SD indices developed before. In section 5 we conclude.

2 Requirements for Sustainable Development Indices

In the literature, the criteria for selecting appropriate SD indicators which form the basis for any SD index have been widely discussed (see e.g. Hodge and Hardi, 1997, OECD, 1994, Atkinson et al., 1997, and Radke, 1999). Key requirements include (i) the rigorous connection to the definitions of sustainability (Pezzey, 1992, pp. 55 et sqq.), (ii) the selection of meaningful indicators (representing holistic fields) which should not be highly correlated, (iii) reliability and availability (measurability) of data for quantification over longer time horizons (Ramachandran, 2000, or Stehling, 1988), (iv) process orientated indicator selection (SRU, 1994, p. 87, Radke 1999, pp. 183 et sqq.) as well as (v) the possibility of deriving political (sub) objectives (Esty et al., 2006).

In this paper we will focus on additional requirements for SD indices such as adequate normalization, aggregation, and weighting of the underlying variables (Cash et al., forthcoming, Parris and Kates, 2003, and Moldan et al., 2004). Normalization is usually applied to single variables² in order to make them comparable, i.e. transforming the various scales of variables into one unique scale. The normalized indicators are then aggregated using specific formulas (e.g. arithmetic mean). If one indicator is more “important” than another, the former is assigned a stronger weight than the latter within the aggregation procedure.

Scientifically sound methods for normalization (to make data ‘comparable’), weighting (to specify the ‘correct’ interrelationships), and aggregation (to get the ‘right’ functional relationship) are obviously pre-requisite for the construction of meaningful SD indices. However, as pointed out by Nardo et al. (2005, p. 21), already the normalization of data implies a value judgment, as different scales could not be harmonized in an meaningful manner. The same applies to weighting since this involves potentially normative ‘quotas of substitution’ (Freudenberg, 2003). For example, if the SD index weights income per capita (in

² Following Ott (1978) we refer to the normalized variables as indicators.

billion US\$) twice as much as emissions of carbon dioxide (in Mtons with negative sign), the SD level remains constant when a country increase per-capita-income by half a billion US\$ and thereby also increases its carbon dioxide emissions by one Mton. Obviously, such a metric must fail when we account for maximum thresholds of carbon dioxide concentrations in the atmosphere. In essence, the derivation of weights often does not comply with scientific criteria.

While there are no unambiguous rules for selecting variables, normalization, and weighting, which are matters of substance to be decided by natural science or/and policy, Ebert and Welsch (2004) have laid out precise requirements regarding the meaningful aggregation of (commensurable) variables, which is a methodological matter. Thus – if variables are selected and weights are given – Ebert and Welsch present meaningful aggregation methods for these variables (without normalization). Meaningfulness is build on the notion that a sustainability index must allow unambiguous orderings of the relevant world states over time independent of the measurement units in which the variables describing the world states are expressed (Welsch, 2005, p. 7): If we consider for example the index of two different toxic emissions at different points in time, a meaningful index should deliver the same ordinal assessment even if the units of one index are altered (e.g. Kg to tons) provided that the underlying emissions remain constant (“commensurability problem”). An index is meaningful if the ordering which is represented is unaffected by the way the commensurability problem is addressed (given approved variables and weights). In their formal analysis, Ebert and Welsch (2004) derive feasible aggregation procedures for variables depending on the measurement scales and the desired properties of the index. As to scales, a distinction between interval scales and ratio scales is made: Interval scales do not have any natural zero point unlike ratio scales, therefore ratios are not meaningful in interval scales.³ The comparability of scales means that the

³ The different scales of temperature (with Kelvin as an exception) do not have a natural zero point, thus the statement “today it is twice as hot as yesterday” makes scientifically no sense.

technical (natural) relationships of every indicator to be aggregated are known and constant.⁴ Ebert and Welsch then identify four generic classes of scales that can be applied to variables: interval-scale non-comparability (INC), interval-scale full comparability (IFC), ratio-scale non comparability (RNC), and ratio-scale full comparability (RFC). Furthermore, they discuss basic mathematical properties for an index – continuity, (strong/weak) monotonicity and separability – the desirability of which depends on the specific circumstances of the analysis and the required robustness.

