

Non-Compensatory/Non-Linear Composite Indicators for Ranking Countries: A Defensible Setting

Munda, Giuseppe; Nardo, Michela

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:

Munda, G., & Nardo, M. (2009). Non-Compensatory/Non-Linear Composite Indicators for Ranking Countries: A Defensible Setting. *Applied Economics*, 41(12), 1513-1523. <https://doi.org/10.1080/00036840601019364>

Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

gesis
Leibniz-Institut
für Sozialwissenschaften

Terms of use:

This document is made available under the "PEER Licence Agreement". For more Information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der

Leibniz-Gemeinschaft



Non-Compensatory/Non-Linear Composite Indicators for Ranking Countries: A Defensible Setting

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-06-0377.R1
Journal Selection:	Applied Economics
JEL Code:	C43 - Index Numbers and Aggregation < C4 - Econometric and Statistical Methods: Special Topics < C - Mathematical and Quantitative Methods, Q01 - Sustainable Development < Q0 - General < Q - Agricultural and Natural Resource Economics, C82 - Methodology for Collecting and Estimating Macro Data < C8 - Data Collection and Data Estimation Methodology < C - Mathematical and Quantitative Methods
Keywords:	Index Numbers and Aggregation, Multi-Criteria Decision Analysis, , Social Choice, Environmental Sustainability Index



Version June, 2006

Non-Compensatory/Non-Linear Indexes for Ranking Countries: A Defensible Setting

Abstract: Composite indicators (or indexes) are very common in economic and business statistics for benchmarking the mutual and relative progress of countries in a variety of policy domains such as industrial competitiveness, sustainable development, globalisation and innovation. The proliferation of the production of composite indicators by all the major international organizations is a clear symptom of their political importance and operational relevance in policy-making. As a consequence, improvements in the way these indicators are constructed and used seem to be a very important research issue from both the theoretical and operational points of view. This paper starts with an analysis of the axiomatic system underlying the mathematical modelling commonly used to construct composite indicators. Then a different methodological framework, based on non-compensatory/non-linear aggregation rules, is developed. Main features of the proposed approach are: (i) the axiomatic system is made completely explicit, and (ii) the sources of technical uncertainty and imprecise assessment are reduced to the minimum possible degree.

Keywords: *Index Numbers and Aggregation, Multi-Criteria Decision Analysis, Social Choice, Environmental Sustainability Index*

JEL Classification Numbers: *C43, C82, Q01*

1. Introduction

Composite indicators are very common in fields such as economic and business statistics (e.g., the OECD Composite of Leading Indicators) and are used in a variety of policy domains such as industrial competitiveness, sustainable development, quality of life assessment, globalisation, innovation or academic performance (see Cox and others 1992, Cribari-Neto et al 1999, Färe et al. 1994, Griliches 1990, Forni et al. 2001, Huggins 2003, Grupp and Mogege 2004, Lovell et al. 1995, Author 2005, Author et al. 2005, Saisana and Tarantola 2002, and Wilson and Jones 2002, among others). The proliferation of these indicators is a clear symptom of their importance in policy-making, and operational relevance in economic statistics in general (see e.g. Granger, 2001). All the major international organizations such as OECD, the EU, the World Economic Forum or the IMF are producing composite indicators in a wide variety of fields (Author et al., 2005). A general objective of most of these indicators is the ranking of countries and their benchmarking according to some aggregated dimensions (see e.g. Cherchye, 2001, Kleinknecht 2002 and OECD, 2003) (Table 4 in the appendix supplies a sample of various types of composite indicators).

A recent report by OECD clearly states that “... *Composite indicators are valued for their ability to integrate large amounts of information into easily understood formats for a general audience. However, composite indicators can be misleading, particularly when they are used to rank country performance on complex economic phenomena and even more so when country rankings are compared over time. They have many methodological difficulties which must be confronted and can be easily manipulated to produce desired outcomes... The proliferation of composite indicators in various policy domains raises questions regarding their accuracy and reliability. Given the seemingly ad hoc nature of their computation, the sensitivity of the results to different weighting and aggregation techniques, and continuing problems of missing data, composite indicators can result in distorted findings on country performance and incorrect policy prescriptions... Despite their many deficiencies, composite indicators will continue to be developed due to their usefulness as a communication tool and, on occasion, for analytical purposes*” (OECD, 2003, p. 3).

As a consequence, the improvement of the way these indicators are constructed and used seems to be a very important research issue from both theoretical and operational points of view. Our main objective in this article is to contribute to the improvement of the overall quality of sustainability composite indicators (or indexes) by looking at one of their technical weaknesses, that is, the aggregation convention used for their construction. For this aim, we first try to clarify the axiomatic system of a linear aggregation rule. Then we check if this is compatible with the objective of a composite indicator using some results of multi-attribute utility theory and measurement theory literature. Finally, starting from concepts coming from multi-criteria decision theory and social choice, a non-compensatory aggregation rule is proposed and is corroborated with numerical examples.

2. On the Use of Linear Aggregation Rules

Although various functional forms for the underlying aggregation rules of a composite indicator have been developed in the literature (e.g. Diewert, 1976, Journal of Economic and Social Measurement, 2002), in the standard practice, a composite indicator I , can be considered a weighted linear aggregation rule applied to a set of variables (OECD, 2003, p. 5):

$$I = \sum_{i=1}^N w_i x_i \quad , \quad \text{where } x_i \text{ is a scale adjusted variable (e.g. GDP per capita)}$$

normalized between zero and one, and w_i a weight attached to x_i , usually with

$$\sum_{i=1}^N w_i = 1 \quad \text{and} \quad 0 \leq w_i \leq 1, \quad i = 1, 2, \dots, N.$$

The main research question to be answered is the following: under which axiomatic conditions a linear aggregation rule can be used?

A first answer to this question is given by the following theorem (Debreu, 1960; Keeney and Raiffa, 1976; Krantz et al., 1971): *given the variables x_1, x_2, \dots, x_n , an additive aggregation function exists if and only if these variables are mutually preferentially independent*. A subset of indicators Y is *preferentially independent* of $Y^c = Q$ (the complement of Y) only if any conditional preference among elements of Y , holding all elements of Q fixed, remain the same, regardless of the levels at which Q are held. The variables x_1, x_2, \dots, x_n are *mutually*

preferentially independent if every subset Y of these variables is preferentially independent of its complementary set of evaluators.

Preferential independence is a very strong condition from both the epistemological and operational points of view, it implies that the trade-off ratio between two variables $S_{x,y}$ is independent of the values of the $n-2$ other variables, i.e.

$$\frac{\partial S_{x,y}}{\partial q} = 0 \quad \forall x, y \in Y, \forall q \in Q \quad (\text{Ting, 1971}). \quad (1)$$

From an *operational point of view* this means that an additive aggregation function permits the assessment of the marginal contribution of each variable separately (as a consequence of the preferential independence condition). The marginal contribution of each variable can then be added together to yield a total value. If, for example, environmental dimensions are involved, the use of a linear aggregation procedure then implies that among the different aspects of an ecosystem there are not phenomena of synergy or conflict. This appears to be quite an unrealistic assumption (Funtowicz et al., 1990). For example, *"laboratory experiments made clear that the combined impact of the acidifying substances SO_2 , NO_X , NH_3 and O_3 on plant growth is substantially more severe than the (linear) addition of the impacts of each of these substances alone would be."* (Dietz and van der Straaten, 1992).¹ What happens if the linear aggregation is nevertheless done? The resulting indicator will be biased, i.e. it will not entirely and truthfully reflect the information of its components. The dimension and the direction of the error are not easily determined, thus the correction of the composite cannot be properly done. Summarising, we can conclude that the assumption of preference independence is essential for the existence of a linear

¹ From an *epistemological point of view*, preferential independence implies the separability of values. This is quite an important issue in philosophy; for example, in ethics the thesis of the unity of the virtues is defended to different degrees by Plato, Aristotle and Aquinas. *"... it is true that a virtue often cannot be treated apart from the company it keeps. Courage is not an excellence when it appears amid vices-for example as a disposition of the dedicated Nazi.... For example, one's appreciation of the value of freedom, "positive" or "negative", in a particular society, cannot be simply treated as separable from what individuals realise with that freedom. Different political and ethical values when applied in a particular context cannot be applied in isolation from one another"* (O'Neill, 1993, p. 114).

aggregation rule. Unfortunately, it is usually never tested whether preference independence applies to a given composite indicator, although this assumption has very strong consequences which often are not desirable in an index². When preference independence cannot be advocated, non-linear aggregation rules are then needed.

