British Journal of Arts and Social Sciences
ISSN: 2046-9578, Vol.11 No.II (2013)

http://www.bjournal.co.uk/BJASS.aspx

Estimating the Weights of a Composite Index Using AHP: Case of the Environmental Performance Index

Nick Dedeke
Supply Chain and Management Information Systems Group, Northeastern University, Boston, U.S.A.
E-mail: a.dedeke@neu.edu

Abstract

Composite indices (CI) are multidimensional tools that are used to quantify complex issues in society. In the area of environmental performance, the Environmental Performance Index (EPI) is emerging as a composite construct that many executives and managers consider when assessing policy decisions relating to the environment. The construct is also used for benchmarking the environmental performances of countries. The allocation of weights to the 2010 EPI construct was largely determined by the judgments of few experts. This study uses a more analytical method for determining the relative weights for the construct. This research conducted a survey of EPI policy categories and used Analytic Hierarchy Process to determine weights for the EPI construct. The results show that the weights that were derived using the AHP approach were quite different from those that the EPI developers originally assigned to the categories of the 2010 EPI construct. The derived AHP weights were used to compare the BRICs countries. The study found that the environmental performances of Brazil and Russia were much better than those of India and China. Sensitivity and uncertainty analyses were used to confirm the robustness of the study's conclusions.

Keywords: Analytic Hierarchy Process, Composite indicators, Sensitivity analysis, Environmental performance

1. Introduction

A fundamental challenge in decision-making is the task of combining multidimensional measures into a manageable number of metrics. A composite indicator is a tool that is suited to meet this challenge. These tools aggregate multiple indicators into few factors, which are ultimately aggregated into a single score. Composite indicators are increasingly used by National and International Statistical Offices as well as by other public, private and nonprofit organizations, to convey information about different disciplines, such as the environment, economy, and technological development (Saisana et al., 2005). In literature, several indices have been proposed for defining and measuring the state of the environment. For example, the Index of Sustainable Economic Welfare (Daly and Cobb, 1989), the Living Planet Index (Loh et al., 2005), the Ecological Footprint (Rees, 1992), and the Environmental Performance Index (Esty, 2006) have been proposed in literature. The Environmental Performance Index is a construct that is used to aggregate multiple environmental indicators into a single metric. In mathematical terms, the EPI composite index score is an outcome, which is a function of indicators and weights. The implicit weights generally represent the contribution of each indicator to the outcome variable.

While much progress has been made in regards to the creation of indices, more work still needs to be done to improve the available tools. While the field of environmental management now has a number of indices that are comprehensive, some of these tools still need refinement. Refinements are needed because the development of composite scores are often based on imperfect information. The very process of summarizing complex issues into a single score often means that information would be lost. During the construction stages, experts have to make judgments about several issues, including the inclusion or exclusion of variables, the handling of missing data, the determination of the model, and the weighting of the components of the model (Tarantola et al., 2002). These realities imply that the process of creating composite indices may or may not be grounded in science and methodical rigor. For example, Neumayer (2000) found that some environmental indices were essentially arbitrary in their selection of the component variables, and in the choices of the method of aggregation and construction.

The high number of methodical and theoretical assumptions that underpin the development of a composite index require developers to take steps to limit the potential weaknesses that might compromise the quality of an index. Saltelli (2006) recommends stringent standards of rigor and robustness in order to make composite indicators more acceptable to audiences. This appeal was made because composite indicators have provoked some debates in literature. Some have argued that composite indicators lack of consistency (Pillarisetti and van den Bergh 2010) and that they manifest methodological problems (York 2009; Siche et al. 2008; Niemeijer 2002). Cox et al. (1992) concluded that work needs to be done in the area of determining weighting schemes for indicators, as some of the schemes that were used to determine weights were often too complex or too subjective. Cherchye et al. (2007) echoed the concern that there was a lack of consensus in regards to the way that relative weights were determined for indices. This lack of a standard approach had made the allocated relative weights quite controversial for the audiences of a number of composite indicators. Given that composite indices continue to be widely used for environmental decision making (Niemeijer and de Groot 2008), it is important that these concerns be addressed for each composite indicator.

This paper aims to contribute to literature by assessing an important composite construct, the Environmental Performance Index (EPI). The EPI construct is a widely known construct for ranking countries according to their performance in the area of environmental policy accomplishments. Specifically, this study will investigate if the allocation of weights to the 2010 EPI construct is consistent with the results that one would obtain if a participatory approach and the Analytic Hierarchy Process (AHP) method were used to derive the weights for the composite construct. To our knowledge, this study is the first one to use the AHP approach to evaluate the EPI construct. The research surveyed experts in the field of environmental management to explore how they would prioritize some of the policy goals of the EPI construct. Furthermore, though the developers of the 2010 EPI allocated weights subjectively to the construct, this paper uses AHP, a multi-criteria decision-making approach, to systematically allocate weights to the construct.

The rest of the paper is organized as follows. The next section gives a brief introduction of why the economic development in Brazil, Russia, India and China, also known as the BRICs or Emerging market countries, is related to the state of the Environment. Section 2 introduces the objectives, policy categories and indicators of the EPI construct. The section also introduces the weighting and aggregation schemes that were used for the EPI construct. Section 3 presents the survey that was executed in this study and describes the attributes of the respondents. Section 4 presents the AHP method and shows how the AHP-weights were derived from the survey data. Section 5 explains the impact of AHP-weights on country ranking by calculating the EPI scores for 4 Emerging Countries. Section 6 presents the methods of uncertainty and the sensitivity analyses as well as the results that the study got in applying it to this study. Section 7 summarizes the findings of the study and presents some of its limitations.

1.2 Growth in Emerging Countries and the Environment

The countries of Brazil, Russia, India and China are also known as—the BRICs economies. The increased focus on these countries is driven by certain economic realities. It is estimated that the BRICS would account for over half of the size of the G6 countries by 2050. The shift in Gross Domestic Product relative to the G6 is expected to take place gradually over the coming years. By the year 2025, the annual increase in US dollar spending by the BRICs could be double that of the G6 countries. Consequently, the BRICs countries are expected to be major drivers of global demand in the coming years as the economic growth in western economics slow down due to the impact of greying populations and public sector debt challenges. All these projections imply that the emerging markets would become increasingly attractive for global investors and become transformed into locations where industrial growth and production occurs at a rapid tempo (Wilson and Purushothaman, 2003).

This expected rapid growth has serious implications for the environment. The high growth of industrial activities in BRICs countries would very likelyput them in the position of being able to do more harm to the environment. The rapid economic growth will bring with it rapid growth in emissions. In the case of countries that rely on wood, such as Brazil, deforestation would be a major issue.

China’s rapid growth has made it the second largest consumer of chemicals that deplete the ozone layer of the atmosphere. China has also refused to sign a number of global agreements related to the environment, including the U.N. fisheries agreement, which was intended to help countries to coordinate the management of fish stocks that migrate or are in international waters. All the other BRICs ratified the agreement. All the BRIC countries signed the agreement at the 1992 Convention on Biological Diversity. The signatories agreed to conserve biodiversity, to adopt sustainable use of the components of biodiversity, and to share the benefits arising from the commercial and other utilization of genetic resources in a fair and equitable manner (Cassara and Prager 2005; Roodam, 2007). Given their potential influence on the state of the environment, the author believes that it is central to ensure that the environmental performances of these countries are closely monitored, in particular against each other. Therefore, the environmental performance scores of the BRICs countries would be used as the cases for this study.

2. Environmental Performance Index (EPI)

The Environmental Performance Index (EPI) construct was developed by the Yale Center for Environmental Law and Policy and the Center for International Earth Science Information Network at Columbia University's Earth Institute. They developed the tool in collaboration with environmental agencies and experts, who worked with international agencies.