Table 1 provides an overview of which functional forms for the aggregation of variables are viable depending on their scales (and the desired mathematical properties): If interval-scaled variables are not meaningfully comparable (INC), they can not be meaningfully aggregated except for a dictatorial ordering towards a continuous and weak monotone index. If interval-scaled variables are comparable (IFC), aggregation based on an arithmetic mean is possible to achieve a continuous, strongly monotone, and separable index. In general, variables which are measured on a ratio scale provide more flexibility for meaningful aggregation: In the case of non-comparability (RNC), a geometric mean can be applied to yield a continuous, strongly monotone, and separable index; in case of comparability (RFC), any homothetic function can be used to deliver a mathematically meaningful index meeting the requirements of continuity and strong monotonicity.⁵ Furthermore, if different scales should be aggregated (e.g. amount of pollutant and temperature), it is impossible to aggregate them in a meaningful way.

The aggregation rules of Table 1 provide minimal methodological requirements to be met by any meaningful SD index. However, as will be laid out in the following section, indices applied in practice typically violate these qualifying conditions: Whereas the aggregation of

⁴ Comparability is not ensured if there is no scientific and prior relationship between the indicators (as e.g. different air polluting substances as CO₂, NO_x, and particles).

⁵ If the index for ratio-scaled comparable indicators should also confer with separability, a constant-elasticity-of-substitution (CES)-type function provides a suitable aggregation method.

variables measured in ratio-scale without being comparable would call for a geometric mean, indices are often based on a (misleading) arithmetic mean.

Table 1: Aggregation rules for variables by Ebert and Welsch

	<u>Non-Comparability</u>	<u>Full comparability</u>
<u>Interval scale</u>	Dictatorial ordering	Arithmetic mean
<u>Ratio scale</u>	Geometric mean	Any homothetic function ⁵

Beyond requirements of formal consistency, there are additional – rather pragmatic – considerations on the operational use of SD indices for policy making. Since the broad majority should accept an index on which policy reforms might be based on, this index should be at least sufficiently transparent in composition (Hammond et al., 1995, Kuik and Gilbert, 1999, Bellagio-Principles, and Jesinghaus, 1999).

3 Survey of Sustainability Indices

Table 2 provides a short characterization of eleven SD indices which are widely used in the SD policy debate. After a brief summary of the indices we will evaluate them against basic requirements for meaningfulness.

Table 2: Characteristics and source of SD indices

<u>Index</u>	<u>Reference</u>	<u>Countries</u>	<u>Variables</u>
Living Planet Index (LPI)	WWF (1998)	n.a. ¹	1100
Ecological Footprint (EF)	Wackernagel and Rees (1997)	148 ²	arbitrary
City Development Index (CDI)	UNCHS (2001)	125 ³	11
Human Development Index (HDI)	UNDP (2005)	177	4
Environmental Sustainability Index (ESI)	Esty et al. (2005)	146	76
Environmental Performance Index (EPI)	Esty et al. (2006)	133	16
Environmental Vulnerability Index (EVI)	SOPAC (2005)	235	50
Index of Sustainable Economic Welfare (ISEW) ⁴	Cobb (1989)	6	25
Well Being Index (WI)	Prescott-Allen (2001)	180	87
Genuine Savings Index (GS)	Hamilton et al. (1997)	104 ⁵	5
Environmental Adjusted Domestic Product (EDP)	e. g. Hanley (2000)	n.a. ⁶	(many)

¹ The LPI measures the number of individuals of a specific species in a certain population (which are not restricted to national borderlines).

² Based on the LPI-Report 2004 (WWF, 2004).

³ The CDI has been applied to cities, regions, and countries.

⁴ Identical with the Genuine Progress Index (GPI).

⁵ See “Little Green Data Book 2005” (Worldbank, 2005).

⁶ The number of countries that are implementing (different kinds of) SEEA to calculate an EDP has been rapidly growing during the last years.

Living Planet Index (LPI)

The global biodiversity indicator Living Planet Index was developed by WWF (1998). It measures trends in over 2000 populations of more than 1100 species of vertebrates in terrestrial, freshwater, and seawater ecosystems⁶. The LPI provides a sub-index for the three spheres: For every species within a sphere, the ratio between its populations in pairs of consecutive years is calculated. The geometric mean of these quotients of different species multiplied with the index value of the former year then delivers the biodiversity index for the respective sphere (1970 serves as a base-year with the index value for 1970 scaled to unity). The geometric mean of these indices is the LPI. As all variables are in changes of numbers of species no normalization is accomplished and all ratios are equally weighted.

