Building an Enlarged Human Development Indicator: Europe and the Southern Mediterranean Basin

by

Ernest Reig

March 2012

Working Papers in Applied Economics

WPAE-2012-03

Departamento de Estructura Económica (Economía Aplicada II), Facultad de Economía, Universitat de València
Avda dels Tarongers, s/n, 46022 – Valencia, Spain
Phone +34 96 382 83 49 / Fax +34 96 382 83 54
http://www.estructuraeconomica.es
e-mail: investigacion.estructura@uv.es

ISSN 2172-3036

Building an Enlarged Human Development Indicator: Europe and the Southern Mediterranean Basin

Ernest Reig (Dept. of Applied Economics II, University of Valencia, Ernest.Reig@uv.es)

Abstract

This paper calculates a human Wellbeing Composite Index (WCI) for 42 countries, belonging to the European Economic Space, North Africa and the Middle East, as an alternative to the shortcomings of other well-known measures of socio-economic development (i.e. Gross Domestic Product per head and Human Development Index, HDI). To attain this goal, different Data Envelopment Analysis (DEA) models are used as an aggregation tool for seven selected socio-economic variables which correspond to the following wellbeing dimensions: income per capita, environmental burden of disease, income inequality, gender gap, education, life expectancy at birth and government effectiveness.

The use of DEA allows avoiding the subjectivity that would be involved in the exogenous determination of weights for the variables included in WCI. The aim is to establish a complete ranking of all countries in the sample, using a three-step process, with the last step consisting in the use of a model that combines DEA and compromise programming, and permits to obtain a set of common weights for all countries in the analysis.

The results highlight the distance that still separates Southern Mediterranean countries from the benchmark levels established by some European countries, and also point to the main weaknesses in individual countries' performance. Nordic countries, plus Switzerland, top the list of best performers, while Mauritania, Libya and Syria appear at the bottom.

Key words: Wellbeing Composite Index (WCI), Human Development, Data Envelopment Analysis (DEA), Compromise Programming, European and Southern Mediterranean Countries

1. Introduction

This paper represents a contribution to a growing literature that aims to overcome the shortcomings of some well-known measures of socio-economic development, like Gross Domestic Product per head or the Human Development Index of the United Nations, by enlarging the number of variables included in the analysis and using novel methodological approaches. We show the potential of a combination of Data Envelopment Analysis models and Compromise Programming to calculate a human Wellbeing Composite Index (WCI) for a group of 42 countries, made up of North Africa and some Middle East countries, all the current members of the European Union (EU), plus Iceland, Norway and Switzerland, and the two countries with the most possibilities of becoming new members of the EU in forthcoming years —Turkey and Croatia. We are able to rank all this countries according to their scoring with WCI, and display their weaknesses concerning specific facets of the index.

The European-Mediterranean partnership was launched in 1995, under the Barcelona Process, and revamped in 2008 with the creation of the Union for the Mediterranean. According to the Joint Declaration of the Paris Summit of the Euro-Mediterranean Heads of States and Government (13 July 2008) this initiative can play an important role in addressing common challenges, including "economic and social development; world food security crisis; degradation of the environment, including climate change and desertification, with the view of promoting sustainable development". In fact the group of European and Southern Mediterranean countries share a geo-political region characterized by strong commercial and migratory flows and increasing political and mutual security ties. The interest of ranking these countries according to a multi-faceted index of human wellbeing lies in the possibilities opened for exploring gaps in a wide range of particular dimensions, and not simply in income per head, and in benchmarking their individual performance against best-practice in the area. The new insights provided by this type of analysis may help to improve the design of social and economic cooperation between both sides of the Mediterranean Basin. The strategic ambition for the Euro-Mediterranean partnership, as described in the aforementioned Joint Declaration, explicitly mention the determination to favor human resource development and employment, including poverty alleviation, and the commitment to strength political participation, human rights and the role of women in society.

Our wellbeing index is based on the idea of human development, but aims to overcome some of the limitations of the current Human Development Index (HDI) of the United Nations, which this institution has calculated on a yearly basis since 1990 (United Nations Development Program [UNDP], 1990). Among the criticisms leveled to HDI (Desai 1991; Sagar and Najam 1998; Alkire 2010) one is that it only addresses three factors —income, education and health— which, while being very important, do not cover all those aspects that contribute to increasing the quality of human life. For this reason, the WCI proposed here also includes other facets related to gender inequality, income distribution, the effects of the environment on human health, and government effectiveness. However, the aggregation of the scores that each country records for each of these partial indicators, in order to obtain a composite index, entails significant methodological problems. This paper provides a solution that consists of calculating the weights of each partial indicator endogenously, making use of a non-parametric approach known as Data Envelopment Analysis (DEA). As a result, it is possible to rank countries according to their WCI scores, without using subjective or arbitrary weighting schemes. Furthermore, using DEA in combination with Compromise Programming allows to overcome some shortcomings of conventional DEA models and allows for a full ranking of all the countries in our sample. After this Introduction, Sections 2 and 3 deal with methodological aspects, Section 4 describes the dataset, Section 5 comments on the findings and Section 6 concludes. An Annex summarizes the models being used for the calculations.

2. A composite indicator for human wellbeing

The HDI is a good example of a composite index that aggregates information referring to three large areas, —income, education and longevity—, and which aims to provide a more complete picture of human wellbeing than one merely based on average levels of income per capita. This paper aims to use this as a starting point to advance in two different directions. Firstly, and more technically, by modifying the way in which the sub-indexes that reflect the various dimensions of wellbeing are aggregated. Secondly, by increasing the number of dimensions the composite index considers in order to integrate aspects that fully affect people's wellbeing, such as income and gender inequality, government effectiveness and the effects of environmental conditions on people's health. Although each of these components is relevant on their own, it is worth trying to construct an index that aggregates the information from all of them and allows comparisons to be made across countries. The final result will be a composite index that offers broader coverage than the HDI and which is constructed using a different methodology.

A review of problems arising in the construction of composite indicators can be found in Nardo *et al.* (2008). One of the most relevant problems, after individual relevant sub-indicators have been selected and normalized, is the selection of the appropriate weights to be applied to these sub-indicators, which are supposed to reflect their relative importance.

Researchers have traditionally opted for one of the following three main approaches to weight sub-indicators. The simplest solution has been to arbitrarily impose specific weights, sometimes granting equal weights to the different components.

A second alternative has been to introduce social preferences regarding the different dimensions of interest. This has frequently led to relying on expert opinions. While there are a variety of ways to summarize experts' judgment, one increasingly popular method is to resort to multi-criteria methods and use the Analytical Hierarchic Process (AHP). The AHP is based on an ordinal pair-wise comparison of attributes (Saaty 1980, 2001), after the problem is structured as a hierarchy, with the global target — human development in this case— at the top and the alternatives being compared (i.e. countries) at the bottom.

Finally, a third stream is represented by those researchers that have preferred to avoid subjectivity in the determination of weights and have proceeded to determine them endogenously. One way of doing so is by performing Principal Components Analysis (PCA) on the original variables or sub-indicators. Essentially, PCA is designed to reduce the number of variables to a smaller number of indicators, called *principal components*, which are linear combinations of the original variables. It is commonly believed that if most of the variance in the data can be attributed to the first few principal components, the original variables can be replaced by those components or indicators with a minimum loss of information.