Another issue connected with the use of a linear aggregation rule is the one of weights. The common practice in attaching weights is well synthesised by a recent OECD document: *"Greater weight should be given to components which are considered to be more significant in the context of the particular composite indicator"* (OECD, 2003, p. 10). In the decision theory literature, this concept of weights is usually referred to as *symmetrical importance*, that is *"... if we have two non-equal numbers to construct a vector in R^2 , then it is preferable to place the greatest number in the position corresponding to the most important criterion."* (Podinovskii, 1994, p. 241).

Author (2006) illustrate that the concept of symmetrical importance is incompatible with a linear aggregation rule, given that in a linear aggregation rule, weights can only have the meaning of a trade-off ratio. As a consequence, since trade-offs always depend on the scales of measurement used, and since weights are connected to the values of trade-offs, weights also depend on the scales of measurement. As clearly shown by Anderson and Zalinski (1988), when weights depend on the range of variable scores, such as in the context of a linear aggregation rule, the interpretation of weights as a measurement of the (psychological) concept of importance is completely inappropriate. Summarizing the discussion we can state that the *use of weights in combination with intensity of preference (given that variables are always supposed to be measured on an interval or ratio scale) within a linear aggregation rule originates compensatory aggregation conventions and gives the meaning of trade-offs to the weights*³.

² Often in practice the test of mutual preferential independence is not done because it is considered extremely time consuming. For M variables, there are $M(M-1)/2$ pairs that must be independent, thus the number of conditions to verify gets astronomically large as M gets even modestly large. However some results due to Leontieff (1947) can save much of the potential work to be done (see also Keeney and Raiffa, 1976, p. 112-114).

³ There is unanimous agreement in the literature that the only method where weights are computed as scaling constants and there is no ambiguous interpretation is the so-called trade-off method starting with revealed preferences. No weight importance judgment is required in this method. The trade-off method can be briefly described as follows. Let's consider two countries A and B , differing only for the scores of variables x_k and x_l . The problem is then to adjust the score of say x_k for B , in such a way that A and B become indifferent. Formally, it is:

In standard composite indicators, compensability among the different individual indicators is always assumed; this implies complete substitutability among the various components considered. For example, in a hypothetical sustainability index, economic growth can always substitute any environmental destruction or inside e.g., the environmental dimension, clean air can compensate for a loss of potable water. From a descriptive point of view, such a complete compensability is often not desirable.

Vansnick (1990) showed that the two main approaches in multi-criteria aggregation procedures i.e., the compensatory and non-compensatory ones can be directly derived from the seminal work of Borda and Condorcet. If one wants the weights to be interpreted as “importance coefficients” (or equivalently *symmetrical importance* of variables) non-compensatory aggregation procedures must be used (Bouyssou, 1986; Bouyssou and Vansnick, 1986). From a social choice point of view, these non-compensatory rules are always Condorcet consistent rules (Author, 2006a); their use in the framework of composite indicators, can be corroborated by referring to a clear result of social choice literature⁴. The majority rule is theoretically the most desirable aggregation rule, but practically often produces undesirable intransitivities, thus “*more limited ambitions are compulsory. The next highest ambition for an aggregation algorithm is to be Condorcet*” (Arrow and Raynaud, 1986, p. 77).

Thus we can conclude that the use of non-compensatory aggregation rules to construct composite indicators is **compulsory** for reasons of theoretical consistency when weights with the meaning of importance coefficients are used or when the assumption of preferential independence does not hold. Moreover the

$$I(A) = I(B) \Leftrightarrow I(x_1, \dots, x'_k, \dots, x'_t, \dots, x_n) = I(x_1, \dots, x''_k, \dots, x''_t, \dots, x_n) \Rightarrow \\ \Rightarrow \sum_{\substack{i=1 \\ i \neq k, t}}^N w_i x_i + w_k x'_k + w_t x'_t = \sum_{\substack{i=1 \\ i \neq k, t}}^N w_i x_i + w_k x''_k + w_t x''_t \Rightarrow w_k x'_k + w_t x'_t = w_k x''_k + w_t x''_t$$

This last equation is an equation in the unknown w_k and w_t . To compute the N weights as trade-offs, it is necessary to assess $N-1$ equivalence relations which together with the usual normalization constraint

$w_1 + \dots + w_n = 1$ determine a linear system of N equations in the N unknown weights. Of course if some uncertainty on the variable scores exists, this method cannot be applied. As one can easily understand to assess weights as trade-offs, as it should be always done when using a linear aggregation rule, it is a much harder job than to use weights as importance coefficients.

⁴ Ebert and Welsch (2004) also propose the use social choice to improve the theoretical framework of environmental indexes.

use of Condorcet consistent rules is also **desirable** in general as advised by social choice literature. Finally, one should note that to use a linear aggregation rule, the assumption that the variable scores are measured on an interval or ratio scale of measurement and no uncertainty exists must always apply. Rarely this happens in the practice of composite indicators, where for instance, sometimes quantitative scores are arbitrarily given to variable scores originally measured on an ordinal measurement scale (see e.g. Nicoletti et al., 2000). On the contrary, by using Condorcet aggregation rules no limitation on the measurement scale of the variable scores exists (Author, 2006).

For all these reasons, we think that in some applications the use of non-compensatory Condorcet consistent aggregation rules is desirable. Since this possibility has almost never been explored in the framework of composite indicators the following Section is devoted to this issue.

3. An Axiomatic Approach for the Construction of Non-Compensatory/Non-Linear Condorcet Consistent Composite Indicators

When various variables are used to evaluate two different countries, some of these variables may be in favour of country *a* while other variables may be in favour of country *b*. As a consequence a conflict among the variables exists. How this conflict can be treated at the light of a non-compensatory logic and taking into account the absence of preference independence? This is the classical multi-criteria discrete problem (Author, 1995; Roy, 1996; Vincke, 1992). With this analogy in mind, we present an aggregation convention for (non-linear and non-compensatory) composite indicators able to rank different countries. For sake of clearness, some basic definitions are first given. These definitions are adapted to the context of composite indicators borrowing concepts from multi-criteria decision theory and complex system theory⁵.

3.1 Basic definitions

⁵ Some of these definitions were inspired by discussions with M. Giampietro.

Dimension: is the highest hierarchical level of analysis and indicates the scope of objectives, individual indicators and variables. For example, a sustainability composite indicator can include economic, social and environmental dimensions.

Objective: an objective indicates the direction of change desired. For example, within the economic dimension GDP has to be maximised; within the social dimension social exclusion has to be minimised; within the environmental dimension CO₂ emissions have to be minimised.

Individual indicator: it is the basis for evaluation in relation to a given objective (any objective may imply a number of different individual indicators). It is a *function* that associates each single country with a variable indicating its desirability according to expected consequences related to the same objective. For example, GDP, saving rate and inflation rate inside the objective “growth maximisation”.

Variable: is a constructed measure stemming from a process that represents, at a given point in space and time, a shared perception of a real-world state of affairs consistent with a given individual indicator. To give an example, in comparing two countries, inside the economic dimension, one objective can be “maximisation of economic growth”; the individual indicator might be R&D performance, the indicator score or variable can be “number of patents per million of inhabitants”. Another example: an objective connected with the social dimension can be “maximisation of the residential attractiveness”. A possible individual indicator is then “residential density”. The variable providing the individual indicator score might be the ratio persons per hectare.

A composite indicator or synthetic index is an aggregate of all dimensions, objectives, individual indicators and variables used. This implies that what formally defines a composite indicator is the *set of properties underlying its aggregation convention*. The rest of this section deals with this issue.

3.2 Problem definition

Given a set of individual indicators $G = \{g_m\}$, $m = 1, 2, \dots, M$, and a finite set $A = \{a_n\}$, $n = 1, 2, \dots, N$ of countries, let's assume that the evaluation of each country a_n with respect to an individual indicator g_m (i.e. the indicator score or variable) is based on an *ordinal, interval or ratio* scale of measurement. For simplicity of

exposition, we assume that a higher value of an individual indicator is preferred to a lower one (i.e. the higher, the better), that is:

$$\begin{cases} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) \\ a_j I a_k \Leftrightarrow g_m(a_j) = g_m(a_k) \end{cases} \quad (2)$$

Where, P and I indicate a preference and an indifference relation respectively, both fulfilling the transitive property.