Ecological Footprint (EF)

The Ecological Footprint (EF) (Wackernagel and Rees, 1997) is based on the quantitative land and water requirements to sustain a (national) living standard into infinity thereby assuming certain efficiency improvements. The ratio of required resources to available resources is interpreted as a measure of ecological sustainability: Ratios exceeding one are seen as unsustainable, i.e. contemporary living standards would violate the principles of sustainable development. Calculation of the EF is based on data from national consumption statistics. Thus, the EF primarily relies on normalization (as any consumption is converted in land use). Weighting is rather implicit in the conversion parameter and aggregation is done by adding up all land and water requirements. There are several approaches similar to the EF, e.g. the MIPS (Material-Input-Per-Service) concept (Schmidt-Bleek, 1994), the Sustainable Process Index (Narodoslawsky and Krotscheck, 2004; Gassner and Narodoslawsky, 2004) or the EcoindexTM (Chambers and Lewis, 2001).

⁶ If species live simultaneously in different ecosystems, the breeding place is chosen as the allocation criteria. This permits a direct and unambiguous allocation of every species to an ecosystem (WWF, 2004).

City Development Index (CDI)

The City Development Index (CDI) suggested by the United Nations Centre for Human Settlements (HABITAT) consists of five sub-indices: (i) an infrastructure index, which builds on four (equally weighted) indicators as percentages of households which are connected to clean water, canalization, electricity and a phone network (without mobiles), (ii) a twofold (equally weighted) waste index, which is composed of the percentage of untreated sewage in total wastewater and the percentage of disposal of solid waste in total solid wastes, (iii) a twofold (diversely weighted) health index, which considers the life expectancy and the infant mortality rate (iv) a twofold (equally weighted) education index which is calculated by adding the percentages of literacy and combined enrolment; and (v) a city product index, which is based on the logarithmic value of a the city's GDP. The CDI employs a common normalization procedure for variables where actual values x_i are transformed according to $(x_i - \underline{x})/(\bar{x} - \underline{x})$ with $\bar{x}(\underline{x})$ denoting exogenous maximum (minimum) target values. Weights for the indicators are based on a multivariate statistical technique (Principal Components Analysis, see e. g. Dunteman, 1989). The five sub-indices are aggregated towards the CDI using an arithmetic mean.

Human Development Index (HDI)

Since 1990 the Human Development Index (HDI) is reported annually as part of the Human Development Report of the United Nations Development Programme (UNDP, 2005). It consists of three (equal weighted) sub-indices which are aggregated by an arithmetic mean: Life Expectancy Index, Education Index (decomposed into an Adult Literacy Index and a Gross Enrolment Ratio Index), and a GNP Index. The HDI has a strong focus on the social dimension of SD. Each sub-index DI_i is calculated as $DI_i = (x_i - \underline{x})/(\bar{x} - \underline{x})$ where x_i denotes the actual value in country i and $\bar{x}(\underline{x})$ refer to exogenous maximum (minimum) values.

(Pilot) Environmental Sustainability Index (ESI)

The “ESI score quantifies the likelihood that a country will be able to preserve valuable environmental resources effectively over the period of several decades” (Esty et al., 2005, p. 23).⁷ The actual ESI 2005 consists of five components which are based on 21 indicators. The 21 indicators are again derived from 76 variables. For normalization the standard deviation is calculated of each (normal distributed) variable. The three aggregation steps consist of arithmetic means with equal weights.

Environmental Performance Index (EPI)

Complementary to the ESI which focuses on the environmental dimension of sustainability, “the EPI addresses the need for a gauge of policy performance in reducing environmental stresses on human health and promoting ecosystem vitality and sound natural resource management. The EPI focuses on current on-the-ground outcomes across a core set of environmental issues tracked through six policy categories for which all governments are being held accountable” (Esty, 2006. pp. 9 et sqq.). The EPI is based on a proximity-to-target approach which measures country performance against an absolute target established by international agreements, national standards, or scientific consensus (Esty, 2006, p. 275). All variables are normalized in a scale from zero to 100. The maximum value of 100 is linked to the target⁸, the minimum value of zero characterizes the worst competitor in the field. Weights are drawn from statistical mechanisms or by consulting experts. Finally, the six policy categories are aggregated to the ESI taking the weighted sum.