The EPI construct was designed to measure the accomplishment of environmental policy goals at a national level. The construct enables policy makers to evaluate environmental performance by specifically focusing on the environmental policy areas in which the outcomes are measurable. The EPI incorporates several environmental performance criteria that were derived from policy assessments, such as the Millennium Ecosystem Assessment, Intergovernmental Panel on Climate Change, the Biodiversity Indicator Partnership, and the Global Environmental Outlook (Emerson et al., 2010). The Environmental Performance Index (EPI) construct enhances global environmental policy monitoring by making it possible to benchmark countries based on their performance in accomplishing environmental policy targets. The EPI's focus on measurable outcomes is different from that of indices, such as the Sustainability Index, which attempted to measure a much broader scope of environmental practices and activities (Esty et al., 2002; 2005). Specifically, the EPI focuses on the actual environmental performance accomplishment rather than the plans, activities and programs that are done in pursuit of a broad concept of sustainability. The EPI approach embraces a data- and results-oriented emphases that have been recommended in the environmental management literature (Esty, 2001).

The 2010 EPI construct is the third version of the tool. The prior versions were published in 2006 and 2008. The 2010 version included indicators that were carefully selected from academic literature after consultation with relevant experts (Emerson et al., 2010).

The construct, at its lowest level, consists of 25 environmental performance indicators which are grouped and linked to ten policy categories (Figure 1). The number of indicators that are associated with each policy group varies from one to four per policy category. Specifically, the policy category for the Environmental Burden of Disease has one indicator, while the category of the Air Pollution (effects on ecosystem) has four indicators.

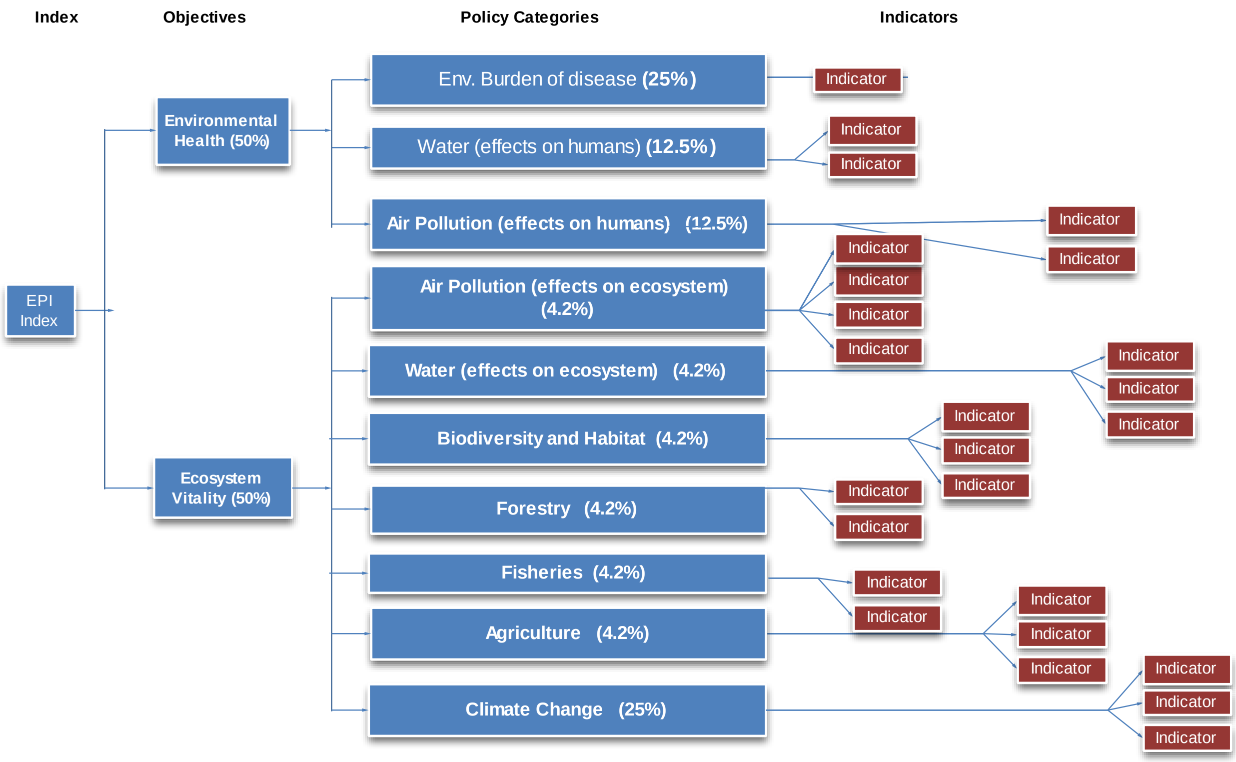

Figure 1: 2010 Environmental Performance Index Construct

The policy categories including, Air Pollution (effects on ecosystem), Water (effects on ecosystem), Biodiversity and Habitat, Agriculture and Climate Change has three indicators each. Each of the remaining policy categories has two indicators. The EPI has two objectives categories, each was weighted 50%. The policy categories are grouped into one of the objectives. For example, Environmental Burden of Disease, Water (effects on humans) and Air Pollution (effects on humans) are grouped into the Environmental Health objective.

The other policy categories, including Air Pollution (effects on ecosystem), Water (effects on ecosystem), Biodiversity and Habitat, Forestry, Fisheries, Agriculture and Climate Change, were grouped under the Ecosystem Vitality objective. Finally, these two objectives were aggregated into the Environmental Policy Index (Figure 1).

2.1 Allocation of Weights to the EPI Construct

The designers of the EPI construct allocated weights to the policy categories based on expert judgments. Specifically, the construct had the following weights allocations: the Environmental Burden of Disease (25%), Air pollution (effects on humans) (12.5%), Water (effects on humans) (12.5%) and Climate Change (25%) (see Table 1). Some categories were assigned higher weights than the other policy categories. The higher weights were primarily assigned based on expert judgment and not on data availability or data quality issues. Within the Environmental Health objective, the Environmental Burden of Disease (EBD) indicator was given a higher weight of 50%, meaning that it contributed 25% to the overall EPI score, because the developers wanted to emphasize its importance as an integrator of a large number of environmental stressors on human health. Similarly, the Climate Change policy category was allocated 50% of the weight (i.e., 25% of the overall EPI), because EPI developers wanted to highlight the global importance of the effects of greenhouse gas emissions across all aspects of ecosystem health and of natural resource management. (Emerson et al., 2010).

The scores for the respective indicators were aggregated into policy categories using weighted linear averaging (this would be explained later in the paper). The scores for the policies were aggregated into the objective categories also by the calculation of weighted averages. The next section will describe the indicators and the scoring method that was used for the EPI.

2.2 Describing EPI's Indicators , and Scoring Scheme

Table 1 presents the 25 indicators that the 2010 EPI construct used to quantify the measurable progress in the respective policy categories. The first column in Table 1 shows the policy categories. The second column presents the indicators that are linked to each of the policy categories. The third column lists those institutional sources from which the data for the respective indicators were gathered. The last column includes the policy targets that were used for each indicator. The target is the level that has to be reached by a country, for each indicator, if the country were to be assigned the maximum proximity-to-target score of 100 points.

Table 1: Environmental Performance Indicators and Data Sources (Emerson et al., 2010)

Policy Categories	Indicators	Data sources	Policy Targets
Environmental burden of disease (25%)	**Environmental burden of disease (25%)**	World Health Organization	10 DALYs (Disability Life Adjusted Years) per 1,000 population
Air pollution (effects on humans) (12.5%)	**Indoor air pollution (6.3%)**	World Development Indicators	0%population using solid fuels
	Outdoor air pollution (Urban Particulates) (6.3%)	World Development Indicators	20 ug/m3 of PM10
Water (effects on humans) (12.5%)	**Access to water (6.3%)**	World Development Indicators	100% population with access
	Access to sanitation (6.3%)	World Development Indicators	100% population with access
Air Pollution (effects on ecosystem) (4.2%)	**Sulfur dioxide emissions per populated land area (2.1%)**	Emissions Database for Global Atmospheric Research (EDGAR) v3.2, United National Framework Convention on Climate Change (UNFCCC), Regional Emissions Inventory in Asia (REAS)	0.01 Gg SO2/sq km
	Nitrogen oxides emissions per populated land area (0.7%)	EDGARv3.2, UNFCCC, REAS	0.01 Gg NOx /sq km
	Non-methane volatile organic compound emissions per populated land area (0.7%)	EDGARv3.2, UNFCCC, REAS	0.01 Gg NMVOC /sq km
	Ecosystem ozone (0.7%)	Model for ozone and Related chemical Tracers (MOZART) II model	0 ppb exceeding above 3000 AOT40. AOT40 is cumulative exceeding above 40 ppb during daylight summer hours