This paper belongs to the methodological stream that aims to compute a weighting scheme for sub-indicators by using an endogenous approach. But this goal will be achieved in this case by making use of DEA, a widely used technique originally designed to allow non-parametric computation of efficiency indexes (Cooper, Seiford and Tone 2007). The typical DEA problem handles, through mathematical programming, a variety of observed quantities of inputs consumed and outputs produced by a number of Decision Making Units (DMUs) to benchmark the performance of individual decision-making units against frontiers of best practices. These frontiers are built up after an examination of the observed behavior of all the units in the data set. Benchmarking against these best practice frontiers establishes the basis for computation of inefficiency measures for each DMU, also paving the way for ranking the DMUs. The following Section is devoted to expounding a summary of DEA methodology and its usefulness as an aggregation tool.

3. DEA as a tool for the construction of composite indicators

3.1. The basics of DEA

DEA defines the efficiency of DMU_0 —the DMU being assessed each time— by the maximum of a ratio that transforms inputs in outputs:

$$\text{Max } \frac{\sum_{r=1}^{s} u_{r0} y_{r0}}{\sum_{i=1}^{m} v_{i0} x_{io}} \tag{1}$$

Subject to

$$\frac{\sum_{r=1}^{s} u_{r0} y_{rj}}{\sum_{i=1}^{m} v_{i0} x_{ij}} \leq 1 \quad j = 1,\ldots,n$$

and

$$u_{r0} \geq 0 \quad r = 1, \ldots, s$$

$$v_{i0} \geq 0 \quad i = 1, \ldots, m$$

This maximization exercise is performed for each of the decision units in the set and each time the weights *u* and *v* in the ratio, respectively applied to outputs (*y*) and inputs (*x*), are chosen in order to place the assessed DMU in the most favorable light, meaning that they are computed by maximizing the efficiency ratio of the unit considered. This maximization is performed subject to the constraints that the efficiency ratio of all *j* units, computed with those same weights, have an upper bound of 1.

Thus, in DEA the weights allowing the aggregation of inputs and outputs are *specific* to each unit, as they are obtained by solving an optimization problem that changes with the DMU that is under consideration each time.

The DEA approach possesses important advantages as a basis for the measurement of any aggregate of socio-economic variables:

a) First, it can deal with a variety of data dimensions, which is important because of the multi-faceted nature of the composite variable

b) Second, it provides an in-built method of data normalization, as DMUs are ranked between 0 and 1 according to their efficiency

c) And finally, exogenous a priori information is not required to calculate weights, which result from solving individual linear program optimization. The relative weight assigned to each attribute or socio-economic variable is *endogenously* determined in the performance evaluation model.

The ability of DEA-based models to build up composite indicators that reflect a variety of economic, social, and environmental factors has been shown by Callens and Tyteca (1999), Zaim et al. (2001) and Zhou et al. (2007), among others. Many authors have later pursued this line of analysis, with some of them purposefully applying this methodology to the construction of wellbeing and quality of life indicators (Hashimoto and Kodama 1997, Zhu 2001, Murias et al. 2005, Jurado and Mayo 2011, Domínguez-Serrano and Blancas 2011, Gonzalez at al. 2011) and human development composite indicators (Mahlberg and Obersteiner 2001, Despotis 2002, 2005, Zhou et al. 2010).

A basic strength of DEA is that it provides a reasonable method to detect inefficiency: a decisional unit is judged inefficient when it performs worse than its peers, *even* if contemplated under the most favorable idiosyncratic set of weights. This *benefit of the doubt* approach (Cherchye *et al.* 2007, 2008) fits international multi-dimensional comparisons particularly well when the alternative practice of using a fixed and predetermined set of weights may prevent acceptance of the entire exercise by the concerned national political authorities. Some national authorities, or domestic experts, may judge that the chosen (fixed) *subjective* common weights are *unfair*, because they give undue importance to those specific topics (i.e. some socio-economic variables), where the country is underperforming. The advantage of using DEA lies, thus, in the fact that this method allows each country to grant higher weights to its best-performing specific features. Using DEA, policy-makers, or other stakeholders, may never argue that an unfair weighting scheme is being used to assess their country, because the most favorable specific weighting scheme —according to the relative performance of the country in any particular dimension— has been used.

In this paper we are not dealing with *efficiency* as such, in the sense of measuring the technical or economical ability to make the best of available resources by transforming inputs into outputs. Instead, we are concerned with comparing several specific *attributes* —related to human development or broadly-defined people's wellbeing— across different countries. Therefore, the objective function in DEA model 1 must change to recognize the change in focus. Instead of having outputs in the numerator and inputs in the denominator, we have only attributes (*outputs*) in the model and all countries are assumed to have a single virtual input of unit value, which can be interpreted as a sort of global manager or *helmsman*. In the context of production economics, this assumption implies that a single input provides varying amounts of several services (Lovell, Pastor and Turner 1995), or, as in this case, gives rise to different intensities in several features concerning socio-economic development and human wellbeing. Restrictions in the optimization program will change accordingly to accommodate these new characteristics. See the Appendix for a summary description of the specific DEA models being used in this paper.

3.2. Using DEA to rank decision making units

DEA models establish a dichotomous classification between efficient and inefficient units. But conventional DEA models fail to discriminate among the sub-set of inefficient units, or among the efficient ones. Each inefficient unit score is calculated in relation to a different set of efficient units acting as a reference, corresponding to a specific segment of the best practice frontier. It renders meaningless any direct comparison among inefficiency levels (Kao and Hung 2005), unless the units being compared share exactly the same reference set, which in practice is highly improbable. Also, the number of efficient units can be excessive and analysts may wish to establish some ordering among them. A variety of methods are currently available to deal with this problem (Adler, Friedman and Sinuany-Stern 2002). Here we have decided to follow three steps to establish different levels of wellbeing performance:

(i) In the first step we use a DEA model with outputs (attributes) but no inputs (a virtual input equal to unity for all countries) to establish an initial difference between *advanced* (efficient) and *not so advanced* (inefficient) countries, according to the information contained in a set of socio-economic variables (attributes). We use model 3, a *Slacks Based Measure* of efficiency suggested by Tone (2001), (see the Appendix) which represents a step forward from model 1 in order to improve the measurement of efficiency, because merely measuring differences in the potential for radial expansions of outputs is not enough to assess DMU's efficiency scores .

(ii) In the second step, we follow an iterative method to detect consecutive layers of *efficiency* (human development or wellbeing). This method will make it possible to *grade* countries, setting up different groups in descending order according to their socio-economic performance.

(iii) Finally, we combine a DEA model with compromise programming —a Multi-Criteria Decision Method (MCDM) — to build a DEA-MCDM model capable of individually ranking all the countries in the sample (see the description of model 4 in the Appendix).

As a result of Step (i), the database is divided into two sets of efficient and inefficient units (countries). We must remember that *efficient* and *inefficient* are synonymous in this case to *well-performing* and *not so well-performing* countries with regards to a set of socioeconomic variables. Steps (ii) and (iii) are complementary, as both have similar goals, but are carried out using different methods. We now briefly describe these methods in non-mathematical terms (see the Appendix for a more detailed and formal description).