⁷ Up to now, the ESI has been calculated three times (ESI, 2001, 2002, and Esty et al., 2005). Since the composition of the indices has been changed from calculation to calculation, it is hardly possible to compare the three rankings on the sustainable performance of countries.

⁸ For each indicator, a policy-relevant long-term goal for public health or ecosystem conditions is identified (drawn from international agreements, standards set by international organizations or national authorities, or prevailing consensus among environmental scientists) (Esty et al., 2006: 9).

Environmental Vulnerability Index (EVI)

The Environmental Vulnerability Index (EVI) comprises 32 indicators of hazards, 8 indicators of resistance, and 10 indicators that measure damage (SOPAC, 2005, p. 7). The EVI scale for normalization ranges between a value of 1 (indicating high resilience / low vulnerability) and 7 (indicating low resilience / high vulnerability). The 50 indicators are given equal weights and then aggregated by an arithmetic mean.

Index of Sustainable Economic Welfare (ISEW) and Genuine Progress Indicator (GPI)

The Index of Sustainable Economic Welfare (ISEW) has been developed by C.W. Cobb (1989) to integrate environmental and social externalities in national welfare accounting. With some modifications to the original accounting method (among others Cobb and Cobb, 1994), the ISEW has been relabeled to the Genuine Progress Indicator (Cobb et al., 1995). Although the ISEW is also calculated for some countries,⁹ these calculations were done by very different institutions and are hardly comparable.

The starting point for the ISEW is the inflation-adjusted consumption of households. The time series of consumption values is adjusted by five categories to obtain a 'GDP' which is more appropriate for measuring social welfare: (i) distribution of income, (ii) economic activities not counted in the conventional gross national income, (iii) time adjustments, (iv) damage caused by economic activity, and (v) the consideration of net capital endowment of foreign investors. As all adjustments are monetarized (normalization and weighting), the sum is used for aggregation.

⁹ Applications of the ISEW include the USA (Cobb, 1989), England (Jackson and Marks, 1994), Austria (Hochreiter et al., 1995), Denmark (Jespersen, 1994), the Netherlands (IMSA, 1995), Italy (Guenno and Tiezzi, 1998), Germany (Diefenbacher, 1995), and Australia (Hamilton, 1999).

Well-Being Assessment (Well-Being Index – WI)

The Well-Being Assessment by Prescott-Allen (2001) is based on the assumption that a healthy environment is necessary for healthy humans. Accordingly, the Well-Being Index (WI) is the arithmetic mean of a Human Well-being Index (HWI) and an Ecosystem Well-Being Index (EWI). The indices HWI and EWI in turn consist of five sub-indices. The HWI comprises a Health and Population, Welfare, Knowledge, Culture and Society, as well as an Equity Index. The EWI comprises indices for land, water, air, species and genes as well as for resources deployment. The five dimensions of the HWI are based on 36 indicators, those of the EWI on 51 indicators. The aggregation of these dimensions is conducted by a weighted¹⁰ arithmetic mean of further sub-indices or variables which are normalized again by a proximity-to-target approach using targets of related indicators (Prescott-Allen, 2001, pp. 298 et sqq.).

Genuine Savings (GS)

Pearce and Atkinson (1993) put forward an index which is based on the Hicksian income concept. In 1997 this index has been enhanced by Hamilton et al. (1997) using the Hartwick rule (Hartwick, 1977) which defines the level of re-investment from resource rents that are reinvested to assure that the (societal) capital stock will never decline.¹¹ The Genuine Savings (GS) are thus an indicator of weak SD. The societal capital stock consists of produced capital, human capital (knowledge, skills etc.) as well as natural capital (resources etc.). As in the ISEW all values are monetarized, such that aggregation is again achieved by simply adding up.

¹⁰ The derivation of weights is not explained in detail.

¹¹ However Asheim et al. (2003) showed that the Hartwick rule is not a measure of sustainability as not all external effects are internalized and thus resource productivity is not represented appropriately. The authors show that an SD in the sense of the “Hartwick rule” is not at all sustainable in practice.