Water (effects on ecosystem) (4.2%)	**Water quality index (2.1%)**	United Nations Environment Program (UNEP) Global Environmental Monitoring System (GEMS)/Water\	Dissolved oxygen: 9.5mg/l (Temp<20ºC), 6mg /l (Temp>=20ºC); pH: 6.5 - 9mg/l; Conductivity: 500µS; Total Nitrogen: 1mg/l; Total phosphorus: 0.05mg/l; Ammonia: 0.05mg/l
	Water stress index (1%)	University of New Hampshire Water Systems Analysis	0% territory under water stress
	Water scarcity index (1%)	Fand and Agriculture Organization (FAO)of the UN	0 fraction of water overuse
Biodiversity & Habitat (4.2%)	**Biome protection (2.1%)**	International Union for Conservation of Nature (IUCN), CIESIN	10% weighted average of biome areas
	Marine protection (1%)	Sea Around Us Project, Fisheries Centre, University of British Columbia	10% of Exclusive Economic Zone (EEZ)
	Critical habitat protection (1%)	Alliance for Zero Extinction, The Nature Conservancy	100% AZE sites protected
Forestry (4.2%)	**Growing stock change (2.1%)**	FAO	Ratio >= 1 n cubic meters / hectare
	Forest cover change (2.1%)	FAO	% no decline
Fisheries (4.2%)	**Marine trophic index (2.1%)**	UBC, Sea Around Us Project	no decline of slope in trend line
	Trawling intensity (2.1%)	UBC, Sea Around Us Project	0% area with combined bottom
Agriculture (4.2%)	**Agricultural water intensity (0.8%)**	FAO	10% water resources
	Agricultural subsidies (1.3%)	Yale Center for Environmental Law & Policy, World Development Report, Organization of Economic Cooperation and Development (OECD)	0 Nominal Rate of Assistance (NRA)
	Pesticide regulation (2.1%)	UNEP-Chemicals	22 points
Climate Change (25%)	**Greenhouse gas emissions per capita (including land use emissions) (12.5%)**	World Resources Institute (WRI), Climate Analysis Indicator Tool (CAIT), Houghton World Development Indicators (WDI) 2009,	2.5 Mt CO2 eq. (Estimated value associated with 50% reduction in global GHG emissions by 2050, against 1990 levels)
	CO2 emissions per electricity generation (6.3%)	International Energy Agency	0 g CO2 per kWh
	Industrial greenhouse gas emissions intensity	WRI-CAIT, WDI, Central Intelligence Agency	36.3 tons of CO2 per $mill (USD, 2005, PPP) of industrial GDP

	(6.3%)		(Estimated value associated with 50% reduction in global GHG emissions by

The policy targets for the EPI were derived from relevant sources, such as global, international, supranational agencies and research centers at universities. A country's performance rating on each indicator was determined based on the degree of the closeness of its indicator value to value of the policy target, which was also called proximity-to-target value. For each country and each indicator, a proximity-to-target value was calculated as a measure of the gap between a country's indicator measurement and the policy target. For each indicator and country, the valid range of proximity-to-target scores was between 0 (worst observed level of performance) and 100 (target level performance).

The EPI indicators included both direct measures as well as proxy measures that offered a rougher gauge of the policy progress. Table 1 reveals that the number of indicators that were aggregated into each policy category was different. Also, the table shows that unequal weights were assigned to the indicators within each policy category. For example, in the policy category of Water (effects on ecosystem) Water Scarcity Index and Water Stress Index indicators were assigned a weight of 1% each, while the indicator for Water Quality Index was assigned 2.1%. This was done because the EPI developers subjectively distributed the weight of each policy category among the respective indicators. Hence, if Water Quality Index (weighted 2.1%), Water Stress Index (weighted 1 %) and Water Scarcity Index (weighted 1%) were given the proximity-to-target scores of 80%, 60% and 80% respectively, the aggregated proximity-to-target score for Water (effects on ecosystem) policy category would be 75% ({80+80+60+80}/4) score. Readers are referred to Emerson et al. (2010) for more details about the data.

3. Description of the Survey Instrument and the Survey Procedure

The research project began with the creation of a survey instrument for soliciting expert opinions about policy categories of the EPI construct. The survey instrument included pairwise comparisons of the policy categories. One of the drawbacks of pair-wise comparisons is that the survey became very long when all the 10 policy categories of the 2010 EPI were included. Therefore, the decision was made to cut down the categories covered to 6, namely the Environmental Burden of Disease, Water (effects on humans), Air pollution (effects on ecosystem), Water (effects on ecosystem), Biodiversity & Habitat, and Climate Change policy categories. After the survey instrument development was completed, it was sent to EPI experts for comments and feedback. The survey was changed and updated based on the feedback from the experts.

Then the research project moved into the phase of creating a participant's list. The names and contact information of experts in the field of environmental management were gathered from research papers, research presentations and from environmental research-oriented web sites. The participant's list included experts who had been involved in the development of the 2010 EPI construct, as well as those who had not been involved in the development of the construct. After this participant's list was completed, those on the list were invited to participate in the survey by email. The survey was sent to the participants as an email attachment. The filled survey forms were returned to the researcher by email. A total of 108 experts were sent emails. Four emails were undeliverable. Each potential respondent was sent two follow-up emails to encourage them to participate. The participants were given up to eight weeks to respond to the survey. 20 usable surveys were received after the time limit for the response expired. This constituted an effective response rate of 19%. The biographical data of the respondents are shown in Table 2.

The survey included both respondents who helped to develop the EPI construct (50%) and those who had not been part of the process (50%). 85% of the respondents had worked in the field of environmental management for over 8 years. The respondents were resident in twelve different countries of the world. Hence, the survey data were gathered from stakeholders, who were active in the field and who had legitimate interest in the value of the EPI construct. Literature recommends the inclusion of such stakeholders when developing composite constructs (Nardo et al., 2005).

Table 2. Biographical Data of the Survey Respondents

Attributes	Proportion
Respondent helped develop the EPI construct	50%
Respondent has used the EPI in decision-making	30%
Respondent has over 8 years' experience in field	85%
Country of Origin of Respondents:	
USA	20%
Canada	25%
Belgium	5%
China	10%
UK	5%
South Africa	5%
Netherlands	5%
Italy	5%
South Korea	5%
Guatemala	5%
Republic of Korea	5%
India	5%

4. Using AHP to Determine Weights

One of the quality attributes of the composite indicator is its degree of coherency. Nardo et al. (2005) suggested the use of various analytical and multivariate statistical tools for building coherent composite indicators. Analytic Hierarchy Process (AHP) is an approach that permits the integration of data, experience, and a scientific approach. In this sense, the use of AHP as the method for deriving weights is better than the arbitrary assignment of weights to a construct (Forman, 1983). The Analytic Hierarchy Process (AHP) is a widely used technique for multi-attribute decision-making. It was developed by Saaty (1980). AHP is a participatory and multi-criteria decision making approach in which the relative importance of a factor or indicator is derived from pairwise comparisons data.

In the survey, individual experts were presented with environmental policy categories and required to perform pairwise comparison of them. First, they responded to the question - Which of the two is the more important? Next, they responded to the question - By how much is the dominant indicator more important than the less important one? The preference ratings were transformed into a semantic scale that had a range from 1 (equality of indicators) to 9 (one indicator is extremely more important than the other).