Step (ii) groups countries into efficient categories, according to consecutive *iso-efficiency tiers*. Tier 1 is made up of those countries selected as efficient in the first run of the DEA model. Then, efficient countries are withdrawn from the sample and the DEA model is run a second time. Countries classified as efficient in this second run are taken aside to form Tier 2. The DEA model is again applied to the remaining countries in the sample to determine the components of Tier 3. The process can be repeated *n times* until only efficient countries remain in the sample. These countries make up the last iso-efficiency tier. Efficient countries are always assigned to the current tier and the inefficient ones become the dataset of interest for the next highest tier (Barr, Durchholz and Seiford 2000; Gonçalves *et al.* 2009).

Step (iii) aims to fully rank the countries in the sample. In order to do so, a distinction must be drawn between the idiosyncratic set of weights that allows countries to achieve their highest efficiency scores (the aforementioned *benefit of the doubt* approach) and a new set of common weights that are required to compare the performance of each country against a common benchmark. Efficiency scores obtained under the first set of weights may be called *conventional efficiency scores*, while efficiency scores obtained under common weights are called *global efficiency scores.* In order to compute these global efficiency scores, a linear programming model is used, with an objective function that seeks to minimize the distance between both efficiency measures, conventional and global. In fact, this objective function is a linear combination of two different definitions of this distance. The first term of the objective function, when considered on its own, represents the mean deviation between the *conventional DEA-efficiency scores*, computed with the most self-favorable weights for each DMU, and the global efficiency scores, whereas the second term represents the maximum deviation between the aforementioned efficiency scores. This combination of DEA and compromise programming was pioneered in the construction of human development indicators by Despotis (2002, 2005), and was also suggested for the computation of common weights in a DEA context by Kao and Hung (2005). Compromise programming is grounded in the concept of best-compromise solution in multiobjective programming, as the solution that represents the minimum distance from the ideal solution, which under different distance metrics give rise to a family of distance norms (Cohon, 1978). This family of distance norms was used by Romero (1996) to address the reconciliation between environmental and economic objectives in the management of environmental assets, and by Díaz-Balteiro and Romero (2004) to explore compromises between alternative definitions of sustainability for natural systems. Within a DEA context this general approach has been recently put to use to construct a sustainable energy index and a human development index (Hatefi and Torabi, 2010), to rank farms according to their scoring in a composite sustainability indicator (Reig et al., 2011), and to compute a gender wellbeing composite indicator (Domínguez-Serrano and Blancas 2011).

4. Constructing a Wellbeing Composite Index

4.1. Selected variables

The methodology described in the previous section is now applied to construct a Wellbeing Composite Index (WCI) for the 42 aforementioned countries. The index to be built here must cover the following areas:

a) Average level of income of the population as a determinant of the power to purchase market goods and services
b) Ability to enjoy a healthy life
c) Ability to gain access to basic services, such as education, which increase people's range of opportunities
d) Equality/inequality in income distribution
e) Gender equality/inequality
f) Environmental quality
g) Government effectiveness, in order to reflect the influence of institutional factors allowing for human wellbeing

The seven areas listed above are covered by seven selected variables, which are detailed below.

Gross Domestic Product (GDP) per capita

This variable represents average purchasing power. In this case, this variable is included in the same format as it appears in the construction of the United Nations HDI. Income per capita is adjusted, taking logarithms, because 'achieving a respectable level of human development does not require unlimited income' (UNDP 2009, p. 355). The effect is to soften the differences in income per capita at high income levels.

Life Expectancy at Birth (LEB)

LEB summarizes a complex set of conditions that exert influence on the ability of people to live a healthy life, a component of paramount importance in the definition of human development. We take the life expectancy index from the United Nations Human Development Report.

Education (EDU)

Our education index measures country-relative achievement in both literacy and combined primary, secondary and tertiary gross enrolment and also comes from the United Nations Report on Human Development. The importance of education is apparent in two ways: it is valuable in itself, because knowledge and not only income enters the definition of a good life according to the human development concept, and also due to its instrumental relevance. It is a tool, through human capital formation, to achieve other positive goals linked to economic development (fostering innovation and labor productivity, raising income per capita).

All figures for GDP, LEB and EDU refer to 2007.

Government Effectiveness (GE)

The quality of public institutions is extremely important for economic and social development and for achieving a reasonable level of personal security and protection, and as such is receiving a growing amount of attention in international comparisons of domestic conditions. The World Bank has sponsored a project that measures six different dimensions of governance: voice and accountability, political stability and absence of violence and terrorism, government effectiveness, regulatory quality, rule of law and control of corruption. The indicators are based on perceptions of governance drawn from 35 separate data sources constructed by 33 different organizations from around the world, and currently covers a total of 212 countries and territories (Kauffmann, Kraay and Mastruzzi 2009). Within this framework, governance is broadly defined as the traditions and institutions by which authority in a country is exercised. We have chosen a single indicator of government effectiveness to remain in line with the number of variables selected in the other relevant areas and have decided that *government effectiveness* summarizes some of the most important qualities people expect from Government at all levels. The indicator of government effectiveness captures the perception of the quality of public services (schools, basic health, public transportation system, and tax administration), the quality of the civil service and its independence from political pressures, the quality of policy formulation and implementation and the credibility of the government's commitment to such policies (Kauffman, Kraay and Mastruzzi 2009, pp. 6 and 75).

The indicator is defined as the percentile rank (0-100) corresponding to each country score, for 2008.

Environmental Burden of Disease (EBD)

The EBD is one of the basic pillars of the *Environmental Performance Index*, which measures the effectiveness of national environmental protection in 163 countries and is jointly computed by the Yale Center for Environmental Law and Policy (Yale University) and the Center for International Earth Science Information Network (Columbia University).

Modifiable environmental factors are responsible, approximately, for one quarter of the global disease burden and more than one third of the burden among children. Major disease burdens belonging to this category include diarrhea, lower respiratory infections, malaria and some cardio-vascular diseases and cancers. Environmental agents are poor water, sanitation and hygiene, indoor and urban air pollution, lead exposure and climate change.

The EBD is measured in disability-adjusted life years (DALYs), a 'health gap measure that extends the concept of potential years of life lost (PYLL) due to premature death to include equivalent years of 'healthy' life lost by virtue of being in a state of poor health or disability' (Emerson *et al.* 2010, Annex 1). We use data published for the Environmental Performance Index, which refer to 2004.

Gini coefficient (GINI)

The Gini coefficient is one of the most frequently used measures of distribution inequality, ranging from a value of 0 expressing total equality, to a value of 1 that denotes a situation of maximum inequality. It is frequently multiplied by 100, as in this case, to range from 0 to 100.

Comparing different countries, an increasing value of the Gini coefficient means that income distribution is becoming more and more unequal. As all the indicators in this dataset have been defined using *the more the better* criterion, GINI has been defined as 100 —Gini coefficient (in %), in order to preserve this ranking criterion. It means that according to our GINI variable, distribution becomes more egalitarian the higher the value of GINI.