Let's also assume the existence of a set of individual indicator weights $W = \{w_m\}$, $m = 1, 2, \dots, M$, with $\sum_{m=1}^M w_m = 1$, derived as importance coefficients. The mathematical problem to be dealt with is then how to use this available information to rank in a complete pre-order (i.e. without any incomparability relation) all the countries from the best to the worst one. In doing so the following properties are desirable:

1. The sources of uncertainty and imprecise assessment should be reduced as much as possible.
2. The manipulation rules should be the more objective and as simple as possible, that is all *ad hoc* parameters should be avoided.
3. A theoretical guarantee that weights are used with the meaning of "symmetrical importance" must exist. As a consequence, complete compensability should be avoided. This entails that variables have to be used with an ordinal meaning. This is not a problem since no loss of information is implied (Arrow and Raynaud, 1986). Moreover, given that often the measurement of variables is imprecise (see OECD, 2003, p.7), it seems even desirable to use indicator scores with an ordinal meaning.
4. Desirable ranking procedures using ordinal information are always of a Condorcet type (Arrow and Raynaud, 1986, Moulin, 1988). A problem inherent to this family of algorithm is the presence of cycles, i.e. cases where aPb , bPc and cPa . This problem has been widely studied among others by Fishburn, 1973; Fishburn et al., 1979; Kemeny, 1959; Moulin, 1985; Truchon, 1995; Young and Levenglick, 1978, Vidu, 2002; Weber, 2002. The probability $\pi(N, M)$ of obtain a cycle with N countries and M

individual indicators increases with N as well as with the number of indicators. With many countries and individual indicators, cycles occur with an extremely high frequency. Therefore, the ranking procedure used has to deal with the cycle issue properly.

5. Arrow's impossibility theorem (Arrow, 1963) clearly shows that no perfect aggregation convention can exist. Then, it is essential to check not only which properties are respected by a given ranking procedure, but also if any essential property for the problem tackled is lost.

3.3 The proposed composite indicator

The mathematical aggregation convention we are proposing can be divided into two main steps:

1. Pair-wise comparison of countries according to the whole set of individual indicators used.
2. Ranking of countries in a complete pre-order.

For carrying out the pair-wise comparison of countries the following axiomatic system is needed (adapted from Arrow and Raynaud, 1986, p. 81-82).

Axiom 1: Diversity. Each individual indicator is a total order on the finite set A of countries to be ranked, and there is no restriction on the individual indicators; they can be any total order on A .

Axiom 2: Symmetry. Since individual indicators have incommensurable scales, the only preference information they provide is the ordinal pair-wise preferences they contain⁶.

Axiom 3: Positive Responsiveness. The degree of preference between two countries a and b is a strictly increasing function of the number and weights of individual indicators that rank a before b ⁷.

⁶ In our case, this axiom is needed since the intensity of preference of individual indicators is not considered an useful preference information given that compensability has to be avoided and weights have to be symmetrical importance coefficients. Moreover, thanks to this axiom, a normalisation step is not needed. This causes a further reduction of the sources of uncertainty and imprecise assessment.

⁷ In social choice terms then the *anonymity* property (i.e. equal treatment of all individual indicators) is broken. Indeed, given that full decisiveness yields to dictatorship, Arrow's impossibility theorem forces us to make a trade-off between *decisiveness* (an alternative has to be chosen or a ranking has to be made) and anonymity. In our case the loss of anonymity in favour of decisiveness is even a positive property. In general, it is essential that no individual indicator

Thanks to these three axioms a $N \times N$ matrix, E , called *outranking matrix* (Arrow and Raynaud, 1986, Roy, 1996) can be built. Any generic element of E : e_{jk} , $j \neq k$ is the result of the pair-wise comparison, according to all the M individual indicators, between countries j and k . Such a global pair-wise comparison is obtained by means of equation (2).

$$e_{jk} = \sum_{m=1}^M \left(w_m(P_{jk}) + \frac{1}{2} w_m(I_{jk}) \right) \quad (3)$$

where $w_m(P_{jk})$ and $w_m(I_{jk})$ are the weights of individual indicators presenting a preference and an indifference relation respectively. It clearly holds

$$e_{jk} + e_{kj} = 1. \quad (4)$$

Property (4), although obvious, is very important since it allows us to consider the outranking matrix E as a *voting matrix* i.e., a matrix where instead of using individual indicators, alternatives are compared by means of voters' preferences (with the principle one agent one vote). This analogy between a multi-criterion problem and a social choice one, as noted by Arrow and Raynaud (1986), is very useful for tackling the step of ranking the N countries in a consistent axiomatic framework.

The issue is now to exploit the information contained in the outranking matrix in order to rank all countries in a complete pre-order. A problem connected with the use of Condorcet consistent rules is that of cycles. A cycle breaking rule normally needs some arbitrary choices such as to delete the cycle with the lowest support. Now the question is: Is it possible to tackle the cycle issue in a more general way?

Condorcet himself was aware of the problem of cycles in his approach; he built examples to explain it and he was even close to find a consistent rule able to rank any number of alternatives when cycles are present. However, attempts to fully understand this part of Condorcet's voting theory have arrived at

weight is more than 50% of the total weight; otherwise the aggregation procedure would become lexicographic in nature, and the indicator would become a dictator in Arrow's term.

conclusions like “... *the general rules for the case of any number of candidates as given by Condorcet are stated so briefly as to be hardly intelligible ... and as no examples are given it is quite hopeless to find out what Condorcet meant*” (E.J. Nanson as quoted in Black, 1958, p. 175). Or “*The obscurity and self-contradiction are without any parallel, so far as our experience of mathematical works extends ... no amount of examples can convey an adequate impression of the evils*” (Todhunter, 1949, p. 352 as cited by Young, 1988, p. 1234).

Attempts of clarifying, fully understanding and axiomatizing Condorcet’s approach for solving cycles have been mainly done by Kemeny (1959) who made the first intelligible description of the Condorcet approach, and by Young and Levenglick (1978) who made its clearest exposition and complete axiomatization.

In the version presented by Young and Levenglick (1978), the main methodological foundation is the maximum likelihood concept. In fact, “*Condorcet’s argument proceeds along the following lines. People differ in their opinions because they are imperfect judges of which decision really is best. If on balance each voter is more often right than wrong, however, then the majority view is very likely to identify the decision that is objectively best.*” (Young, 1988, p. 1232).

The maximum likelihood principle selects as a final ranking the one with the maximum pair-wise support. This selected ranking is the one which involves the minimum number of pair-wise inversions. Since Kemeny (1959) proposes the number of pair-wise inversions as a distance to be minimized between the selected ranking and the other individual profiles, the two approaches are perfectly equivalent. A formal proof of this equivalence can be found in Truchon (1988b, pp. 6-10). The selected ranking is also a median ranking for those composing the profile (in multi-criteria terminology it is the “compromise ranking” among the various conflicting points of view), for this reason the corresponding ranking procedure is often known as the Kemeny median order.

Arrow and Raynaud (1986, p. 77) arrive at the conclusion that the highest feasible ambition for an aggregation algorithm building a multi-criterion ranking is to be Condorcet. These authors discard the Kemeny median order, on the grounds that preference reversal phenomena may occur inside this approach (Arrow and Raynaud, 1986, p. 96). However, although the so-called Arrow-Raynaud’s method does not present rank reversal, it is not applicable if cycles

exist. Since in the context where composite indicators are built, cycles are very probable to occur, here the only solution is to follow Kemeny rule (or the equivalent maximum likelihood ranking procedure), thus accepting that rank reversals might appear⁸.

The acceptance of rank reversals phenomena implies that the famous axiom of independence of irrelevant alternatives of Arrow's theorem is not respected. Anyway, Young (1988, p. 1241) claims that the maximum likelihood ranking procedure is the "*only plausible ranking procedure that is locally stable*". Where *local stability* means that the ranking of alternatives does not change if only an interval of the full ranking is considered.