Let us denote the composite's categories that were rated by experts as $\{I_1, I_2, \dots\dots\dots, I_n\}$, where n is the number of categories compared. Based on the responses of the experts, a preference matrix was derived for each respondent, which had the following format:

$$A = [a_{ij}] = \begin{pmatrix} 1 & a_{12} & a_{13} & a_{1j} & a_{1n} \\ 1/a_{12} & 1 & a_{22} & a_{2j} & a_{2n} \\ 1/a_{13} & 1/a_{22} & 1 & a_{3j} & a_{3n} \\ 1/a_{1j} & 1/a_{2j} & 1/a_{3j} & 1 & a_{4n} \\ 1/a_{1n} & 1/a_{2n} & 1/a_{3n} & 1/a_{4n} & 1 \end{pmatrix} \quad [\ 1\ ]$$

Where a_{ij} is the preference for indicator I_i over I_j when both were compared as a pair of alternatives, for i, j = 1, 2,……n. Notice that if an evaluator's preference of I_i over I_j is a_{ij}, the preference of I_j over I_i is a_{ji} = 1/a_{ij} . Where a_{ij} > 0.

Using this preference matrix information from each expert, the AHP method was used to derive the relative weights of each indicator using the Eigenvector technique. The weight matrix had a form similar to the following:

$$W = [w_i/w_j] = \begin{pmatrix} w_1/w_1 & w_1/w_2 & w_1/w_3 & w_1/w_j & w_1/w_n \\ w_2/w_1 & w_2/w_2 & w_2/w_3 & w_2/w_j & w_2/w_n \\ w_3/w_1 & w_3/w_2 & w_3/w_3 & w_3/w_j & w_3/w_n \\ w_j/w_1 & w_j/w_2 & w_j/w_3 & w_j/w_j & w_j/w_n \\ w_n/w_1 & w_n/w_2 & w_n/w_3 & w_n/w_j & w_n/w_n \end{pmatrix} \quad [\ 2\ ]$$

The calculation of the weights matrix was done using the MATLAB software. The Eigenvectors and Eigenvalues were derived using the following equation:

[W, λ] = eig(A) [3]

The weights matrix was derived as the normalized Principal Eigenvector, also called the Priority Vector.

For a matrix of size N × N, only N–1 comparisons were required to establish weights for N indicators. However, using the AHP method a total of N(N-1)/2 comparisons were done. This redundancy permits an analyst to derive judgment errors by calculating the Consistency Ratio for each respondent. Hence, AHP provides a method for estimating the consistency of each evaluator. The Consistency Ratio (CR) of each respondent was calculated using the following equation:

$$CR = \frac{CI}{RI} \quad [\ 4\ ]$$

Where CI and RI are the Consistency Index and Random Consistency Index respectively. CI is given by the following relationship:

$$CI = \frac{\lambda_{max} - n}{n-1} \quad [\ 5\ ]$$

Where n is the number of comparisons and λ_{max} is the largest Eigenvalue or the Principal Eigenvalue. According to Saaty (1980), a small Consistency Index, of less than 0.1, is considered to be good. However, this is merely a rule-of-thumb. There is no consensus in literature in regards to what is an optimal Consistency Ratio. For more details about the AHP method the reader is referred to Lootsma (1991), Murphy (1993), and Saaty (1980).

5. Results of the AHP Analysis

Figure 2 summarizes the descriptive statistics for the AHP-weights that were calculated for all evaluators. The Figure 2 presents the statistics in graphical format. The histograms show the means while the lines show the minimum and maximum vales.

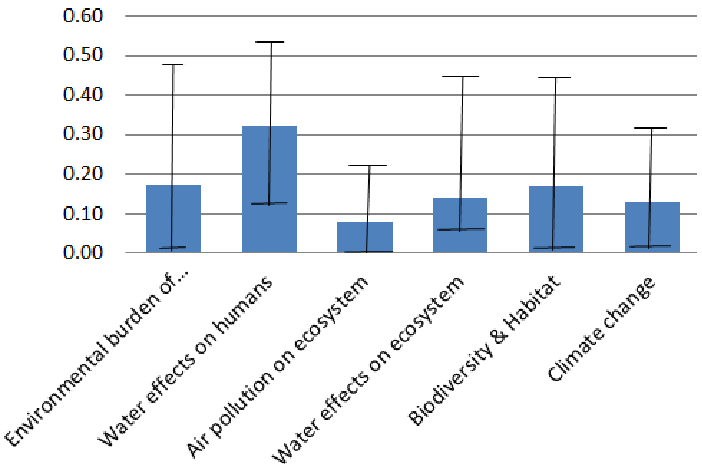

Figure 2. AHP-weights for Policy Categories

Table 3 shows details of the analysis of the AHP-weights by the type of participants. The table includes four main columns, namely results for EPI Collaborators, for Non-Collaborators, for all Respondents and data for the normalized weights of the 2010 EPI categories.

Table 3 has some interesting results. First, we would compare the mean values for all respondents to the normalized original EPI 2010 weights. The mean values of the AHP-weights for all respondents were significantly lower than the original weights for the Environmental Burden of Disease (EBD) and for the Climate Change (CH) policy categories. On the other hand, for all the other categories the mean AHP-weights were up to three-times higher, except the category of Air Pollution (effects on ecosystem), than the original EPI 2010 weights.

Table 3. Comparing AHP-weights to Original Weights of the EPI 2010 Construct

	EPI Collaborators		**EPI Non-Collaborators**		**All Respondents**		**Original Weights of EPI 2010 (Normalized)**
	Mean	**StDev**	**Mean**	**StDev**	**Mean**	**StDev**	
EBD	0.14	0.11	0.20	0.16	0.17	0.14	0.33
WEH	0.33	0.12	0.30	0.12	0.32	0.12	0.167
APE	0.08	0.037	0.08	0.07	0.08	0.05	0.056
WEE	0.19	0.14	0.12	0.13	0.15	0.14	0.056
BH	0.13	0.11	0.17	0.16	0.15	0.14	0.056
CH	0.12	0.09	0.12	0.08	0.12	0.08	0.33

The data also permits a deeper exploration of the weights by type of respondents. Interestingly, both those respondents, who helped develop the EPI 2010 construct, and the Non-collaborators rated Water (effects on humans) (WEH) as the policy category with the highest relative weight. Similarly, both groups assigned similar relative importance to Climate Change and Air Pollution (effects on ecosystem) categories. Only in the areas of Environmental Burden of Disease and Water (effects on the ecosystem) were the differences greater than 6%. The results show that the AHP-weights that were derived from this study were quite different from those allocated to the original EPI construct. It also seems clear that the AHP-weights were mostly similar across the two types of respondents. The Consistency Ratio (CR) ranged from 0.07 to 0.47. The median of the CR series was 0.25 and the Mean of the CR data series was 0.24. These are acceptable values, though they are higher than 0.1 For the CR calculation, the Random Consistency Index (RI) was defined as 1.24 based on the random index study that was conducted by Alonso and Lamata (2006).

5.1 Impact of AHP-Weights on EPI Scores

This study further explored the impact of the derived AHP-weights by calculating partial EPI scores for four Emerging Countries, using the six policy categories that were included in the survey.

To do this, the study downloaded and used the proximity-to-target scores that were used to calculate scores for the BRICs by the 2010 EPI proximity-to-target scores. However, the AHP-weights that were derived in this study were substituted for the original EPI weights. The aggregation of the composite scores was made by using a linear weighted average function of the proximity-to-target scores. In Equation [6], $X_{m,c}$ represents the proximity-to-target scores that were aggregated for each policy category and for each country c. The w_m is the AHP-weight associated with each policy category and linked to the objective category. pEPIx_c is the new composite, score for country c. It is a weighted arithmetic average score given by the following equation:

$$\mathrm{pEPIx}_c = \sum_{j=1}^{j=m} X_{m,c} * w_m \qquad [\ 6\ ]$$

The weighted linear average was used to aggregate the six policy categories into the new partial pEPIx score. Equation [7] was used to calculate the composite score difference for any two countries data set. pEPI_c denotes the pEPI (partial EPI score based on original weights) score for country c, while pEPIx_c denotes the new pEPIx (partial EPI scores using AHP-weights) score for country c. N is the total number of countries.

$$EPIDiff = \sum_{c=1}^{c=N} \left(\mathrm{pEPI}_c - \mathrm{pEPIx}_c\right) \qquad [\ 7\ ]$$

We used the 2010 country proximity-to-target scores data, which were downloaded from the EPI web site, to calculate pEPI (partial EPI based on original weights) and pEPIx (partial EPI scores using AHP-weights) in Table 4.