It is difficult to find easy-to-access and reliable international information concerning income distribution. We have used the most recent (2008) Eurostat data for European countries, with some exceptions: only 2000 data were available for Switzerland, 2006 data for Turkey and 2007 for Croatia. As regards the countries in North Africa and the Middle East, we have resorted to the most recent information from the World Income Inequality Database (V2 0c, May 2008), produced by the World

Institute for Development Economics, United Nations University (UNU-WIDER). Unfortunately no data on income distribution can be found for Libya in any database, and to avoid having to exclude this country to build the composite indicator, we have assumed that Libya has a GINI equal to the arithmetic mean of all the other North African and Middle Eastern Countries in the sample.

Poverty rates could represent an alternative if more data concerning this dimension were made available and a consensus was formed in order to compare countries with very different standards of living. Currently, the two poverty indicators (HPI-2 and HPI-1) coined by the Human Development Report respond to different criteria and it is not possible to measure OECD and developing countries using the same scale

Global Gender Gap (GGG) Index

Gender inequalities of income and different opportunities to live a full and healthy life for men and women should appear prominently in any exercise addressing the measurement of human development. We use the GGG index introduced by the World Economic Forum (Hausmann, Tyson and Zahidi 2009) as a tool for tracking gender-based inequalities in four fundamental categories: economic participation and opportunity, educational attainment, political empowerment and health and survival.

Data are truncated at the equality benchmark, which is considered to be 1, with the exception of the two health variables. This means that the same score is assigned to a country that has reached parity between men and women as to one where women have surpassed men. The type of scale being employed measures how close women are to reaching parity with men in each category, but does not reward or penalize countries for having a gender gap in the other direction (females favored over males).

An average of each category index is used to obtain the overall *Gender Gap Index* score. The final value of GGG is bound between 0 (inequality) and 1 (equality). Data refer to 2009.

Unfortunately the GGG Index does not supply information for Lebanon and Libya. In order to avoid having to do without these countries, for which information is available for our other variables, we have resorted to a rough extrapolation. The World Bank supplies data for female participation rates in the labor market in its World Development Indicators Database and we use this information as a proxy for the social status of women. Taking into account that female participation rates in Lebanon stand between Syria and Egypt, we take the average of the GGG Index values of both countries to compute a GGG

Index score for Lebanon. We proceed in the same way for Libya, whose female participation rate occupies an intermediate position between those of Tunisia and Jordan.

The original data corresponding to these seven variables are displayed in Table 1.

5. Results

Results obtained from calculating the WCI by making use of the methodology described in Section 3 and the data contained in Table 1 are discussed below for each stage of the analysis:

Table 1. Country data for selected variables

	LEB	EDU	GDP	GINI	GGG	EBD	GE
Belgium	0.91	0.97	0.98	72.00	0.72	81.00	88.60
Bulgaria	0.80	0.93	0.79	64.00	0.71	65.50	58.30
Czech Republic	0.86	0.94	0.92	75.00	0.68	76.00	82.50
Denmark	0.89	0.99	0.98	75.00	0.76	81.00	99.50
Germany	0.91	0.95	0.98	70.00	0.74	82.80	93.40
Estonia	0.80	0.96	0.89	69.00	0.71	59.40	84.40
Ireland	0.91	0.99	1.00	70.00	0.76	84.80	91.90
Greece	0.90	0.98	0.94	67.00	0.67	82.80	70.60
Spain	0.93	0.98	0.96	69.00	0.73	84.80	80.10
France	0.93	0.98	0.97	72.00	0.73	82.80	90.00
Italy	0.94	0.97	0.95	69.00	0.68	86.90	66.40
Cyprus	0.91	0.91	0.92	72.00	0.67	86.90	85.30
Latvia	0.79	0.96	0.85	62.00	0.74	59.40	70.10
Lithuania	0.78	0.97	0.86	66.00	0.72	60.30	71.60
Luxembourg	0.91	0.98	1.00	72.00	0.69	82.80	92.40
Hungary	0.81	0.96	0.87	75.00	0.69	66.60	73.00
Malta	0.91	0.89	0.91	73.00	0.66	86.90	85.80
Netherlands	0.91	0.99	0.99	72.00	0.75	86.90	95.70
Austria	0.92	0.96	0.99	74.00	0.70	86.90	93.80
Poland	0.84	0.95	0.85	68.00	0.70	70.30	68.20
Portugal	0.89	0.93	0.91	64.00	0.70	79.20	82.00
Romania	0.79	0.92	0.80	64.00	0.68	64.40	50.20
Slovenia	0.89	0.97	0.93	77.00	0.70	77.50	82.90
Slovakia	0.83	0.93	0.89	76.00	0.68	70.30	77.30
Finland	0.91	0.99	0.98	74.00	0.83	82.80	98.10
Sweden	0.93	0.97	0.99	76.00	0.81	86.90	98.60
United Kingdom	0.91	0.96	0.98	66.00	0.74	81.00	94.30
Croatia	0.85	0.92	0.85	71.00	0.69	74.50	69.70
Turkey	0.78	0.83	0.81	61.00	0.58	65.50	62.60
Iceland	0.95	0.98	0.98	73.00	0.83	91.50	91.50
Norway	0.93	0.99	1.00	75.00	0.82	82.80	97.60
Switzerland	0.95	0.94	1.00	70.00	0.74	89.10	99.10
Mauritania	0.53	0.54	0.49	61.00	0.61	25.70	14.70
Morocco	0.77	0.57	0.62	59.10	0.59	64.40	51.70
Tunisia	0.81	0.77	0.72	59.20	0.62	67.80	65.40
Algeria	0.79	0.75	0.73	64.70	0.61	61.30	37.00
Libya	0.81	0.90	0.83	63.18	0.62	70.30	18.00
Egypt	0.75	0.70	0.66	67.90	0.59	61.30	43.10
Israel	0.93	0.95	0.93	69.50	0.70	91.50	88.20
Jordan	0.79	0.87	0.65	61.20	0.62	70.30	64.00
Lebanon	0.78	0.86	0.77	63.00	0.60	64.40	31.30
Syria	0.82	0.77	0.64	63.00	0.61	73.00	28.00
Sample Average	0.85	0.91	0.87	68.69	0.70	75.01	73.50
Standard Deviation	0.08	0.11	0.12	5.10	0.06	12.46	22.68
Coefficient of Variation	0.09	0.12	0.14	0.07	0.09	0.17	0.31

Source: cited in the text.

Step (i)

Table 2 summarizes the most relevant findings. It can be observed that 9 countries record maximum WCI scores, namely Switzerland, Norway, Iceland, Sweden, Denmark, Finland, Slovenia, Ireland and the Netherlands. This group is mainly made up of Scandinavian countries and a few other highly developed European countries with a high level of income per capita, and a solid Welfare State. Only the presence in this select group of Slovenia, probably the most advanced country to join the EU in the first decade of the 21st century, was in some way unexpected.

Initially, with the DEA approach it is not possible to ascertain differences between countries when all of them record an identical score equal to unity. However, it is useful to employ an indirect discrimination procedure, based on comparing the number of times that each of these countries acts as a benchmark for countries that register scores lower than one. By doing so, we discover that only three countries are part of the best practice frontier for the rest of the countries in the sample, namely Sweden (SE), Norway (NO) and Iceland (IC), and in particular Sweden, which always appears either on its own or in the company of the others as part of the wellbeing reference frontier for the 33 countries whose WCI score is lower than one. As a result, the socioeconomic characteristics of Sweden in relation to the seven variables included in the analysis will have a large influence on the calculation of the wellbeing gap that separates each of the least advantaged countries from their frontier of reference.