Green Net National Product (EDP) and SEEA

The Green Net National Product or likewise the Environmentally Adjusted Net Domestic Product (EDP) has been developed within the scope of SEEA (System of Integrated Environmental and Economic Accounting – UNEP, 2000 and UN et al., 2003). Following inter alia Hanley (2000) three different versions of the EDP can be distinguished: (i) the EDPI which subtracts depreciations of natural resources caused by their extraction from the net national income (NNI), (ii) the EDPII, which subtracts from the NNI the costs necessary to reach the same state of the environment at the end of the period as existed at the beginning of the period, and (iii) the EDPIII, which subtracts the costs of environmental pressure and destruction (calculated by willingness-to-pay methods). Again aggregation takes place by simply adding up the monetarized values.

Most indicators underlying the aggregate indices are based on variables measured in pieces or weights. Thus, they are measured in ratio scale. Only in the EVI two indicators are measured in Fahrenheit, which is an interval scale. The holistic nature of SD is reflected by the fact that all indices feature (some) incomparable variables. Table 3 summarizes the approaches taken by the various SD indices regarding the central issues of normalization, weighting and aggregation.

Table 3: Methods of SD indices regarding scale, normalization, weighting, and aggregation

<u>Index</u>	<u>Scale</u>	<u>Normalization</u>	<u>Weighting</u>	<u>Aggregation</u>
Living Planet Index	RNC	$\left(\frac{X_{i,t}}{X_{i,t-1}} \right)$	equal	$\sqrt[N]{\prod_{i=1}^N \frac{X_{i,t}}{X_{i,t-1}}}$
Ecological Footprint	RNC	transformation in square km	equal	$\sum_{i=1}^N x_i$
City Development Index	RNC	$\frac{x_i - \underline{x}}{\bar{x} - \underline{x}}$	2 steps PCA/experts	$\frac{1}{N} \sum_{i=1}^N w_i x_i$
Human Development Index	RNC	$\frac{x_i - \underline{x}}{\bar{x} - \underline{x}}$	equal	$\frac{1}{N} \sum_{i=1}^N x_i$
Environmental Sustainability Index 2005	RNC	standard deviation	equal / experts	$\frac{1}{N} \sum_{i=1}^N x_i$
Environmental Performance Index	RNC	best = 100 worst = 0	PCA and experts	$\sum_{i=1}^N w_i x_i$
Environmental Vulnerability Index	RNC / INC	aim = 1 worst = 7	equal	$\frac{1}{N} \sum_{i=1}^N x_i$
Index of Sustainable Economic Welfare	RNC	monetarized	equal	$\sum_{i=1}^N x_i$
Well Being Index	RNC	best = 100 worst = 0	subjective (not derived)	$\frac{1}{N} \sum_{i=1}^N (w_i) x_i$
Genuine Savings Index	RNC	monetarized	equal	$\sum_{i=1}^N x_i$
Environmentally Adjusted Domestic Product (EDP)	RNC	monetarized	equal	$\sum_{i=1}^N x_i$

With variables represented by x_i , weights by w_i , and countries by i and years by t .

4 Evaluation of Sustainability Indices

When we assess the eleven indices with respect to fundamental scientific requirements, we find major shortcomings. Firstly, the selection of SD variables which form the basis for subsequent SD indices are far from reflecting the entire holistic nature of sustainability. For applied analysis, the selection of variables obviously depends on data availability which might improve in the future. In addition, the proper selection of variables can be quite country-specific as sustainability requirements may be viewed differently across countries (Booyesen, 2002, p. 139). On the other hand, this trades off with the political desire for cross-country comparison.

Secondly, as to weighting, there is no generally accepted procedure. On the one hand, experts could be consulted in a rather open discussion process with the risk of rather subjective weightings. On the other hand, econometrically derived weights might be even less acceptable from a policy-making perspective since politically insignificant variables could be assigned high values. Similar to normalization, weighting poses a genuine problem as it ostensibly aims at the comparability of variables even though these are obviously not comparable (Nardo et al., 2005, pp. 44 et sqq.). The SD indices examined in this paper proceed either by transforming variables' values into a new unique scale (e.g. 0 to 1, 1 to 7, 0 to 100 or -2 to +2) by translation and expansion (CDI, HDI, EVI, EPI, WI, ESI) or convert all the variables into another unit by expansion (square meters or monetary values – EF, ISEW, GS, EDP).