Table 4. Comparing EPI Scores for Emerging Countries

	EBD	WEH	APE	WEE	BE	CH	pEPI*	pEPIx
BRAZIL	58.5	79.3	39.3	85.6	61.3	46.4	58.60	66.09
RUSSIA	44.2	90.1	54.6	84.5	80.3	45.3	57.10	70.86
INDIA	39.4	50.1	37.1	68.3	38.6	60.2	49.58	48.97
CHINA	62.3	70.0	30.2	66.0	57.2	40.2	54.38	58.70
Original EPI weights (normalized)	0.333	0.167	0.056	0.056	0.056	0.333		
AHP-weights	0.17	0.32	0.08	0.15	0.15	0.12		

*the ranking here was the same as the published rankings for original EPI 2010 which used 10 indicators.

The Table 4 shows that the pEPIx scores, which were based on AHP-Weights, were higher than pEPI scores. This result created the desirable effect that the gaps between old and new EPI scores are now more pronounced than they were before. For example, the EPI difference gap between Brazil and Russia changed from 1.5 to close to 4.77 percentage points. The difference between the scores of China and Russia also increased from 2.72 to 12.16 percentage points. The minimum and maximum EPIDiff was - 13.76 and -0.61 percentage points respectively. Furthermore, Russia index score became the largest of the four, according to the new score. In the next section we will run uncertainty and sensitivity analyses for the BRICs, to investigate the robustness of these conclusions.

6. Uncertainty and Sensitivity Analysis

One could imagine the scenario in which many more persons participated in the survey and hence a wider range of AHP-Weights would be available for the analysis. What would happen to the pEPIx scores of the Emerging Countries in such a case? Uncertainty and sensitivity analyses are useful for exploring such questions. Uncertainty analysis provides investigators with a means for introducing variability into the input variables of a composite indicator. This process transforms the output variable into a non-linear function whose attributes ould be explored. Uncertainty analyses measure the degree to which conclusions are still valid given uncertainties in input variables (Kennedy 2008; Saisana 2008; Saltelli et al. 2008).

There are several methods that could be used to execute the uncertainty analysis. This study adopts an approach that has been used elsewhere in literature (Cherchye et al., 2008; Saisana 2008; Saisana et al., 2007).

The approach assumes that one is given an output model $Y_c = f(X_1, X_2, ...X_j)$ that is a function of several input variables X_j. Both of the Equations [6] and [7], which were presented in Section 5.1, were used for our uncertainty and sensitivity analyses.

Table 5. The List of 30 Uncertain Input Factors for the Analysis

The Sobol (1967) LPτ (quasi random) sampling technique was used to perform uncertainty analyses of the model. In all, 30 uncertain variables were used for the analysis. Variables X_1 to X_{24} (see Table 5) were the input variables that were used to estimate the impact of proximity-to-target inaccuracies on the model. The uncertain variables X_{25} to X_{30} were used to introduce uncertainty into the AHP-weights (see Table 5).

The values that were used for the uncertainty analysis are shown in Table 6. The table shows the mean values and the standard deviations that were used for the uncertain factors. The pEPIx values for each country (from Table 4) were used as the mean values for the proximity-to-target scores in Table 6. To derive the standard deviation values for the uncertainty analysis, the population standard deviations of the proximity-to-target scores for each of the policy categories were calculated, using 2010 proximity-to-target data for 163 countries that were downloaded from the EPI/Yale University web site. These values were divided by a factor of 4. For example, for the Climate Change policy category, the population standard deviation of the 163 original scores was about 16. Hence, a sample standard deviation of 4 percentage points (see Table 6) was used for the uncertainty analysis of the Climate Change category. This meant that the uncertainty analysis investigated what would happen to the outcome function if a fourth of the variance in the input factors was due to measurement errors. A similar approach was used to derive the standard deviations for the AHP-Weights. The standard deviation of the AHP-weights series for the 20 respondents of the survey was calculated and this value was divided by 2. We did not divide this latter value by 4, because the variance values were too small to make an impact on uncertainty analysis.

Table 6: Probability Distribution Function for Input Factors

	Burden of disease (EBD)	Water (effect of humans) (WEH)	Air Pollution effect of Ecosystem (APE)	Water (effects on ecosystem) (WEE)	Biodiversity & Habita (BHH)	Climate Change (CH)
BRAZIL	N [58.49,8]	N [79.33,8]	N [39.3,4]	N [85.63,4]	N [61.31,8]	N [46.44,4]
RUSSIA	N [44.18,8]	N [90.11,8]	N [54.62,4]	N [84.51,4]	N [80.26,8]	N [45.27,4]
INDIA	N [39.35,8]	N [50.11,8]	N [37.07,4]	N [68.34,4]	N [38.64,8]	N [60.24,4]
CHINA	N [62.31,8]	N [70.01,8]	N [30.18,4]	N [65.95,4]	N [57.21,8]	N [40.17,4]
AHP-Weight	N [0.17,0.07]	N [0.32,0.06]	N [0.08,0.03]	N [0.15,0.07]	N [0.15,0.07]	N [0.12,0.04]
AHP2010 Weights	N [0.33,0.07]	N [0.17,0.07]	N [0.056,0.02]	N [0.056,0.02]	N [0.056,0.02]	N [0.33,0.07]

Key: Normal~[Mean, Sample standard deviation]

The uncertainty analysis was run for each of the countries independently, using 12 uncertain variables per run, namely the respective uncertainty inputs for each country's six proximity-to-target scores and the six AHP-weights. The sample size was N = 13,312. Hence, 13, 312 composite scores were calculated per country. The uncertainty analysis was done in the following manner. For each country, a sample of size N was generated. $X_j = [X_{j1,} X_{j2} X_{j3,} .....X_{jk}]$, where j = 1 to N. K represents the number of input factors that were included in the analysis. For each country, the model for pEPIx (see Equation [7]) was estimated at each sample point, yielding not a composite number but a non-linear composite score function. $f(X_j) = f(X_{j1,} X_{j2}, X_{j3,} .....V_{jk})$. In the last step, the values for the 5^{th}, 50^{th} and 95^{th} percentiles of $f(X_j)$ were derived. The distance between the 5^{th} and 95^{th} percentile was used to judge the uncertainty of the model (pEPIx function). The next section discusses the results of the uncertainty analyses.

6. Results of the Uncertainty Analysis

Table 7 shows the results of the uncertainty analysis for the BRICs. It presents the percentiles for $g(Y_{c,\ j})$. Figure 3 shows the probability distribution (pdf) graph for the four countries. In the Figure 3, if the pdf for two countries overlap, it implies that they have similar performance. If the pdfs of two countries do not overlap, it implies that the difference in their pEPIx values is robust and independent of input uncertainties.

Table 7: Uncertainty Analysis Results for the Composite Output Function

NAME	5TH	50TH	95TH
Brazil	49.1	65.2	82.68
Russia	53.2	70.3	88.1
India	35.8	48.23	61.67
China	43.8	58	73.5

Both Table 7 and the pdf in Figure 3 show that Russia and Brazil were the most similar in environmental performance. This is also evident from an evaluation of the distance between the 5^{th} and the 95^{th} percentiles for both countries. This implies that the magnitudes of their composite scores were similar, if one considered the input uncertainties. Hence, the countries should be treated as more or less similar on the pEPIx construct. Figure 3 and Table 7 also show that both Russia and Brazil are clearly better than India.