The specific situation of each country for which WCI<1 is quite different. Some countries, such as Austria, France, Luxembourg and Germany fall only slightly short of the maximum score, the gap being less than 5%. Most of the remaining European countries display larger gaps, which are wider still in Central and Eastern European countries. Finally, Rumania and Bulgaria, Turkey, and all the countries in North Africa and the Middle East, with the exception of Israel, record the largest gaps between their socioeconomic situation and frontier projection.

Table 2. WCI: scores and ranking

Country	Rank	Score	Reference	3 Main Weaknesses
Switzerland	1	1		
Norway	1	1		
Iceland	1	1		
Sweden	1	1		
Denmark	1	1		
Finland	1	1		
Slovenia	1	1		
Ireland	1	1		
Netherlands	1	1		
Austria	10	0.97	SE. IC. NO	GGG (16). GE (2). LEB (1)
France	11	0.96	SE. IC. NO	GGG (11). GE (7). EBD (5)
Luxembourg	12	0.96	NO	GGG (19). GE (5). GINI (4)
Germany	13	0.95	SE	GGG (9). GINI (8). GE (5)
Israel	14	0.95	IC	GGG (17). GDP (5). GINI (5)
Belgium	15	0.94	SE	GGG (13). GE (11). EBD (7)
United Kingdom	16	0.94	SE	GINI (15). GGG (9). EBD (7)
Spain	17	0.93	SE. NO	GE (23). GGG (10). GINI (10)
Cyprus	18	0.92	SE	GGG (21). GE (15). GDP (7)
Malta	19	0.92	SE	GGG (22). GE (14). EDU (9)
Czech Republic	20	0.90	SE	GGG (19). GE (19). EBDI (14)
Italy	21	0.90	SE. IC	GE (45). GGG (20). GINI (8)
Portugal	22	0.89	SE	GE (20). GINI (18). GGG (16)
Greece	23	0.89	SE. NO	GE (39). GGG (22). GINI (12)
Slovakia	24	0.88	SE	GE (27). EBD (23). GGG (18)
Croatia	25	0.86	SE	GE (41). GGG (17). EBD (16)
Hungary	26	0.86	SE	GE (35). EBD (30). GGG (18)
Estonia	27	0.86	SE	EBD (46). GE (16). LEB (16)
Poland	28	0.85	SE	GE (44). EBD (23). GDP (16)
Lithuania	29	0.83	SE	EBD (44). GE (37). LEB (19)
Latvia	30	0.82	SE	EBD (46). GE (40). GINI (22)
Bulgaria	31	0.79	SE	GE (69). EBD (32). GDP (25)
Turkey	32	0.77	SE	GE (57). GGG (39). EBD (32)
Jordan	33	0.77	SE	GE (54). GDP (51). GGG (31)
Tunisia	34	0.76	SE	GE (50). GDP (36). GGG (30)
Romania	35	0.76	SE	GE (96). EBD (34). GDP (22)
Egypt	36	0.68	SE	GE (128). GDP (48). EBD (41)
Morocco	37	0.67	SE	GE (90). EDU (69). GDP (59)
Algeria	38	0.67	SE	GE (166). EBD (41). GDP (35)
Lebanon	39	0.66	SE	GE (215). GGG (36). EBD (34)
Syria	40	0.62	SE	GE (252). GDP (55). GGG (34)
Libya	41	0.55	SE	GE (447). GGG (31). EBD (23)
Mauritania	42	0.38	SE	GE (570). EBD (238). GDP (99)

Source: Author's calculations.

The information stemming from this first stage of the analysis also includes the identification of the three variables for which each country displays the most vulnerability or relative weakness. In 31 of the 33 countries in which WCI<1, *GE* is one of the three variables to perform the worst. This is not initially surprising if we take into account that Sweden, the most frequent benchmark, records a very high score of 98.6 in this facet of government effectiveness, only surpassed by Denmark and Switzerland, while the average value is only 73.4. The coefficient of variation for this variable, at 0.30, is the highest of all seven, which explains why almost all countries are poorly placed in relative terms.

Table 2 highlights the three variables that performed the worst in relative terms for each country and provides a distance estimate in the form of a percentage (in brackets), which separates the value observed for the variable in question and the value projected onto the frontier. Some interesting patterns emerge:

(i) The largest distances in the case of the countries whose gap in regard to the frontier is smallest, such as Austria, France, Luxembourg, Germany, Israel and Belgium are recorded in gender equality (GGG).

(ii) The United Kingdom is the only country where income distribution inequality is the main obstacle to overcome in order to record a higher score in socioeconomic wellbeing. Other countries that also perform relatively poorly in this sense are Luxembourg, Germany, Lithuania and those in Southern Europe (Portugal, Spain, Italy and Greece).

(iii) Government effectiveness (GE) emerges as the greatest relative weakness of the group of developing countries at the bottom of the table. It is also relatively frequent that these countries display a significant gap in terms of income per capita (GDP) as a sub-indicator of wellbeing. These countries also often record gaps with regard to their best practice frontier in gender inequality (GGG) and environmental conditions that have a negative impact on human health (EBD). Morocco, however, is the exception to this rule, displaying education deficiencies (EDU) instead as a comparatively important factor hampering its ability to attain a higher WCI score.

Step (ii)

As indicated previously, an alternative procedure to determine differences in the socioeconomic performance of countries is to employ repeated iterations, every time selecting the countries with the top wellbeing score. Four successive applications of the same DEA model used in Step (i) have been performed and this has made it possible to determine 5 successive tiers of countries in descending order, according to their WCI:

Tier 1

Comprising the nine countries that appear in Step (i) with a WCI equal to one, namely Switzerland, Norway, Iceland, Sweden, Denmark, Finland, Slovenia, Ireland and the Netherlands. Sweden stands out among these countries, as mentioned previously, due to frequently being a reference for other countries in the sample.

Tier 2

The following countries are included in this group: France, Germany, Austria, Slovakia, Spain, Greece, Hungary, Israel, Italy, Latvia, Luxembourg and the United Kingdom. The country in this group that most frequently appears as a reference for other countries with a WCI<1 is Austria.

Tier 3

The group is made up of Belgium, Cyprus, Lithuania, Malta and the Czech Republic. Belgium is the most frequent reference in this group.