Thirdly, as to aggregation, we showed that RNC scaled variables could be meaningfully aggregated by a geometric mean. Unfortunately, all but one index, fail to comply with the scientific aggregation rules elaborated by Ebert and Welsch (2004): Only the LPI, which is based on populations of some species and thus is measured in RNC, uses the appropriate geometric mean for aggregation. While the LPI is therefore formally correct, the index appears nevertheless rather questionable since it presumes substitutability of species.

For most of the considered indices it would be straightforward to aggregate indicators by a geometric mean to assure consistency and meaningfulness. For example, if the variables of the considered SD indices are properly selected (all RNC), they could easily be aggregated by the geometric mean instead of the used arithmetic mean (the latter being inappropriate for ratio-scales) without any subjective normalization.¹²

Choosing variables, normalization methods, and weightings will in general be associated with subjective judgments (if one does not decide a priori that various problem dimensions are incomparable) contrary to aggregation, where the necessary clear-cut methodological requirements by Ebert and Welsch (2004) guarantee consistent and meaningful aggregation functions. The latter are widely neglected in the SD index practice.

5 Conclusions

We have surveyed eleven indices that are used in policy practice to measure national sustainable development. Our main contribution is the critical assessment concerning to what extent the three central steps of indices formation – normalization, weighting, aggregation – satisfy fundamental scientific requirements. We find that normalization and weighting of indicators – which in general are associated with subjective judgments – reveal a high degree of arbitrariness without mentioning or systematically assessing critical assumptions. As to aggregation, there are scientific rules which guarantee consistency and meaningfulness of composite indices. Yet, these rules are often not taken into account. As a consequence, SD indices currently employed in policy practice are doomed to be useless if not misleading with respect to concrete policy advice.

¹² This is not true for the EVI, as this SD index is based on variables measured in RNC and INC (Fahrenheit). These variables could therefore not be meaningful aggregated.

References

- Alberti, M. and J.D. Parker, 1991. Indices of Environmental Quality: the Search for Credible Measures. *Environmental Impact Assessment Review*, 11 (2): 95-101.
- Asheim, G.B., W. Buchholz, and C. Withagen, 2003. The Hartwick Rule: Myths and Facts. *Environmental and Resource Economics*, 25 (2): 129-150.
- Atkinson, G., W.R. Dubourg, K. Hamilton, M. Munasinghe, D.W. Pearce, and C.E.F. Young, 1997. *Measuring Sustainable Development: Macroeconomics and Environment*. Edward Elgar, Cheltenham.
- Booyesen, F., 2002. An Overview and Evaluation of Composite Indices of Development. *Social Indicators Research*, 59: 115-151.
- Cash, D.W., W. Clark, F. Alcock, N.M. Dickson, N. Eckley, and J. Jaeger, forthcoming. *Saliency, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making*. Policy Sciences.
- Chambers, N. and K. Lewis, 2001. *Ecological Footprint Analysis: Towards a Sustainability Indicator for Business*. ACCA Research Paper, 65, Glasgow.
- Cobb, C.W., 1989. The Index for Sustainable Economic Welfare. In: H. Daly and J.B. Cobb (Editors), *For the Common Good – Redirecting the Economy toward Community, the Environment, and a Sustainable Future*. Beacon Press, Boston, pp. 401-457.
- Cobb, C., T. Halstead, and J. Rowe, 1995. *The Genuine Progress Indicator: Summary of Data and Methodology, Redefining Progress*, Washington DC.
- Cobb, C.W. and J.B. Cobb, 1994. *The Green National Product: a Proposed Index of Sustainable Economic Welfare*. University Press of America, Maryland.
- Dalal-Clayton, B. and S. Bass (Editors), 2002. *Sustainable Development Strategies: A Resource Book*. Organization for Economic Co-Operation and Development, Paris.
- Diefenbacher, H., 1995. Der "Index of Sustainable Economic Welfare" – eine Fallstudie für die BRD 1950-1992. *Texte und Materialien der FEST*, B 24, Heidelberg.
- Dunteman, G.H., 1989. *Principal Component Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences, No. 07-069, Newbury Park, CA.
- Ebert, U. and H. Welsch, 2004. Meaningful Environmental Indices: a Social Choice Approach. *Journal of Environmental Economics and Management*, 47: 270-283.
- ESI (Environmental Sustainability Index), 2002. *2002 Environmental Sustainability Index*. Yale Center for Environmental Law & Policy, New Haven.