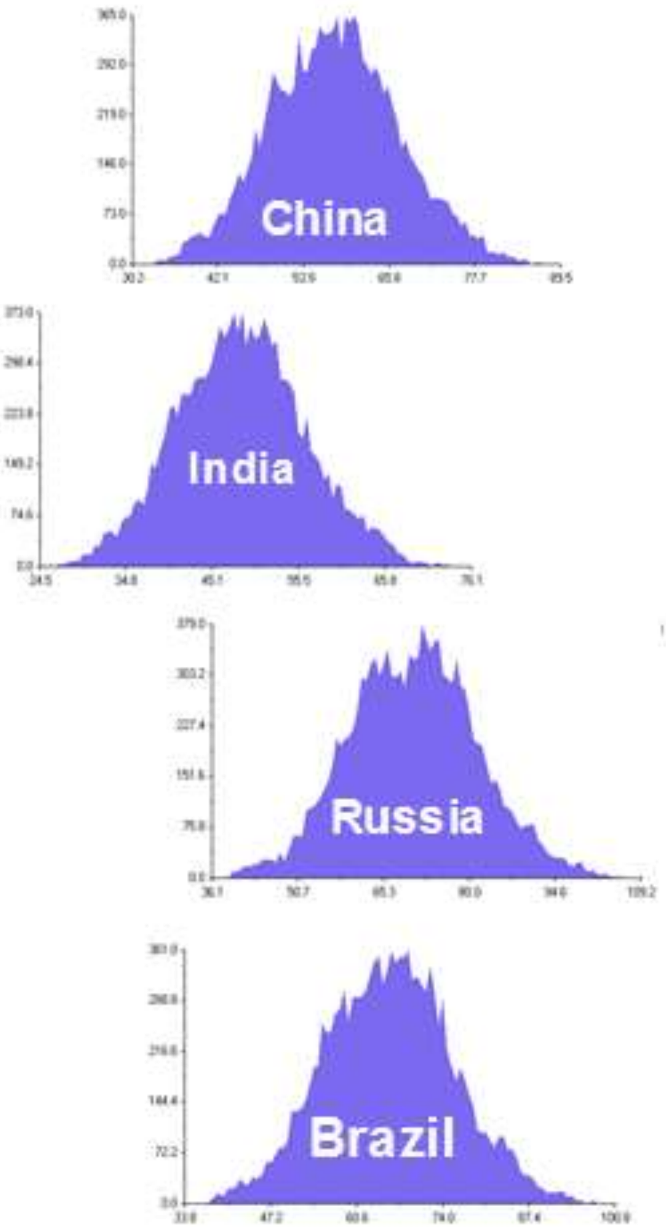

Figure 3. Probability Distribution Function (pdf) for the Uncertainty Analysis.

Similarly, while China has an edge over India, it somewhat trails behind Brazil. Overall, the results of the uncertainty analysis confirms the prior conclusion that pEPIx has higher values and of the ranking of the countries.

Table 8 below shows the uncertainty analyses for the difference between the composite scores, see Equation [7]. The results indicate that up to the 75th percentile, the new partial composite score was higher than the original one, for India, Russia and Brazil. For China, the number of cases in which the new partial composite score was also higher than the score based on the original weights, was up to the 68th percentile.

Table 8: Uncertainty Analysis Results for the Variable EPIDiff: for Difference Original pEPI scores and new pEPIx scores for Russia

NAME	5TH	50TH	75TH	95TH	Mean
Brazil	-29.46	-7.85	0.246	8.35	-7.54
Russia	-36.28	-14.44	-5.7	7.40	-13.7
India	-31.04	-16.13	-10.2	-1.86	-16
China	-24.0	-4.99	2.45	14.0	-4.39

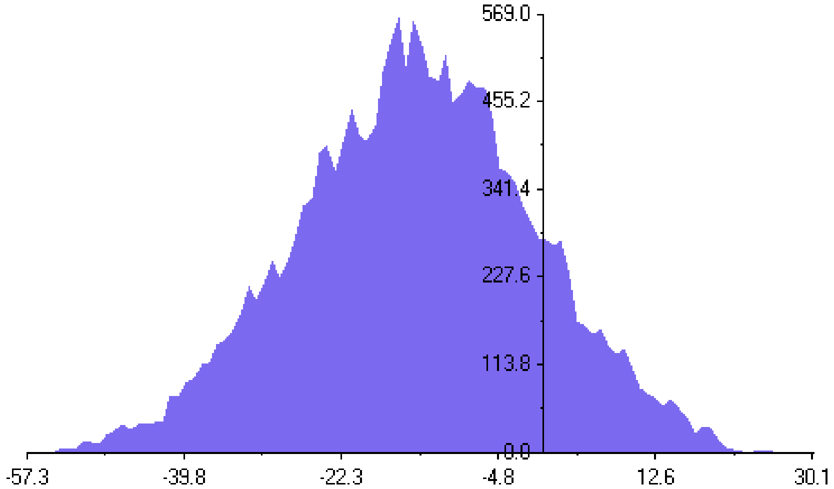

Figure 4. Uncertainty Analysis Graph for the Output Variable EPIDiff: for Difference Original pEPI and pEPIx Scores for Russia

Therefore, in all the cases, the difference in scores was robust to uncertainty in the input variables. Figure 4 shows the graphical plot of the uncertainty analysis for difference in composite scores for the country of Russia. In the next section, the results of the sensitivity analyses would be presented.

6.1 Variance-based Sensitivity Analyses

In the preceding section, we analyzed the impact that uncertain input factors had on the output variable. The purpose of the sensitivity analysis is to apportion the variation in the output composite function to specific input factors (Saltelli et al., 2004; 2008), Saltelli et al. (2008) identified few of the advantages and the disadvantages of using the variance-based approach for sensitivity analysis. Its advantages include the capability to capture a full range of effects of the input factors and its potential to expose the effect of input factor interactions on the output composite function. A main disadvantage is the (time) cost of running the simulations. Given its benefits, this paper used the variance-based method for sensitivity analysis. The variance-based sensitivity approach apportions the variance of the output function to the input factors through the process of decomposition. Variance-based methods enable the estimation of the two sensitivity measures, that is, the first-order sensitivity (S_i) measure and the total order sensitivity index (S_{Ti}). S_i is given by the following equation:

$$S_i = \frac{V_{xi}(E_{x-i}(Y \mid x_i))}{V(Y)} \qquad [8]$$

In the equation above, $E_{x-i}(Y \mid x_i)$ is the conditional expectation and $V_{xi}(E_{x-i}(Y \mid x_i))$ is the variance of the conditional expectation. V(Y) is the unconditional variance, it is also depicted as $V(Y) = V(E(Y \mid X_i)) + E(V(Y \mid X_i))$. S_i measures the fractional contribution that an input variable X_i is responsible for in the variance of Y, without accounting for the interactions among the input factors X_i and other input factors. In other words, the first-order sensitivity identifies the degree of reduction of the variation in Y that would be achieved if X_i were kept fixed. Equation [8] enables one to derive the sensitivity of Y to each input factor, the Equation [9], shows the derivation of the total-order sensitivity:

$$S_{Ti} = \frac{E_{x-i}(V_{xi}(Y \mid x_{-i}))}{V(Y)} \qquad [\,9\,]$$

In many cases, the variance in the output factor is attributable not only to the individual input factors, but also to their interactions to other factors. The derivation of the sensitivities of each of the possible interactions is rather extensive and time consuming. Hence, most studies using sensitivity analysis have the practice of deducing the impact of the interactive effects from the difference between the total-order and the first-order sensitivity values. The total-order sensitivity describes the contribution of each individual factor and of all its possible interactions with other input variables to the output variable. For example, if a model had four input factors, the total-order sensitivity (S_{T1}) for it will capture the individual and all the interactive influences $S_{T1} = S_1 + S_{12} + S_{13} + S_{14} + S_{123} + S_{134} + S_{1234}$.

For a detailed discussion of variance-based methods, the reader is referred to Saltelli et al. (2008). This study used the Sobol (1993) method of sampling and used the same sample that was used for the uncertainty analysis for the sensitivity analysis. The Sobol method and sampling requirement as implemented in SIMLAB software was also used. The sensitivity analyses were run for each country separately.

6.2 Results of the Sensitivity Analyses

Table 8 presents the results of the sensitivity analyses. The table presents the first-order and total-effect sensitivities for our data. In Table 8, one is able to discern four important input factors, if when held fixed, would reduce most of the variance of the output. While there is no universal threshold for determining when an input becomes important, in practice, Saisana at al. (2005) recommends that one consider a factor important if it explains 10% or more of the variance in the output, that is when its first order sensitivity is greater than 0.1 ($S_{\mathbf{k}} > 0.1$). Furthermore, an input factor is considered to be involved in interactions with other inputs, the greater the difference of $S_{\mathbf{Tk}}$ - $S_{\mathbf{k}}$. Based on results given in Table 9, none of the weight factors are involved in significant interactions with other factors. Also, the sensitivity analyses show that the AHP-weights are more influential than the variance in the proximity-to-target scores for the scenario examined.