Tier 4

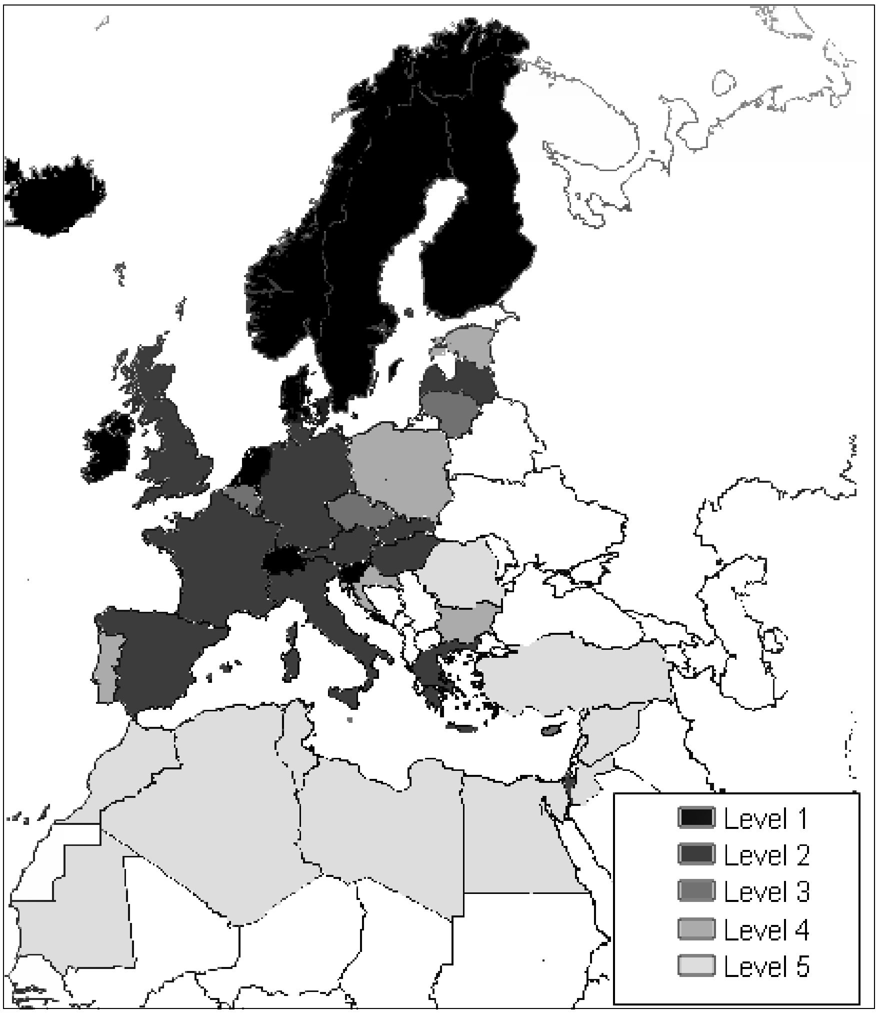
This group includes Bulgaria, Croatia, Estonia, Poland and Portugal, the latter appearing most as a reference for the other countries.

Tier 5

This tier comprises all the countries that recorded a WCI<1 in the previous tier. It is made up of Algeria, Egypt, Jordan, Lebanon, Libya, Mauritania, Morocco, Rumania, Syria, Tunisia and Turkey. As observations now only refer to 11 countries and calculations are made for 7 variables, it is advisable not to proceed with further applications of DEA.

In order to appreciate the different tiers more clearly, they have been represented using different colors in Map 1.

Map 1. WCI levels



Source: Graphing author's calculations.

Step (iii)

The aim of this step is to obtain a complete ranking of the 42 countries analyzed. This is not strictly possible using only a DEA model, and for this reason we have resorted to model 4, which allows obtaining common weights, and which characteristics are described in the Appendix. By applying this model to five basic scenarios (values of the parameter t respectively of 0, 0.25, 0.50, 0.75 and 1) and calculating the mean, we obtain the ranking in Table 3, which provides full discrimination with no ties among the scores recorded by the different countries.

Table 3. WCI: ranking and scores using common weights

	Rank	Scores (with common weights)
Sweden	1	1.00
Denmark	2	0.99
Norway	3	0.98
Switzerland	4	0.98
Finland	5	0.98
Netherlands	6	0.97
Iceland	7	0.97
Austria	8	0.97
Luxembourg	9	0.95
Ireland	10	0.95
Germany	11	0.95
France	12	0.94
United Kingdom	13	0.94
Israel	14	0.94
Belgium	15	0.93
Malta	16	0.92
Cyprus	17	0.92
Slovenia	18	0.91
Spain	19	0.89
Czech Republic	20	0.89
Portugal	21	0.87
Estonia	22	0.86
Slovakia	23	0.86
Greece	24	0.84
Italy	25	0.83
Hungary	26	0.83
Croatia	27	0.82
Poland	28	0.80
Lithuania	29	0.79
Latvia	30	0.77
Jordan	31	0.75
Tunisia	32	0.74
Bulgaria	33	0.73
Turkey	34	0.73
Romania	35	0.69
Morocco	36	0.64
Egypt	37	0.63
Algeria	38	0.60
Lebanon	39	0.59
Syria	40	0.59
Libya	41	0.55
Mauritania	42	0.38

Source: Author's calculations.

It can now be observed that Sweden's place at the top of the Table is unquestionable, followed by four countries with a WCI of more than 0.98, which are the other three Nordic countries (Denmark, Finland and Norway) and Switzerland. The rest of the countries in the EU-15 are ranked from 6th, in the case of the Netherlands, to 25th, in the case of Italy. Among the countries that have joined the EU later, Malta, Cyprus and Slovenia register the highest wellbeing scores and are respectively ranked 16th, 17th and 18th. The lowest score for an EU-27 country was recorded by Rumania, which was ranked 35th. As regards the level of development in the countries in North Africa and the Middle East, the table is led by Israel, ranked 14th, while Mauritania was ranked last at 42nd. Also Tunisia was ranked 32nd, Turkey 34th, Morocco 36th, Egypt 37th and Syria 40th.

These results confirm that the countries on the Southern Mediterranean Basin still have a long way to go to reach the average levels of wellbeing in other nearby countries that belong to the EU or participate in the European Single Market. The gap varies depending on the base indicator being focused on. This means that different development strategies can be adopted to target the specific weaknesses detected in each country and trying to correct them.

6. Conclusions

GDP per capita has been for a long time the main basis for international comparisons of economic development and wellbeing. Nowadays, however, the limitations of this simplistic approach are widely recognized. Such limitations refer to the characteristics of the concept of GDP itself, which does not take into account a large number of non-market goods and services, while including others which purchases do not necessarily raise wellbeing. Also, the linkages between GDP and wellbeing are less direct than what they might initially seem, as this macroeconomic variable omits aspects that affect people's quality of life significantly. Such aspects include people's health, greater or lesser opportunities to take part in social life and political activity, gender inequalities and others.

United Nations has reported the shortfalls of GDP per capita as an indicator of wellbeing and has sponsored the idea of human development, which essentially tends to highlight that the level of income of each person is a means rather than an end to expand their opportunities as a human being. Development is directly linked to the real freedoms that individuals are able to enjoy (Sen 1999).

Although the indicators associated to the HDI Project of the United Nations represent a considerable advance towards a multi-dimensional conception of the development process, they are not exempt of problems that restrict their validity (i.e. they fail to pay enough attention to income distribution inequalities and environmental problems).