The adaptation of the maximum likelihood ranking procedure to the ranking problem we are dealing with is very simple (Author, 2005). The maximum likelihood ranking of countries is the ranking supported by the maximum number of individual indicators for each pair-wise comparison, summed over all pairs of countries considered. More formally, all the $N(N-1)$ pair-wise comparisons compose the outranking matrix E , where $e_{jk} + e_{kj} = 1$, with $j \neq k$. Call R the set of all $N!$ possible complete rankings of alternatives, $R = \{r_s\}$, $s = 1, 2, \dots, N!$. For each r_s , compute the corresponding score φ_s as the summation of e_{jk} over all the $\binom{N}{2}$ pairs j, k of alternatives, i.e.

$$\varphi_s = \sum e_{jk} \quad (5)$$

where $j \neq k$, $s = 1, 2, \dots, N!$ and $e_{jk} \in r_s$

The final ranking (r^*) is the one which maximises equation (6), which is:

$$r^* \Leftrightarrow \varphi_* = \max \sum e_{jk} \quad \text{where } e_{jk} \in R. \quad (6)$$

Other properties of this ranking procedure are the following (Young and Levenglick, 1978).

⁸ Anyway a Condorcet consistent rule always presents smaller probabilities of the occurrence of a rank reversal in comparison with any Borda consistent rule. This is again a strong argument in favour of a Condorcet's approach in our framework.

- *Neutrality*: it does not depend on the name of any country, all countries are equally treated.
- *Unanimity* (sometimes called *Pareto Optimality*): if all individual indicators prefer country *a* to country *b* then *b* should not be chosen.
- *Monotonicity*: if country *a* is chosen in any pair-wise comparison and only the individual indicator scores (i.e. the variables) of *a* are improved, then *a* should be still the winning country.
- *Reinforcement*: if the set *A* of countries is ranked by 2 subsets G_1 and G_2 of the individual indicator set *G*, such that the ranking is the same for both G_1 and G_2 , then $G_1 \cup G_2 = G$ should still supply the same ranking. This general consistency requirement is very important in the framework of composite indicators, since one may wish to apply the individual indicators belonging to each single dimension first and then pool them in the general model. It has to be noted that the maximum likelihood ranking procedure is the only Condorcet consistent rule which holds the reinforcement property and as noted by Arrow and Raynaud, reinforcement “... has definite ethical content and is therefore relevant to welfare economics and political science.” (Arrow and Raynaud, 1986, p. 96).

3.4 The computational problem

Moulin (1988, p. 312) clearly states that the Kemeny method (or equivalently the maximum likelihood approach) is “*the correct method*” for ranking alternatives, and that the “*only drawback of this aggregation method is the difficulty in computing it when the number of candidates grows*”. In fact the number of permutations can easily become unmanageable; for example when 10 countries are present, it is $10! = 3,628,800$. Indeed this computational drawback is very serious since the Kemeny median order is NP-hard to compute. This NP-hardness has discouraged the development of algorithms searching for exact solutions, thus the majority of the algorithms which have been proposed in the literature; are mainly heuristics based on artificial intelligence, branch and bound approaches and multi-stage techniques (see e.g., Barthélemy et al., 1989; Charon et al., 1997; Cohen et al., 1999; Davenport and Kalagnam, 2004; Dwork et al.,

2001; Truchon, 1998b). Recently, a new numerical algorithm aimed at solving the computational problem connected to linear median orders by finding exact solutions has been developed too (Author, 2006b).

Thanks to the existence of all these computational algorithms, the maximum likelihood (or Kemeny) ranking procedure can always be applied in the context of composite indicators, where a high number of countries to be ranked is the normal state of affairs.

3.5 A sensitive issue: is information on intensity of preference complete lost in a Condorcet framework?

Given that the preference structure is based on equation (2), one might wonder if information on intensity of preference (when variables are measured on an interval or ratio scale) is completely lost in a Condorcet framework (since small and big intensities are treated equally). The problem is indeed very old. Its origins may be found in the famous bold paradox in Greek philosophy: how many hairs one has to cut off to transform a person with hairs to a bold one? Luce (1956) was the first one to discuss this issue formally in the framework of preference modelling. He introduced the idea of the existence of a sensibility threshold below which an agent either does not sense the difference between two elements, or refuses to declare a preference for one or the other. Mathematical characterisations of preference modelling with thresholds can be found in Roubens and Vincke (1985).

By introducing a positive indifference threshold q the resulting preference model is the so-called *threshold model*:

$$\left\{ \begin{array}{l} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) + q \\ a_j I a_k \Leftrightarrow |g_m(a_j) - g_m(a_k)| \leq q \end{array} \right\} \quad (7)$$

where a_j and a_k belong to the set A of countries and g_m to the set G of individual indicators. If one wishes to take into account the possible uncertainty around the value of the threshold q , sensitivity analysis and robustness analyses can be used (Saltelli et al., 2004), another possibility is the use of mathematical sophisticated concept such as the one of fuzzy preference modelling (Author, 1995).

Finally, one should note that the intensity of preference can easily be used in the benchmarking step, where it is not the ranking but the distance from a reference point that matters. For the majority of indicators used in assessment exercises no clear reference point is available, for instance, when GDP is used nobody knows the ideal value of a Country GDP, thus it is quite common to compare with other Countries GDP, e.g. the USA one. A first very simple benchmarking procedure can be the application of a normalisation rule known as “*distance from the group leader*”, which assigns 100 to the leading country in that particular individual indicator and other countries are ranked as percentage points away from the leader (OECD, 2003). More elegant approaches can be based on the so-called ideal point approaches, which is a well established technique in multi-criteria evaluation literature (see e.g. Yu, 1985; Zeleny, 1982). In this framework, to get a set of reference values, an “ideal point” can be defined by choosing the best values reached in any single indicator, and then mathematical distances are used.

3.6 An Illustrative numerical example

Let's take into consideration a simple hypothetical example where there are three countries (A, B, C) to be ranked according to a composite sustainability indicator. Let's assume that three dimensions have to be considered, i.e. the economic, the social and the environmental ones, and that each dimension should have the same weight, that is 0.3333.

The following individual indicators are used

Economic dimension

Indicator: GDP per capita. Weight: 0.165. Objective: maximisation of economic growth. Variable: US dollar per year.

Indicator: Unemployment rate. Weight: 0.165. Objective: minimisation of unemployed people. Variable: percentage of population.

Environmental dimension

Indicator: Solid waste generated per capita. Weight: 0.333. Objective: minimisation of environmental impact. Variable: tons per year.

Social dimension

Indicator: Income disparity. Weight: 0.165. Objective: minimisation of distributional inequity. Variable: Q5/Q1.

Indicator: Crime rate. Weight: 0.165. Objective: minimisation of criminality.
Variable: robberies per 1000 inhabitants.

The impact matrix described in Table 1 can then be constructed.

	Indicators	GDP	Unemp. rate	Solid waste	Inc. dispar.	Crime rate
Countries						
A		22,000	0.17	0.4	10.5	40
B		45,000	0.09	0.45	11.0	45
C		20,000	0.08	0.35	5.3	80

Table 1. Impact matrix of the illustrative numerical example

Considering axioms 1, 2 and 3, and equation (3), the following pair-wise comparisons hold:

$AB=0.333+0.165+0.165=0.666^9$

$BA=0.165+0.165=0.333$

$AC=0.165+0.165=0.333$

$CA=0.165+0.333+0.165=0.666$

$BC=0.165+0.165=0.333$

$CB=0.165+0.333+0.165=0.666$

These results can be synthesised in the outranking matrix:

$$E = \begin{bmatrix} & A & B & C \\ A & 0 & 0.666 & 0.333 \\ B & 0.333 & 0 & 0.333 \\ C & 0.666 & 0.666 & 0 \end{bmatrix}$$

By applying the C-K-Y-L rule to the 3! possible rankings it is:

$ABC \varphi_1 = 0.666 + 0.333 + 0.333 = 1.333$

$BCA \varphi_2 = 0.333 + 0.333 + 0.666 = 1.333$

⁹ Comparing A with B it turns out that A stands out according to the indicators Solid wastes (weight 0.333), Income disparity (weight 0.165) and Crime rate (weight (0.165)).

Manuscript submitted to Applied Economics

$$\text{CAB } \varphi_3 = 0.666 + 0.666 + 0.666 = 2$$

$$\text{ACB } \varphi_4 = 0.333 + 0.666 + 0.666 = 1.666$$

$$\text{BAC } \varphi_5 = 0.333 + 0.333 + 0.333 = 1$$

$$\text{CBA } \varphi_6 = 0.666 + 0.666 + 0.333 = 1.666$$

The final ranking r^* is then CAB.