Specifically, the sensitivity analyses show that four AHP-weights explained a high proportion of the variance in the pEPIx output for the Emerging countries analyzed. Namely, the uncertainties in AHP-weight for Environmental Burden of Disease (EBD), Water (effects on humans) (WEH), Water (effects on ecosystem) (AEE), and Biodiversity and Habitat (BH) policy categories accounted for 88% of the variance in Brazil's pEPIx score, for 74% of variance in India's pEPIx score and for 85% in variance in China's pEPIx score.

Table 9. First-Order and Total-Effect Sensitivity Values for Composite Construct pEPIx

BRAZIL RESULTS	RUSSIA RESULTS	INDIA RESULTS	CHINA RESULTS

	First order Sobol	Total effect Sobol	First order Sobol	Total effect Sobol	First order Sobol	Total effect Sobol	First order Sobol	Total effect Sobol
EBD	0.018	0.019	0.016	0.016	0.03	0.03	0.022	0.023
WEH	0.05	0.06	0.05	0.05	0.098	0.1	0.072	0.072
APE	0	0	0	0	0.001	0.001	0.001	0.001
WEE	0.005	0	0	0.005	0.008	0.009	0.006	0.007
BE	0.01	0.01	0.01	0.013	0.021	0.023	0.01	0.02
CH	0.002	0.002	0.001	0.001	0.003	0.004	0.002	0
wgtEBD	**0.165**	**0.166**	0.08	0.086	**0.13**	**0.131**	**0.23**	**0.23**
WgtWEH	**0.22**	**0.22**	**0.252**	**0.253**	**0.149**	**0.149**	**0.21**	**0.21**
WgtAPE	0.013	0.014	0.023	0.023	0.02	0.02	0.009	0.009
WgtWEE	**0.323**	**0.324**	**0.28**	**0.283**	**0.345**	**0.346**	**0.225**	**0.236**
WgtBE	**0.17**	**0.175**	**0.266**	**0.269**	**0.116**	**0.12**	**0.187**	**0.19**
WgtCH	0.029	0.03	0.025	0.025	0.086	0.087	0.027	0.027

Furthermore, the uncertainties in the AHP-weights for Water (effects on humans) (WEH), Water (effects on ecosystem) (AEE), and Biodiversity and Habitat (BH) policy categories accounted for 80% of variance in Russia's pEPIx score.

The results, in Table 9, provide more insights about the sources of variance in the pEPIx scores across countries. For example, Russia's score was not significantly impacted by changes in EBD's AHP-Weight as much as Brazil's score. This suggests that this country's high score was dependent on three areas whereas Brazil's score was based on good scores in four of the important categories. The results show that the uncertainty in the AHP-Weights accounted for most of the variance in the composite score.

Table 10. First-Order and Total-Effect Sensitivity for Differences Between pEPIx and pEPIx

	BRAZIL RESULTS		**RUSSIA RESULTS**		**INDIA RESULTS**		**CHINA RESULTS**	
	First order Sobol	**Total effect Sobol**	**First order Sobol**	**Total effect Sobol**	**First order Sobol**	**Total effect Sobol**	**First order Sobol**	**otal effect Sobol**
EBDIndia	0.01	0.02	0.01	0.02	0.02	0.03	0.01	0.02
WEHIndia	0.0 1	0.0 2	0.0 1	0. 02	0 .02	0 .03	0 .01	.02
APEIndia	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
WEEIndia	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BEIndia	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01
CHIndia	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.01
WgtEBD	**0.10**	**0.10**	0.05	0.05	0.09	0.09	**0.14**	**0.14**
WgtWEH	**0.14**	**0.14**	**0.16**	**0.16**	**0.11**	**0.11**	**0.13**	**0.13**
WgtAPE	0.01	0.01	0.02	0.02	0.02	0.02	0.01	0.01
WgtWEE	**0.22**	**0.22**	**0.20**	**0.20**	**0.28**	**0.28**	**0.16**	**0.16**
WgtBE	**0.14**	**0.14**	**0.20**	**0.20**	**0.10**	**0.10**	**0.14**	**0.14**
WgtCH	0.02	0.02	0.02	0.02	0.08	0.08	0.02	0.02
WEBD2010	**0.10**	**0.11**	0.05	0.06	0.09	0.10	**0.14**	**0.15**
WWEH2010	**0.19**	**0.19**	**0.22**	**0.23**	**0.15**	**0.16**	**0.18**	**0.18**

WAPE2010	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.00
WWEE2010	0.02	0.02	0.02	0.02	0.03	0.03	0.02	0.02
WBE2010	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.01
WCH2010	0.09	0.09	0.09	0.09	0.03	0.03	0.09	0.09

This suggests that the experts, who develop the EPI composite construct, should consider the process of allocating weights to the EPI construct as a crucial aspect of the development process. The more thorough the process of allocating weights is, the better the resulting composite is likely to be.

Table 10 presents the results of the sensitivity analyses for the difference between the composite scores. It reveals that the AHP-weights for Water (effects on humans) (WgtWEH), Water (effects on ecosystem) (WgtWEE), Biodiversity and Humanity (WgtBE), and the original weights for Water (effects on humans) (WWEH2010) has a significant impact on the variance in EPIDiff variable for all the countries. In addition, the AHP-weight for Environmental Burden of Disease (WgtEBD) and the original weights for Environmental Burden of Disease (WEBD2010) had a significant impact on the EPIDiff variable, difference in original and new composite scores, for Brazil and China respectively. In sum, the results of the sensitivity analyses indicate that the weights associated with EBD and WEH categories accounted for 50% of the variance in EPIDiff for both Brazil and China, while the WEE, BE and WEH weights accounted for over 60% of the variance in EPIDiff for India and Russia. This means that Brazil and China as one pair and India and Russia as a second pair, seem to be achieving environmental progress in the same areas.

7. Conclusion

This paper explored the development of weights for the EPI construct using the AHP approach. The survey included a globally representative sample of experts in the field of environmental management. Six policy categories of the 2010 EPI were included in the preference rating process. The derived AHP-Weights were contrasted with the original weights that were allocated to the original 2010 EPI score. The AHP-Weights that were derived from preference ratings were mostly different from the weights that were assigned to the 2010 EPI construct by its developers. Interestingly, the study found that the mean AHP-weights that were derived from the preference ratings of those experts who helped to develop the EPI construct, was similar to those that were derived from the opinions of experts who were not involved in the development of the EPI construct. This coherence adds credibility to the derived AHP-weights.

This study is the first one, to our knowledge, which has made an effort to derive weights for the EPI construct using the AHP methodology. The study used the weights that were derived using the AHP method (AHP-Weights) to calculate partial EPI scores (pEPIx) for four Emerging countries, namely Brazil, Russia, India, and China. The calculated composite scores for these countries, using the AHP-Weights, were higher than the scores that were derived using the 2010 EPI weights. The pEPIx scores also enabled one to better differentiate between the scores of the countries. The results of sensitivity and uncertainty analysis confirmed these conclusions.

The findings of this study leads to a few proposals in regards to future enhancement of the EPI construct. There seems to be some consensus across respondents in regards to the AHP-weights that are assigned to EPI categories. The 2010 weights were quite different from those weights. The results strongly suggest that the weights that were assigned to the 2010 EPI construct needs to be reconsidered to establish if the level of difference between these subjective weights assignments and those that are derived using a more analytical approach, such as the AHP method, is warranted. Such a review of the 2010 EPI weights allocation by its developers is necessary, given that its credibility is dependent on the degree to which the stakeholders in the community accept its assumptions as being close to theirs. If the significant gaps between opinions of stakeholders and of the designers of a composite construct is not addressed, the stakeholders may start considering the EPI construct to be less applicable to their particular situations.