This paper provides an alternative by building up a WCI, based on a different methodological approach to that adopted in the construction of the HDI and on the inclusion of a larger number of development facets. Various procedures have been followed in the specialized literature to weight each of the variables or facets of development in order to build composite indicators, and this paper belongs into the methodological stream that aims to compute a weighting scheme for sub-indicators by using an endogenous approach. To attain this goal, Data Envelopment Analysis (DEA) is used as an aggregation tool. One advantage of DEA-computed scores is that they allow each entity or Decision Making Unit (DMU) to display their best side, because the best-performing idiosyncratic weights are being used in the calculation of scores.

In this particular case, solving some mathematical programming models, according to DEA, helps to understand and compare the performing ability of different countries concerning some human wellbeing attributes. 42 countries have been included in the data set, covering current members of the EU, two candidate countries, other European countries that share the Single Market with full members of the EU and countries in North Africa and the Middle East. The aim was to establish a complete ranking of all this countries, which belong to the same geo-political area, using a three-step process as has been described in Section 3. The ranking is based on a composite indicator (WCI) that uses seven variables or 'attributes' corresponding to the following dimensions: income per capita, environmental burden of disease, income inequality, gender gap, education, life expectancy at birth and government effectiveness.

The first stage of the calculation process, Step (i), revealed that the countries in the most favorable situation according to the WCI are Switzerland, Norway, Iceland, Sweden, Denmark, Finland, Slovenia, Ireland and the Netherlands. Basically, these countries are from Northern Europe and are usually among those which enjoy the highest level of wellbeing as a result not only of their high level of income per capita, but to their social policies. In contrast, the countries whose current socioeconomic situation is furthest from their frontiers of efficiency make up a group of 12, including Rumania and Bulgaria, both EU member states, Turkey, which is currently a candidate country, and all the countries in the sample from North Africa and the Middle East, with the exception of Israel. This first stage of the calculation also allowed us to determine which specific aspect recorded the largest gap with respect to the desired situation or, in other words, what each country's main weaknesses were in relative terms. Logically, the profile of each country is different, but as regards some of the Mediterranean Partners of the EU —Morocco, Tunisia, Egypt and Syria— the largest gaps are normally found in income per capita, environmental conditions that affect human health and gender inequality.

The second stage of the procedure, Step (ii), made it possible to visualize the different levels of wellbeing reached by each of the countries in the sample by a process of calculation of the WCI that involves successive iterations. In this case, results have been displayed in a Map, and do not differ from the fundamental aspects discussed above. North Africa and Middle East countries, with the exception of Israel and plus Romania, occupy the less favorable positions.

Finally, Step (iii), we combined the DEA approach with a Multiple Criteria Decision Model (MCDM), a procedure that has allowed us to rank all 42 countries in the sample. Sweden is now seen to unquestionably top the Table, followed by four countries, which are the remaining Scandinavian countries (Denmark, Finland and Norway) plus Switzerland. The lowest score of EU-27 countries is recorded by Rumania, which is ranked 35th. As regards the countries from North Africa and the Middle East, Israel is highest up the table at 14th, while Mauritania is the lowest ranked country in the sample at 42nd. These results coincide quite closely with initial expectations and also with those obtained in the first two stages of the calculation, thereby lending considerable robustness to the approach employed.

In conclusion, this study highlights the usefulness of constructing a multi-dimensional index of wellbeing to cover more aspects than those traditionally considered by the HDI and discusses how to go about doing so. DEA has been used to avoid the frequent subjectivity involved in assigning weightings to each of the dimensions of wellbeing that can be measured, making it possible to fully rank the countries in North Africa, the Middle East and Europe, all of which share the same geopolitical region. The results highlight the distance that still separates both shorelines of the Mediterranean Basin for the variables or dimensions under consideration, which are not limited to income per head, as in many international comparisons. They also allow detecting shortcomings in particular performance areas at individual country level (i.e. governance, gender inequality, etc.). As the Union for the Mediterranean aims to evolve a partnership between European and Northern Mediterranean countries to increase the potential for regional integration and cohesion and promote sustainable development, these findings can be of use for a better design of specific policies in the Mediterranean Region.

References

Adler, N., Friedman, L. & Sinuany-Stern, Z. (2002). Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research,* 140, 49-265.

Alkire, S. (2010). Human Development: Definitions, Critiques and Related Concepts. Human Development Research Paper, No. 2010/01, United Nations Development Program.

Barr, R.S., Durchholz, M.L. & Seiford, L. (2000). Peeling the DEA Onion: Layering and Rank-Ordering DMUs Using Tiered DEA. Technical Report, Southern Methodist University, USA.

Callens, I. & Tyteca, D. (1999). Towards indicators of sustainable development for firms. A productive efficiency perspective. *Ecological Economics*, 28, 41-53.

Cherchye, L., Moesen, W., Rogge, N. & van Puyenbroek, T. (2007). An introduction to 'benefit of the doubt' composite indicators. *Social Indicators Research,* 82, 111-145.

Cherchye, L., Moesen, W. Rogge, N., van Puyenbroeck, T., Saisana, M., Saltelli, A., Liska, R. & Tarantola, S. (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.

Cohon, J.L. (1978) *Multiobjective Programming and Planning.* Academic Press (New York – San Francisco – London).

Cooper, W.W., Seiford, L.M. & Tone, K. (2007). *Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References ad DEA-Solver Software.* 2nd edition, Springer.

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

Despotis, D.K. (2002). Improving the discriminating power of DEA: focus on globally efficient units. *Journal of the Operational Research Society*, 53, 314-323.

Despotis, D.K. (2005). A reassessment of the human development index via data envelopment analysis. *Journal of the Operational Research Society*, 56, 969-980.

Díaz-Balteiro,L. & Romero,C. (2004). In search of a natural systems sustainability index. *Ecological Economics* 49(3), 401-405.

Domínguez-Serrano, M. & Blancas, F.J. (2011). A Gender Wellbeing Composite Indicator: The Best-Worst Global Evaluation Approach. *Social Indicators Research* 102, 477-496.

Emerson, J., Esty, D.C., Srebotnjak, T., Levy, M.A., Mara, V., de Sherbinin, A. & Jaiteh, M. (2010). *Environmental Performance Index 2010.* Yale Center for Environmental Law and Policy, Yale University, and Center for International Earth Science Information-Network, Columbia University, in collaboration with World Economic Forum and Joint Research Center (European Commission).

Gonçalves, E., Correia, J.C., da Silva, G., Angulo, L. & de Carvalho, J.A. (2009). Efficiency and sustainability assessment for a group of farmers in the Brazilian Amazon. *Annals of Operational Research*, 169, 167-181.

González, E., Cárcaba, A., Ventura,J. (2011). The Importance of the Geographic Level of Analysis in the Assessment of the Quality of Life: The Case of Spain. *Social Indicators Research* 102, 209-228.

Hashimoto, A. & Kodama, M. (1997).Has livability of Japan gotten better for 1956-1990? : A DEA approach. *Social Indicators Research* 40, 359-373.

Hatefi, S. M., Torabi, S.A. (2010). A common weight MCDA-DEA approach to construct composite indicators. *Ecological Economics* 70, 114-120.

Hausmann, R., Tyson, L.D. & Zahidi, S. (2009). *The Global Gender Gap Report 2009.* Geneva: World Economic Forum.

Jurado, A. & Perez-Mayo, J. (2011). Construction and Evolution of a Multidimensional Well-Being Index for the Spanish Regions. *Social Indicators Research.* DOI 10.1007/s11205-011-9835-4. Published online: 13 April 2011.