Notice that using one of the standard ways to produce a composite indicator the result would be different. If for each country the composite indicator is calculated as difference from the group leader (which assigns 100 to the leading country and ranks the others as percentage points away from the leader), the impact matrix becomes:

	Indicators	GDP	Unemp. rate	Solid waste	Inc. dispar.	Crime rate
Countries						
A		48.9	47.05	87.5	50.5	100
B		100	88.9	77.8	48.2	88.9
C		44.4	100	100	100	50

Table 2. Impact matrix: distance from the leader

The index will be calculated averaging each indicator (with the same weights as in the multi-criterion matrix) obtaining $I_A=69.8$, $I_B=79.7$, and $I_C=81.9$. The ranking would be CBA, different from what found with our algorithm.

It is interesting to note that the only sources of uncertainty and imprecision assessment left in the mathematical modelling of the composite indicator proposed are:

1. The scale adjustment technique used (neither normalisation nor a common measurement unit are needed).
2. The values of weights attached to dimensions and individual indicators.

As a consequence, these sources of uncertainty and imprecision assessment are reduced at the minimum possible level and a sensitivity analysis of the results is much easier to carry out than any other composite indicator.

4. A Real-World Application: the Environmental Sustainability

Index

The "Environmental Sustainability Index" (ESI) for 2005 is produced by a team of environmental researchers from the Yale and Columbia Universities, in co-operation with the World Economic Forum and the EU's Joint Research Centre.

The aim of the ESI is that of benchmarking the ability of 146 nations to protect the environment over the next decades, by integrating 76 data sets into 21 indicators of environmental sustainability (see Esty at al., 2005). The data base used to construct the ESI covers a wide range of aspects of environmental sustainability ranging from variables measuring the physical state and stress of the environmental systems (like natural resource depletion, pollution, ecosystem destruction), to the more general social and institutional capacity to respond to environmental challenges. Poverty, short-term thinking and lack of investment in capacity and infrastructure committed to pollution control and ecosystem protection thus compete in determining the measure of a country's sustainability.

Although the official ESI ranking is based upon the linear aggregation of 21 equally weighted indicators, in the methodological appendix, an attempt has been made to apply the non-compensatory approach presented here, in order to tackle the issue of weights as "importance measure" and the compensability of different and crucial dimension of environmental sustainability (see the Methodological Appendix in Esty et al., 2005).

Figure 1 compares the ranking obtained by means of the non-compensatory aggregation rule with the one of the ESI2005. In both cases all 21 indicators are equally weighted. From this figure it is clear that the aggregation method used affects principally the middle-of-the-road and, to a lesser extent, the leader and the laggard countries. Overall for the set of 146 countries, the assumption on the aggregation scheme has an average impact of 8 ranks and a rank-order correlation coefficient of 0.962. In particular, while the top 50 countries only move on

average by 5 positions, the following 50 countries move on average by 12 positions and the latter 46 countries by 8 positions.

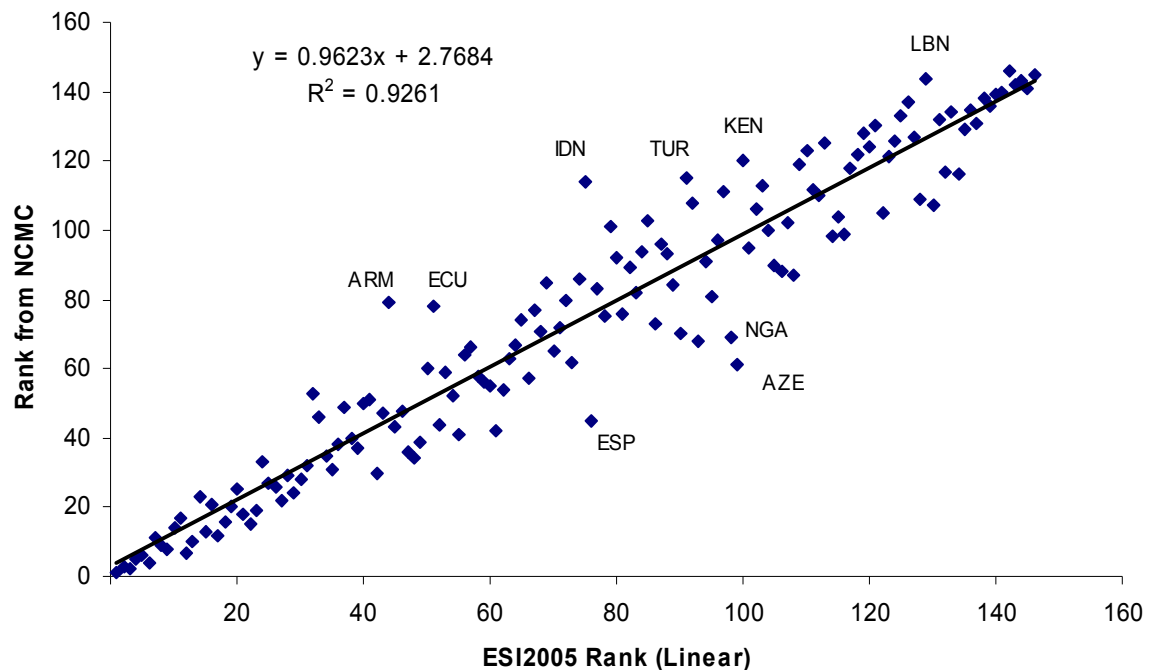


Figure 1. Comparison of rankings obtained by the linear aggregation (ESI2005) and the non-compensatory (NCMA) rules

It is important to underline that although both aggregation schemes seem to produce consistent rankings (the R^2 is 0.92), those rankings do not nevertheless coincide. Using the non-compensatory approach, 43 out of 146 countries display a change in rank higher than 10 positions (none before the 30th ESI rank). When compensability among indicators is not allowed, countries having very poor performance in some indicators, such as Indonesia or Armenia worsen their rank with respect to the linear yardstick, whereas countries that have less extreme values improve their situation, such as Azerbaijan or Spain. Table 3 shows the countries displaying the largest variation in their ranks.

	Aggregation	ESI rank with LIN	rank with NCMC	Change in Rank
Improvement	Azerbaijan	99	61	38
	Spain	76	45	31
	Nigeria	98	69	29
	South Africa	93	68	25
	Burundi	130	107	23
Deterioration	Indonesia	75	114	39
	Armenia	44	79	35
	Ecuador	51	78	27
	Turkey	91	115	24
	Sri Lanka	79	101	22
Average change over 146 countries				8

Table 3. ESI rankings obtained by linear aggregation (LIN) and non-compensatory rule (NCMC): countries that largely improve or worsen their rank position

5. Conclusion

This article has firstly investigated the axiomatic system behind the mathematical modelling of standard composite indicators. The following main conclusions can be drawn:

1. The assumption of preference independence is essential for the existence of a linear aggregation rule. Given to its computational complexity, the test of mutual preferential independence is often not done in real-world practice. Moreover, this assumption has very strong consequences which often are not desirable in a composite indicator.
2. Weights in linear aggregation rules have always the meaning of trade-off ratio. The specification of weights as trade-off is a very complex procedure, for this reason often in practice they are treated as importance coefficients. If weights with the meaning of importance coefficients are used, then a Condorcet approach is needed for axiomatic consistency reasons.
3. In standard composite indicators based on the linear aggregation rule, compensability among the different individual indicators is always

assumed; this implies complete substitutability among the various components considered; this characteristics may not be desirable in many operational frameworks.

A Condorcet consistent non-compensatory/non-linear mathematical aggregation convention for the construction of composite indicators aimed at ranking countries has been developed here. This mathematical aggregation convention can be divided into two main steps:

- pair-wise comparison of countries to be ranked,
- ranking of countries in a complete pre-order.

Weights are never combined with intensities of preference, as a consequence the theoretical guarantee they are importance coefficients is assured. Since intensities of preference are not used the degree of compensability connected with the aggregation model is at the minimum possible level. Given that the summation of weights is equal to one, the pair-wise comparisons can be synthesised in an outranking matrix, which can be interpreted as a voting matrix.

The information contained in the voting matrix is exploited to rank all alternatives in a complete pre-order by using a Condorcet consistent rule. A problem connected with the use of Condorcet consistent rules is the one of cycles. A cycle breaking rule normally needs some arbitrary choices such as to delete the cycle with the lowest support, and so on. A non-arbitrary cycle breaking rule is the so-called Kemeny median order which coincides with the maximum likelihood ranking proposed by Young and Levenglick.