- ESI (Environmental Sustainability Index), 2001. 2001 Environmental Sustainability Index. Yale Center for Environmental Law & Policy, New Haven.
- Esty, D. C., M.A. Levy, T. Srebotnjak, A. de Sherbinin, Ch. H. Kim, and B. Anderson, 2006. Pilot 2006 Environmental Performance Index. Yale Center for Environmental Law & Policy, New Haven.
- Esty, D.C., M.A. Levy, T. Srebotnjak, and A. de Sherbinin, 2005. 2005 Environmental Sustainability Index: Benchmarking National Environmental Stewardship. Yale Center for Environmental Law & Policy, New Haven.
- Freudenberg, M., 2003. Composite Indicators of Country Performance: A Critical Assessment. STI Working Paper 2003/16, Paris.
- Guenno, G. and S. Tizzi, 1998. The Index of Sustainable Economic Welfare (ISEW) for Italy. *Nota Di Lavoro*, No. 5.98.
- Gassner, J. and M. Narodoslawsky, 2004. Sustainable Economy and their application to Austria. *International Journal of Environment and Sustainable Development*, 3 (2): 120-144.
- Hamilton, C., 1999. The Genuine Progress Indicator Methodological developments and results from Australia. *Ecological Economics*, 30: 13-28.
- Hamilton, K., G. Atkinson, and D.W. Pearce, 1997. Genuine Savings as an Indicator of Sustainability. CSERGE Working Paper GEC97-03, Norwich.
- Hammond, A., A. Adriaanse, E. Rodenburg, D. Bryant, and R. Woodward, 1995. Environmental Indicators: A Systematic Approach to Measuring and Reporting on Environmental Policy Performance in the Context of Sustainable Development. World Resource Institute, Washington DC.
- Hanley, N., 2000. Macroeconomic Measures of 'Sustainability'. *Journal of Economic Surveys*, 14 (1): 1-30.
- Hartwick, J.M., 1977. Intergenerational Equity and the Investing of Rents from Exhaustible Resources. *American Economic Review*, 67 (5): 972-974.
- Hochreiter, H., B. Obermayr, K. Steiner, and E. Stockhammer, 1995. Der Index of Sustainable Economic Welfare (ISEW) – eine empirische Studie zur Wohlstandsentwicklung in Österreich von 1955 bis 1992. Interdisziplinäres Institut für Umwelt und Wirtschaft, Wien.
- Hodge, R.A. and P. Hardi, 1997. The Need for Guidelines: The Rationale Underlying the Bellagio Principles for Assessment. In: P. Hardi and T. Zdan (Editors), *Assessing Sustainable Development: Principles in Practice*. International Institute for Sustainable Development, Winnipeg, pp. 7-20.