Kao, C. & Hung, H.T. (2005). Data envelopment analysis with common weights: the compromise solution approach. *Journal of the Operational Research Society*, 56, 1196-1203.

Kaufmann, D., Kraay, A. & Mastruzzi, M. (2009). Governance Matters VIII. Aggregate and Individual Governance Indicators 1996-2008. Policy Research Working Paper No. 4978, The World Bank. Development Research Group. Macroeconomics and Growth Team.

Lovell, C.A.K., Pastor, J.T. & Turner, J.A. (1995). Measuring macroeconomic performance in the OECD: A comparison of European and non-European countries. *European Journal of Operational Research,* 87(3), 507-518.

Mahlberg, B. & Obersteiner, M. (2001). Remeasuring the HDI by Data Envelopment Analysis. *Interim Report IR-01-069. International Institute for Applied Systems Analysis* (Laxenburg, Austria).

Murias, P., Martínez, F. & de Miguel, C. (2006). An Economic Wellbeing Index for the Spanish Provinces: A Data Envelopment Analysis Approach. *Social Indicators Research,* 77, 395-417.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A. & Giovannini, E. (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD and Joint Research Center (European Commission).

Reig-Martínez, E., Gómez-Limón, J.A. & Picazo-Tadeo,A.J. (2011). Ranking farms with a composite indicator of sustainability. *Agricultural Economics* 42, 561-575.

Romero, C. (1996). Multicriteria decision analysis and environmental economics: An approximation. *European Journal of Operational Research* 96, 81-89.

Saaty, T.L. (1980). *The Analytic Hierarchy Process. Planning, Priority Setting, Resource Allocation.* McGraw-Hill.

Saaty, T.L. (2001). *Decision Making for Leaders. The Analytic Hierarchy Process for Decisions in a Complex World.* 3rd edition, RWS Publications.

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

Sen, A. (1999). *Development as Freedom.* Oxford University Press.

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130, 498-509.

United Nations Development Program (UNDP, 1990). *Human development report 1990: Concept and Measurement of human development.* New York: Oxford University Press.

United Nations Development Program (UNDP, 2009). *Linking Climate Change Policies to Human Development Analysis and Advocacy. A Guidance Note for Human Development Report Teams.* Washington, DC: United Nations Development Programme, Human Development Report Office.

Zaim, O., Färe, R. & Grosskopf,S. (2001) An economic approach to achievement and improvement indexes. *Social Indicators Research* 56, 91-118.

Zhou, P., Ang, B.W. & Poh, K.L. (2007) A mathematical programming approach to constructing composite indicators. *Ecological Economics* 62, 291-297.

Zhou, P., Ang, B.W. & Zhou, D.Q. (2010) Weighting and Aggregation in Composite Indicator Construction: a Multiplicative Optimization Approach. *Social Indicators Research* 96, 169-181.

Zhu, J. (2001). Multidimensional quality-of-life measure with an application to *Fortune*'s best cities. *Socio-Economic Planning Sciences* 35, 263-284.

Appendix

Step (i) model

For DMU_o, the decisional unit under analysis (or any individual country in this case), the following model can be computed:

$$\begin{aligned} &Max_{\mu_{ro}} \quad h_o = \sum_{r=1}^{R} \mu_{ro} I_{ro} \\ &Subject\ to: \quad \sum_{r=1}^{R} \mu_{ro} I_{rk} \leq 1 \qquad k = 1,\ldots,K \\ &\qquad\qquad\qquad \mu_{ro} \geq 0 \qquad\qquad\quad r = 1,\ldots,R \end{aligned} \tag{2}$$

where I_{rk} stands for the value of indicator r for unit k, and μ_{ro} is the weight attached to indicator r in the assessment of the efficiency of DMU_o. This model represents a linearization of model (1) and also the assumption of a single virtual input for each decisional unit, with value equal to one. Notice also that we use I as a substitute for y as a reminder that we are no longer using outputs in our objective function, but partial indexes corresponding to measurable attributes of the decisional units.

The problem with model (2) is that the measure of efficiency obtained represents the maximum achievable radial expansion of outputs (attributes) and does not include additional expansions that can only be achieved for some outputs (attributes), but not for all of them in the same proportion. For that reason we use another model that summarizes in a single measure the effect of radial expansions and differences among DMUs concerning the existence of slacks for some attributes in the projection onto the frontier (see Cooper, Seiford and Tone [2007] for technical details). To overcome this problem we use the *Slacks Based Measure* (SBM) of efficiency, introduced by Tone (2001), which is computed according to the following model (output-oriented SBM model):

$$\begin{array}{lll}
h_o^* = Min_{\lambda_k, s_r^+} \dfrac{1}{1+\dfrac{1}{R}\sum_{r=1}^{R} s_r^+ \big/ I_{ro}} & & \\
Subject\ to: & x_o \geq \sum_{k=1}^{K} \lambda_k x_k & \\
 & I_{ro} = \sum_{k=1}^{K} \lambda_k I_{rk} - s_r^+ & r = 1, \dots, R \\
 & \lambda_k \geq 0 & k = 1, \dots, K \\
 & s_r^+ \geq 0 & r = 1, \dots, R
\end{array} \qquad (3)$$

where s_r^+ is the slack in the socioeconomic indicator or attribute r, and λ_k measures the intensity with which DMU_k enters in the composition of the efficient reference set to which DMU_o is being compared. Furthermore, the parameter h_o^* is upper bounded to one, with a unity score indicating the best performance.

Step (ii) model

The same model (3) as in Step (i) is used, following an iterative procedure that has been described in the main text.

Step (iii) model

In the first place, the scores obtained according to Model (3) are employed to establish the reference, or *ideal*, idiosyncratic efficiency scores. Then, the following model is employed to obtain common weights and global efficiency scores:

$$
\begin{array}{llll}
Min_{d_k,\mu_r,z}\; t\dfrac{1}{K}\sum_{k=1}^{K} d_k + (1-t)z & & & \\
Subject\ to: & \sum_{r=1}^{R}\mu_r I_{rk} + d_k = h_k^* & k = 1,\dots,K & \\
 & (d_k - z) \le 0 & k = 1,\dots,K & (4) \\
 & d_k \ge 0 & k = 1,\dots,K & \\
 & \mu_r \ge \varepsilon & & \\
 & z \ge 0 & &
\end{array}
$$

ε being a non-Archimedean small number which assures that all attributes I_r are used in the computation of the scores.

The first term of the objective function represents the mean deviation between the DEA-efficiency scores, or *ideal* scores, namely h_k^* and the global efficiency scores for all units, whereas the second term represents, through the non-negative variable *z*, the maximum deviation between the aforementioned efficiency scores (see Despotis [2002, 2005] for technical details). Different sets of common weights μ_r are generated by varying the parameter *t* between 0 and 1, thus granting more or less relative importance to the norms respectively implied by the first and second terms of the objective function. Each value of the parameter *t* may produce a different set of common weights thus generating a different global efficiency pattern. A series of alternative ranks can thus be obtained, according to different *t* values, and all DMUs can then be ranked with regards to their average global efficiency score.