Of course the approach we are proposing is not the “first best” possible approach for the construction of composite indicators. It is a “second best” approach based on theoretical and empirical grounds, which makes assumptions explicit and thus easier to be evaluated in relation with a particular use. This increases both the coherence and the transparency of the ranking obtained. One has to remember, however, that Arrow’s theorem implies that some useful and desirable properties in an aggregation convention are always lost. Here the properties lost are anonymity (which does not constitute a problem since weights are explicitly allowed and decisiveness is gained) and independence of irrelevant alternatives (which on the contrary is a serious loss since rank reversals may

appear). Furthermore, information on intensity of preference of variables is never used. This loss of information is a necessary price to pay for guaranteeing that compensability is reduced and that weights can be considered as symmetrical importance coefficients. However, information on the intensity of preference can be partly used if indifference thresholds are used. Moreover, ranking procedures can be complemented with benchmarking exercises which are fully based on the use of intensity of preference (of course when variables are measured on an interval or ratio scale).

Finally, one has to note that the overall quality of any composite indicator depends on the following elements:

1. *information available* (often data to measure variables are unreliable or simply missing; thus *garbage in, garbage out* phenomena may easily occur (Funtowicz and Ravetz, 1990));
2. *individual indicators and variable chosen* (i.e. which representation of reality is used). This, to a large extent, determines *interpretability* of an index.
3. *direction of each indicator* (i.e. the bigger the better or vice versa, sometimes this decision is not that easy);
4. *relative importance of these indicators* (weights attached which are one of the main sources of technical uncertainty of the results provided; methodological discussion and sensitivity analysis are necessary to tackle this uncertainty by making it explicit);
5. *mathematical aggregation convention used* (the present paper has dealt with this step).

The quality of the aggregation convention is, yet, an important ingredient to guarantee the consistency between the assumptions used and the ranking obtained. Indeed, the overall quality of a composite indicator depends crucially on the way this mathematical model is embedded in the social, political and technical structuring process (Munda, 2004). This is especially true for the choice of weights that remains the most important source of uncertainty and debate. We have the firm conviction that weights are and must be *context-dependent* in that they reflect political, social and economic priorities and depend on the development model a country (or a group of countries) wants to pursue. Often we have seen policymakers using the umbrella of “science” to disguise lobbies,

Manuscript submitted to *Applied Economics*

individual interests, or incompetence. Again, whatever aggregations procedure is used, we think that transparency must remain one of the main ingredients: without *transparency* the *interpretability* and *coherence* of indicators are difficult to achieve.

For Peer Review

APPENDIX

Area	/ Name of Composite Indicator
Economy	
	Composite of Leading Indicators (OECD)
	OECD International Regulation database (OECD)
	Economic Sentiment Indicator (EC)
	Internal Market Index (EC)
	Business Climate Indicator (EC)
Environment	
	Environmental Sustainability Index (World Economic Forum)
	Wellbeing Index (Prescott-Allen)
	Sustainable Development Index (UN)
	Synthetic Environmental Indices (Isla M.)
	Eco-Indicator 99 (Pre Consultants)
	Concern about Environmental Problems (Parker)
	Index of Environmental Friendliness (Puolamaa)
	Environmental Policy Performance Index (Adriaanse)
Globalisation	
	Global Competitiveness Report (World Economic Forum)
	Transnationality Index (UNCTAD)
	Globalisation Index (A.T. Kearny)
	Globalisation Index (World Markets Research Centre)
Society	
	Human Development Index (UN)
	Overall Health Attainment (WHO)
	National Health Care Systems Performance (King's Fund)
	Relative Intensity of Regional Problems (EC)
	Employment Index (Storrie and Bjurek)
Innovation/ Technology	
	Summary Innovation Index (EC)
	Networked Readiness Index (CID)
	National Innovation Capacity Index (Porter and Stern)
	Investment in Knowledge-Based Economy (EC)
	Performance in Knowledge-Based Economy (EC)
	Technology Achievement Index (UN)
	General Indicator of Science and Technology (NISTEP)
	Information and Communications Technologies Index (Fagerberg)
	Success of Software Process Improvement (Emam)

Table 4. Examples of composite indicators

Source: Saisana and Tarantola, 2002, cited in OECD 2003, p. 4

References

- Author** (1995) - *Multicriteria evaluation in a fuzzy environment*, Physica-Verlag, Contributions to Economics Series, Heidelberg.
- Author** (2004) – “Social multi-criteria evaluation (SMCE)”: methodological foundations and operational consequences, *European Journal of Operational Research* Vol. 158, Issue 3: 662-677.
- Author** (2005) – Multi-Criteria Decision Analysis and Sustainable Development, in J. Figueira, S. Greco and M. Ehrgott (eds.) – *Multiple-criteria decision analysis. State of the art surveys*, Springer International Series in Operations Research and Management Science, New York, pp. 953-986.
- Author** (2006a) – Solving the Discrete Multi-Criterion Problem: An Overview of Lessons Learned from Voting Theory, manuscript submitted to *Journal of Multi-Criteria Decision Analysis*.
- Author** (2006b) – A Solution to the Computational Problem of linear median orders, unpublished manuscript.
- Author** (2006) - Constructing Consistent Composite Indicators: the Issue of Weights, *manuscript submitted to Journal of Applied Statistics*.
- Author** ., **Saisana M.**, **Saltelli A.**, **Tarantola S.**, **Hoffman A.**, **Giovannini E.** (2005) – *Handbook on constructing composite indicators: methodology and user guide*, OECD Statistics Working Paper, Paris.
- Anderson N.H.**, **Zalinski J.** (1988) – Functional measurement approach to self-estimation in multiattribute evaluation, *Journal of Behavioural Decision Making* (1), pp. 191-221.
- Arrow K.J.** (1963) - *Social choice and individual values*, 2d edition, Wiley, New York.
- Arrow K.J.**, **Raynaud H.** (1986) - *Social choice and multicriterion decision making*, M.I.T. Press, Cambridge.
- Barthelemy J.P.**, **Guenoche A.**, **Hudry O.** (1989) – Median linear orders: heuristics and a branch and bound algorithm, *European Journal of Operational Research*, 42, pp. 313-325.
- Black D.** (1958) – *The theory of committees and elections*, Cambridge University Press, Cambridge.
- Borda J.C. de** (1784) – Mémoire sur les élections au scrutin, in *Histoire de l' Académie Royale des Sciences*, Paris.
- Bouyssou D.** (1986) – Some remarks on the notion of compensation in MCDM, *European Journal of Operational Research*, 26 pp. 150-160.
- Bouyssou D.**, **Vansnick J.C.** (1986) – Noncompensatory and generalized noncompensatory preference structures, *Theory and Decision*, 21, pp. 251-266.
- Charon I.**, **Guenoche A.**, **Hudry O.**, **Woirgard F.** (1997) – New results on the computation of median orders, *Discrete Mathematics*, 165/166, pp. 139-153.
- Cherchye L.** (2001) – Using data envelopment analysis to assess macroeconomic policy performance, *Applied Economics*, 33, pp. 407-416.
- Cribari-Neto F.**, **Jensen M.J.**, **Novo A.** (1999) - Research in econometric theory: quantitative and qualitative productivity rankings, *Econometric Theory*, Volume 15, Issue 05, pp 719-752.
- Cohen W.**, **Schapiro R.**, **Singer Y.** (1999) – Learning to order things, *Journal of Artificial Intelligence Research*, 10, pp. 213-270.
- Condorcet, Marquis de** (1785) – *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la probabilité des voix*, De l' Imprimerie Royale, Paris.
- 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 Statistical Society*, 155(3), pp. 353-393.
- Davenport A.**, **Kalagnanam J.** (2004) – *A computational study of the Kemeny rule for preference aggregation*, in McGuinness D.L., Ferguson G. – Proceedings of the Nineteenth National Conference on Artificial Intelligence, Sixteenth Conference on Innovative Applications of Artificial Intelligence, July 25-29, 2004, San Jose, California, AAAI Press/The MIT Press, USA.
- Debreu G.** (1960) - Topological methods in cardinal utility theory, in Arrow K.J., Karlin S. and Suppes P. (eds.)- *Mathematical methods in social sciences*, Stanford University Press, Stanford.