The discovery of a gap between the weights that were assigned to the categories of the 2010 EPI construct and those that were derived from a field survey, illustrates the gains and improvements that could be achieved through the use of a dual approach for determining weights for a composite construct. Such a dual approach would integrate the use of analytical methods, such as the AHP, with expert judgments, in determining weights. This could be achieved in the following manner. The developers of the EPI construct could use an analytical approach and participatory survey to determine the initial weights for a construct. Thereafter, the developers would review such AHP-Weights, and adjust (increase and decrease) the weight allocations to specific policy categories to send a signal to the environmental community about global/regional policy emphases. It is hoped that such adjustments of weights would be done in a judicious manner, in order not limit possible schisms between perceptions of experts on the field and those of developers of the EPI construct.

Future work in this area could pursue additional questions. Such studies might seek to learn if the potential users of the EPI are influenced by the allocated weights to a construct when they decide to adopt or to ignore the construct. Future work could investigate if new EPI policy categories need to be added to the EPI construct. The study has at least two limitations. First, only six (covering about 75% of the total weight) of the ten policy categories of the EPI were included in the AHP rating process. Hence, the weights that were derived in the study should be seen as one guideline relating primarily to the allocation of the weights to the included policy categories of the EPI construct. Future studies might include more policy categories. Also, the sample variances that were used for sensitivity analyses were not determined for each country separately, but were derived based on the attributes of the population data for each policy category. It is plausible that the variance of a country, due to its particular data management situation, could be higher or lower than that which was used in this study. Finally, the study only included one expert from several countries. Future studies may also try to increase participation in each country. If the number of respondents per country were increased, one would be able to explore if derived AHP-weights vary across experts in different countries.

Acknowledgments:

The author is grateful to the senior editor of the reviewers of this journal for their constructive feedback and suggestions on earlier versions of this paper. The author is grateful to the Emerging Markets Center, Northeastern University, Boston, Massachusetts, for their financial support of this research.

References

Alonso, J., and M. T. Lamata. 2006. Consistency in the analytic hierarchy process: A new approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 14, 445-459.

Cassara, A., and D. Prager. 2005. *An Index of Rich-Country Environmental Performance: 2005 Edition*. Center for Global Development, Washington, DC.

Cherchye L, Lovell, K., Moesen, W., Van Puyenbroeck, T. 2007. One market, one number? A composite indicator assessment of EU internal market dynamics, *European Economic Review,* 51, 749–779.

Cherchye, L., W. Moesen, N. Rogge, T. Van Puyenbroeck, M. Saisana, and A. Saltelli. 2008. Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index, *Journal of the Operational Research* Society, 59, 239-251.

Cox, D., Fitzpatrick, R., Fletcher, A., Gore, S., Spiegelhalter, D. and Jones, D. 1992. Quality-of-life assessment: can we keep it simple? *Journal of the Royal Statist. Society*, 155(3), 353-393.

Daly, H. E., and J. B. Cobb. 1989. For the common good: Redirecting the economy towards community, the environment and a sustainable future. Boston: Beacon Press.

Emerson, J., D. C. Esty, M.A. Levy, C.H. Kim, V. Mara, A. de Sherbinin, and T. Srebotnjak. 2010. *2010 Environmental Performance Index*. New Haven: Yale Center for Environmental Law and Policy.

Esty, D.C. 2001. Toward data-driven environmentalism: The Environmental Sustainability Index. *The Environmental Law Reporter* 31(5):10603–10613.

Esty, D.C. 2002. “Why Measurement Matters.” In Esty D.C. and P. Cornelius. *Environmental Performance Measurement: The Global 2001–2002 Report*. New York: Oxford University Press.

Esty, D.C., M. Levy, T. Srebotnjak and A. de Sherbinin. 2005. *The 2005 Environmental Sustainability Index: Benchmarking National Environmental Stewardship*. New Haven: Yale Center for Environmental Law and Policy.

Esty, D.C., M.A. Levy, T. Srebotnjak, A. de Sherbinin, C.H. Kim and B. Anderson. 2006. *Pilot 2006 Environmental Performance Index*. New Haven: Yale Center for Environmental Law & Policy.

Forman E.H. 1983. *The analytic hierarchy process as a decision support system*, Proceedings of the IEEE Computer society.

Kennedy, P. 2007. *A Guide to Econometrics*, fifth Edition, Blackwell.

Loh, J., R.E. Green, T. Ricketts, J. Lamoreux, M. Jenkins, V. Kapos, and J. Randers. 2005. The Living Planet Index: using species population time series to track trends in biodiversity, *Phil. Trans. R. Soc. B.*, 360, 289–295.

Lootsma, F.A. 1991. Scale Sensitivity and Rank Preservation in a Multiplicative Variant of the Analytic Hierarchy Process, *Proceedings of the 2nd International Symposium on The Analytic Hierarchy Process*, (Pittsburgh, PA, 71-83.

Murphy, C.K. 1993. Limits of the Analytical Hierarchy Process from its consistency index, *European Journal of Operational Research*, 65, 138-139.

Nardo, M., M. Saisana, A. Saltelli and S. Tarantola. 2005. Tools for composite indicators building. EUR 21682 EN. Ispra: *Econometrics and Statistical Support to Antifraud Unit, Institute for the Protection and Security of the Citizen*, Joint Research Centre (JRC).

Neumayer, E. 2000. On the methodology of ISEW, GPI and related measures: Some constructive suggestions and some doubt on the threshold hypothesis. *Ecological Economics*, 34, 347–361.

Niemeijer, D. 2002. Developing indicators for environmental policy: Data-driven and theory-driven approaches examined by example. Environmental Science & Policy, 5(2), 91–103.

Niemeijer, D., & de Groot, R. S. 2008. A conceptual framework for selecting environmental indicator sets. *Ecological Indicators*, 8(1), 14–25.

Pillarisetti, J., & van den Bergh, J. 2010. Sustainable nations: What do aggregate indexes tell us? Environment, *Development and Sustainability*, 12, 49–62.

Rees, W. E. 1992. Ecological footprints and appropriated carrying capacity: what urban economics leaves out, *Environment and Urbanisation*, 4, 121–130.

Roodman, D. 2007. How Do the BRICs Stack Up? Adding Brazil, Russia, India, and China to the Environment Component of the Commitment to Development Index, *Center for Global Development Working Paper No. 127*, Washington, DC.

Saaty T. L. 1980. *The Analytic Hierarchy Process*, New York: McGraw-Hill.

Saisana M., 2008. The 2007 Composite learning index: Robustness issues and critical assessment, Report 23274, European Commission, JRC-IPSC, Italy.

Saisana M., A. Saltelli, and S. Tarantola 2005. Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators, *Journal of the Royal Statistical* Society - A, 168(2), 307–323.

Saltelli A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola 2008. *Global Sensitivity Analysis. The Primer*, John Wiley & Sons, England.

Saltelli, A. 2006. Composite indicators between analysis and advocacy. *Social Indicators Research*, 81, 65-77.

Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto 2004. *Sensitivity analysis in practice - A guide to assessing scientific models*. New York: John Wiley &

Sons. SIMLAB 2007. European Commission Joint Research Center, Ispra. http://sensitivity-analysis.jrc.cec.eu.int/

Siche, J. R., Agostinho, F., Ortega, E., and Romeiro, A. 2008. Sustainability of nations by indices: Comparative study between Environmental Sustainability Index, ecological footprint and the energy performance indices. *Ecological Economics*, 66(4), 628–637.

Sobol' I. M. 1967. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Physics*, 7, 86–112.

Sobol', I. M. 1993. Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling & Computational Experiment*, 1, 407–414.

Tarantola, S., Saisana, M., Saltelli, A., Schmiedel, F. and Leapman, N. 2002. *Statistical techniques and participatory approaches for the composition of the European Internal Market Index 1992-2001*, EUR 20547 EN, European Commission: JRC-Italy.

Wilson, D., and R. Purushothaman. 2003. Dreaming With BRICs: The Path to 2050, *Goldman Sachs, Global Papers 99*,

York, R. 2009. The challenges of measuring environmental sustainability: Comment on "Political and social foundations for environmental sustainability". *Political Research Quarterly*, 62(1), 205–208.