- IMSA (Instituut voor Milieu- en Sastemanalyse), 1995: A Pilot ISEW for the Netherlands. IMSA, Amsterdam.
- Jackson, T. and N. Marks, 1994. Measuring Sustainable Economic Welfare – A Pilot Index: 1950 – 1990. Stockholm Environment Institute, Stockholm.
- Jesinghaus, J., 1999. Indicators for Decision Making, European Commission. European Commission, Ispra, http://esl.jrc.it/envind/idm/idm_e_.htm.
- Jespersen, J., 1994. Et Velfærdsindeks for Danmark, 1965 – 1990, Roskilde.
- Kuik, O.J. and A.J. Gilbert, 1999. Indicators of Sustainable Development. In: J.C.J.M. Van den Bergh, (Editor), Handbook of Environmental and Resource Economics, Edward Elgar, Cheltenham, pp. 722-730.
- Moldan, B., T. Hák, J. Kovanda, M. Havránek, and P. Kušková, 2004. Composite indicators of environmental sustainability. Invited paper to ‘Statistics, Knowledge and Policy: OECD World Forum on Key Indicators’, Palermo Italy, 10-13 November 2004.
- Nardo, M., M. Saisana, A. Saltelli, and S. Tarantola, 2005. Tools for Composite Indicators Building. European Commission, Ispra.
- Narodoslawsky, M. and Ch. Krotscheck, 2004. What can we learn from ecological valuation of processes with the sustainable process index (SPI) – the case study of energy production systems. Journal of Cleaner Production, 12: 111-115.
- Nordhaus, W. and J. Tobin, 1971. Is Growth Obsolete?. Cowles Foundation, Yale University, Cowles Discussion Papers: 319, New Haven.
- OECD (Organisation for Economic Cooperation and Development), 1994. Environmental Indicators. OECD, Paris.
- Osberg, L., 1985. The Measurement of Economic Welfare. In: D. Laidler (Editor), Approaches to Economic Well-Being, 26. University of Toronto Press, Toronto, pp. 49-87.
- Ott, W., 1978. Environmental Indices: Theory and Practice. Ann Arbor Science, Ann Arbor.
- Parris, T.M. and R.W. Kates, 2003. Characterizing and Measuring Sustainable Development. Annual Review of Environmental Resources, 28 (13): 1-28.
- Pearce, D.W. and G. Atkinson, 1993. Capital Theory and the Measurement of Sustainable Development: An Indicator of Weak Sustainability. Ecological Economics, 8. 103-108.
- Pezzey, J., 1992. Sustainable Development Concepts – An Economic Analysis. World Bank, Washington DC.
- Prescott-Allen, R., 2001. The Wellbeing of Nations. Island Press, Washington DC.

- Radke, V., 1999. Nachhaltige Entwicklung. Springer, Heidelberg.
- Ramachandran, N., 2000. Monitoring Sustainability: Indices and Techniques of Analysis. Concept Publishing Company, New Delhi.
- Schmidt-Bleek, F., 1994. Wieviel Umwelt braucht der Mensch? MIPS, das Maß für ökologisches Wirtschaften. Birkhäuser Verlag, Berlin.
- SOPAC (South Pacific Applied Geoscience Commission), 2005. Building Resilience in SIDS. The Environmental Vulnerability Index (EVI) 2005. SOPAC Technical Report, Suva, Fiji Islands.
- SRU (Rat von Sachverständigen für Umweltfragen), 1994. Umweltgutachten 1994 – für eine dauerhaft umweltgerechte Entwicklung. Metzler-Poeschel, Stuttgart.
- Stehling, F., 1988. Environmental Quality Indices: Problems, Concepts, Examples. In: W. Eichhorn, (Editor). Measurement in Economics. Physica-Verlag, Heidelberg, pp. 349-369.
- UN (United Nations), European Commission, IMF (International Monetary Fund), OECD (Organisation for Economic Cooperation and Development), and World Bank (Editors), 2003. Handbook of National Accounting – Integrated Environmental and Economic Accounting. Studies in Methods, Series F, No.61, Rev.1. UN, New York.
- UNCHS (United Nations Centre for Human Settlements), 2001. The State of the Worlds Cities 2001. UN, Nairobi.
- UNCED (The United Nations Conference on Environment and Development, Earth Summit), 1992. Rio Declaration on Environment and Development, Rio de Janeiro, 3.–14.6.1992.
- UNDP (United Nations Development Programme), 2005. Human Development Report 2005. Oxford University Press, Oxford.
- UNEP (United Nations Environmental Programme), 2000. Integrated Environmental and Economic Accounting – An Operational Manual. Oxford University Press, Oxford.
- Wackernagel, M. and W. Rees, 1997. Unser ökologischer Fußabdruck, Birkhäuser Verlag, Basel.
- Welsch, H., 2005. Constructing Meaningful Sustainability Indices. In: C. Böhringer and A. Lange (Editors), Applied Research in Environmental Economics. Physica Verlag, Heidelberg, pp. 7-22.
- World Bank, 2005. The Little Green Databook 2005. World Bank Publication, Washington DC.
- WWF (World Wildlife Fund), 2004. Living Planet Report 2004. WWF, Gland.
- WWF (World Wildlife Fund), 1998. Living Planet Report 1998. WWF, Gland.
- Zolotas, X., 1981. Economic Growth and Declining Social Welfare. New York University Press, New York.