- Dietz F.J., van der Straaten J.** (1992) - Rethinking environmental economics: missing links between economic theory and environmental policy, *Journal of Economic Issues*, Vol. XXVI No. 1, pp. 27-51.
- Diewert W.E.** (1976) - Exact and superlative index numbers, *Journal of Econometrics*, Volume 4, Issue 2, Pages 115-145.
- Dwork C., Kumar R., Naor M., Sivakumar D.** (2001) – *Rank aggregation methods for the web*, in Proceedings 10th WWW, pp. 613-622.
- Ebert U., Welsch H.** (2004) – Meaningful environmental indices: a social choice approach, *Journal of Environmental Economics and Management*, 47, pp. 270-283.
- Esty, D.C., Levy M., Srebotnjak, T., de Sherbinin A.** (2005) – *Environmental Sustainability Index: Benchmarking National Environmental Stewardship*. New Haven, Conn.: Yale Center for Environmental Law and Policy.
- Färe R., Grosskopf S., Norris M., Zhang Z.** (1994) – Productivity growth, technical progress and efficiency change in industrialised countries, *American Economic Review*, 84(1), pp. 66-83.
- Fishburn P.C.** (1973) – *The theory of social choice*, Princeton University Press, Princeton.
- Fishburn P.C., Gehrlein W.V. and Maskin E.** (1979) – Condorcet's proportions and Kelly's conjecture, *Discrete Applied Mathematics*, 1, pp. 229-252.
- Forni M., Hallin M., Lippi M., Reichlin L.** (2001) Coincident and Leading Indicators for the Euro Area, *The Economic Journal*, Volume 111, Issue 471, Page 62.
- Funtowicz S.O., Munda G., Paruccini M.** (1990) - The aggregation of environmental data using multicriteria methods, *Environmetrics*, Vol. 1(4), pp. 353-368.
- Funtowicz S.O., Ravetz J.R.** (1990) - *Uncertainty and quality in science for policy*, Kluwer Academic Publishers, Dordrecht.
- Griliches Z.** (1990) – Patent statistics as economic indicators, *Journal of Economic Literature*, 28, pp. 1661-1707.
- Grupp H., Mogee M.E.** (2004) - Indicators for national science and technology policy: how robust are composite indicators?, *Research Policy*, vol. 33(9), pages 1373-1384.
- Guerard J.B.** (2001) – A note on the forecasting effectiveness of the US leading economic indicators, *Indian Economic Review*, 36(1), pp. 251-268.
- Huggins R.** (2003) – Creating a UK competitive index: regional and local benchmarking, *Regional Studies*, 37, pp.89-96.
- Journal of Economic and Social Measurement** (2002) - *Special Issue on Index Number Aggregation*, Volume 28, Numbers 1-2.
- Keeney R., Raiffa H.** (1976) - *Decision with multiple objectives: preferences and value trade-offs*, Wiley, New York.
- Kemeny J.** (1959) – Mathematics without numbers, *Daedalus*, 88, pp. 571-591.
- Kleinknecht A., Van Montfort K., Brouwer E.** (2002) – The non trivial choice between innovation indicators, *Economic Innovation and New Technologies*, 11(2), pp. 109-121.
- Krantz D.H., Luce R.D., Suppes P. and Tversky A.** (1971) - *Foundations of measurement, vol. 1, Additive and polynomial representations*, Academic Press, New York.
- Leontief W.** (1947a) – Introduction to a theory of the internal structure of functional relationships, *Econometrica*, 15, pp. 361 – 373.
- Leontief W.** (1947b) – A note on the interrelation of subsets independent variables of a continuous function with continuous first derivatives, *Bulletin of the American Mathematical Society*, 53, pp. 343 – 350.
- 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, pp. 507-518.
- Luce R.D.** (1956) – Semiorders and a theory of utility discrimination, *Econometrica*, 24, pp.178-191.
- Moulin H.** (1985) – From social welfare orderings to acyclic aggregation of preferences, *Mathematical Social Sciences*, 9, pp. 1-17.
- Moulin H.** (1988) - *Axioms of co-operative decision making*, Econometric Society Monographs, Cambridge University Press, Cambridge.
- Nicoletti G., Scarpetta S. and Boylaud O.** (2000) – Summary indicators of product market regulation with an extension to employment protection legislation, *OECD Economics Department Working papers No. 226*.
- OECD** (2003) – *Composite indicators of country performance: a critical assessment*, DST/IND(2003)5, Paris.

- O'Neill, J.** (1993) - *Ecology, Policy and Politics*, Routledge, London.
- Podinovskii V.V.** (1994) - Criteria importance theory, *Mathematical Social Sciences*, 27, pp. 237 - 252.
- Roubens M., Vincke Ph.** (1985) - *Preference modelling*, Springer-Verlag, Heidelberg.
- Roy B.** (1996) - *Multicriteria methodology for decision analysis*, Kluwer, Dordrecht.
- Saisana M. and Tarantola S.** (2002) – *State-of-the-art report on current methodologies and practices for composite indicator development*, EUR 20408 EN Report, European Commission, JRC, Ispra, Italy.
- Saltelli A. Tarantola S., Campolongo, F. and Ratto, M.** (2004) - *Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models*, John Wiley & Sons publishers, New York.
- Ting H.M.** (1971) - *Aggregation of attributes for multiattributed utility assessment*, Technical report N° 66, Operations Research Center, M.I.T. Cambridge, Mass.
- Todhunter I.** (1949) – *A history of mathematical theory of probability*, Chelsea, New York.
- Truchon M.** (1995) – Voting games and acyclic collective choice rules, *Mathematical Social Sciences*, 25, pp. 165-179.
- Truchon M.** (1998a) – *Figure skating and the theory of social choice*, Cahier 9814 du Centre de Recherche en Economie et Finance Appliqués (CREFA).
- Truchon M.** (1998b) – *An extension of the Condorcet criterion and Kemeny orders*, Cahier 9813 du Centre de Recherche en Economie et Finance Appliquées (CREFA).
- Vansnick J. C.** (1990) - Measurement theory and decision aid- in Bana e Costa C.A. (ed.)- *Readings in multiple criteria decision aid*, Springer-Verlag, Berlin, pp. 81-100.
- Vidu L.** (2002) – Majority cycles in a multi-dimensional setting, *Economic Theory*, 20, pp. 373-386.
- Vincke Ph.** (1992) - *Multicriteria decision aid*, Wiley, New York.
- Weber J.** (2002) – How many voters are needed for paradoxes?, *Economic Theory*, 20, pp. 341-355.
- Wilson J.W., and Jones C.P.** (2002) – An analysis of the S&P 500 index and Cowles's extensions: price indexes and stock returns, 1870-1999, *Journal of Business*, 75, pp. 505-533.
- Young H.P. and Levenglick A.** (1978) – A consistent extension of Condorcet's election principle, *SIAM Journal of Applied Mathematics*, 35, pp. 285-300.
- Young H.P.** (1988) – Condorcet's theory of voting, *American Political Science Review*, Vol. 82, No. 4, pp. 1231-1244.
- Yu P. L.** (1985) - *Multi criteria decision making: concepts, techniques and extensions*, Plenum Press, New York.
- Zeleny M.** (1982) - *Multiple criteria decision making*, McGraw Hill, New York.

Version September, 2006

Non-Compensatory/Non-Linear Composite Indicators for Ranking Countries: A Defensible Setting

Giuseppe Munda* and Michela Nardo

European Commission, Joint Research Centre
Institute for the Protection and Security of the Citizen (IPSC)
Econometrics and Statistical Support to Antifraud (G-09)
TP 361
21020 Ispra (VA)
Italy
e_mail: giuseppe.munda@jrc.it
e_mail: michela.nardo@jrc.it

Abstract: Composite indicators (or indexes) are very common in economic and business statistics for benchmarking the mutual and relative progress of countries in a variety of policy domains such as industrial competitiveness, sustainable development, globalisation and innovation. The proliferation of the production of composite indicators by all the major international organizations is a clear symptom of their political importance and operational relevance in policy-making. As a consequence, improvements in the way these indicators are constructed and used seem to be a very important research issue from both the theoretical and operational points of view. This paper aims at contributing to the improvement of the overall quality of composite indicators (or indexes) by looking at one of their technical weaknesses, that is, the aggregation convention used for their construction. For this aim, we build upon concepts coming from multi-criteria decision analysis, measurement theory and social choice. We start from the analysis of the axiomatic system underlying the mathematical modelling commonly used to construct composite indicators. Then a different methodological framework, based on non-compensatory/non-linear aggregation rules, is developed. Main features of the proposed approach are: (i) the axiomatic system is made completely explicit, and (ii) the sources of technical uncertainty and imprecise assessment are reduced to the minimum possible degree.

Keywords: *Index Numbers and Aggregation, Multi-Criteria Decision Analysis, Social Choice, Environmental Sustainability Index*

JEL Classification Numbers: *C43, C82, Q01*

*Corresponding Author, also Universitat Autonoma de Barcelona, dept. of Economics

1. Introduction

Composite indicators (along all this paper the term *composite indicator* is used as a synonymous of *index*) are very common in fields such as economic and business statistics (e.g., Öcal 2006, Binner et al. 1999 and 2006 or the OECD Composite of Leading Indicators) and are used in a variety of policy domains such as industrial competitiveness, sustainable development, quality of life assessment, globalisation, innovation or academic performance (see Cox and others 1992, Cribari-Neto et al 1999, Färe et al. 1994, Griliches 1990, Forni et al. 2001, Huggins 2003, Grupp and Mogege 2004, Lovell et al. 1995, Munda 2005, Nardo et al. 2005, Saisana and Tarantola 2002, and Wilson and Jones 2002, among others). The proliferation of composite indicators is a clear symptom of their importance in policy-making, and operational relevance in economic statistics in general (see e.g. Granger, 2001). All the major international organizations such as OECD, the EU, the World Economic Forum or the IMF are producing composite indicators in a wide variety of fields (Nardo et al., 2005). A general objective of most of these indicators is the ranking of countries and their benchmarking according to some aggregated dimensions (see e.g. Cherchye, 2001, Kleinknecht 2002 and OECD, 2003).

The improvement of the way composite indicators are constructed and used seems to be a very important research issue from both theoretical and operational points of view. Our main objective in this article is to contribute to the improvement of the overall quality of composite indicators by looking at one of their technical weaknesses, that is, the aggregation convention used for their construction. For this aim, we build upon concepts coming from multi-criteria decision analysis, measurement theory and social choice.

Although various functional forms for the underlying aggregation rules of a composite indicator have been developed in the literature (e.g. Diewert, 1976, Journal of Economic and Social Measurement, 2002), in the standard practice, a composite indicator I , can be considered a weighted linear aggregation rule applied to a set of variables (OECD, 2003, p. 5):

$$I = \sum_{i=1}^N w_i x_i, \quad (1)$$

where x_i is a scale adjusted variable (e.g. GDP per capita) normalized between zero and one, and w_i a weight attached to x_i , usually with $\sum_{i=1}^N w_i = 1$ and $0 \leq w_i \leq 1, i = 1, 2, \dots, N$.

Munda and Nardo (2006) analyse the formal axioms behind the linear aggregation rule and their operational implications and conclude that the use of non-linear aggregation rules to construct composite indicators is **compulsory** for reasons of theoretical consistency when weights with the meaning of *importance coefficients*¹ are used or when the assumption of preferential independence does not hold. Moreover, in standard linear composite indicators, compensability among the different individual indicators is always assumed; this implies complete substitutability among the various components considered. For example, in a hypothetical sustainability index, economic growth can always substitute any environmental destruction or inside e.g., the environmental dimension, clean air can compensate for a loss of potable water. From a normative point of view, such a complete compensability **is often not desirable**.

Vansnick (1990) showed that the two main approaches in multi-criteria aggregation procedures i.e., the compensatory and non-compensatory ones can be directly derived from the seminal work in social choice of Borda (father of the compensatory approach) and Condorcet (father of the non-compensatory approach)². If one wants the weights to be interpreted as “importance coefficients” (or equivalently *symmetrical importance* of variables) non-compensatory aggregation procedures must be used (Bouyssou, 1986; Bouyssou and Vansnick, 1986). The majority rule is theoretically the most desirable aggregation rule, but practically often produces undesirable intransitivities, thus

¹ In the decision theory literature, this concept of weights as importance coefficients is usually referred to as *symmetrical importance*, that is “... if we have two non-equal numbers to construct a vector in R^2 , then it is preferable to place the greatest number in the position corresponding to the most important criterion.” (Podinovskii, 1994, p. 241).

² Ebert and Welsch (2004) also propose the use social choice to improve the theoretical framework of environmental indexes.

“more limited ambitions are compulsory. The next highest ambition for an aggregation algorithm is to be Condorcet” (Arrow and Raynaud, 1986, p. 77).

Finally, one should note that to use a linear aggregation rule, the assumption that the variable scores are measured on an interval or ratio scale of measurement and no uncertainty exists must always apply. Rarely this happens in the practice of composite indicators, where for instance, sometimes quantitative scores are arbitrarily given to variable scores originally measured on an ordinal measurement scale (see e.g. Nicoletti et al., 2000). On the contrary, by using Condorcet aggregation rules no limitation on the measurement scale of the variable scores exists.

For all these reasons, we think that in some applications the use of non-linear/non-compensatory Condorcet consistent aggregation rules is desirable. Since this possibility has almost never been explored in the framework of composite indicators the following Section is devoted to this issue.

2. An Axiomatic Approach for the Construction of Non-Compensatory/Non-Linear Condorcet Consistent Composite Indicators

For the sake of clearness, some basic definitions are first given. These definitions are adapted to the context of composite indicators borrowing concepts from multi-criteria decision theory and complex system theory³.

2.1 Basic definitions

Dimension: is the highest hierarchical level of analysis and indicates the scope of objectives, individual indicators and variables. For example, a sustainability composite indicator can include economic, social and environmental dimensions.

Objective: an objective indicates the direction of change desired. For example, within the economic dimension GDP has to be maximised; within the social dimension social exclusion has to be minimised; within the environmental dimension CO₂ emissions have to be minimised.

³ Some of these definitions were inspired by discussions with M. Giampietro.

Individual indicator: it is the basis for evaluation in relation to a given objective (any objective may imply a number of different individual indicators). It is a *function* that associates each single country with a variable indicating its desirability according to expected consequences related to the same objective. For example, GDP, saving rate and inflation rate inside the objective “growth maximisation”.

Variable: is a constructed measure stemming from a process that represents, at a given point in space and time, a shared perception of a real-world state of affairs consistent with a given individual indicator. To give an example, in comparing two countries, inside the economic dimension, one objective can be “maximisation of economic growth”; the individual indicator might be R&D performance, the indicator score or variable can be “number of patents per million of inhabitants”. Another example: an objective connected with the social dimension can be “maximisation of the residential attractiveness”. A possible individual indicator is then “residential density”. The variable providing the individual indicator score might be the ratio persons per hectare.

A composite indicator or synthetic index is an aggregate of all dimensions, objectives, individual indicators and variables used. This implies that what formally defines a composite indicator is the *set of properties underlying its aggregation convention*. The rest of this section deals with this issue.

When various individual indicators are used to evaluate two different countries, some of these individual indicators may be in favour of country *a* while other variables may be in favour of country *b*. As a consequence a conflict among the individual indicators exists. How this conflict can be treated at the light of a non-linear/non-compensatory logic? This is the classical multi-criteria discrete problem (Munda, 1995; Roy, 1996; Vincke, 1992). With this analogy in mind, we present an aggregation convention for (non-linear and non-compensatory) composite indicators able to rank different countries (or regions, cities and so on).

2.2 Problem definition

Given a set of individual indicators $G=\{g_m\}$, $m=1,2,..., M$, and a finite set

Manuscript revised for Applied Economics

$A=\{a_n\}$, $n=1, 2, \dots, N$ of countries, let's assume that the variable (i.e. the individual indicator score) of each country a_n with respect to an individual indicator g_m is based on an *ordinal, interval or ratio* scale of measurement. For simplicity of exposition, we assume that a higher value of a variable is preferred to a lower one (i.e. the higher, the better), that is:

$$\begin{cases} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) \\ a_j I a_k \Leftrightarrow g_m(a_j) = g_m(a_k) \end{cases} \quad (2)$$

Where, P and I indicate a preference and an indifference relation respectively, both fulfilling the transitive property.

Let's also assume the existence of a set of individual indicator weights $W=\{w_m\}$, $m=1, 2, \dots, M$, with $\sum_{m=1}^M w_m = 1$, derived as importance coefficients. The mathematical problem to be dealt with is then how to use this available information to rank in a complete pre-order (i.e. without any incomparability relation) all the countries from the best to the worst one. In doing so the following operational properties are desirable:

1. The sources of uncertainty and imprecise assessment should be reduced as much as possible.
2. The manipulation rules should be the more objective and as simple as possible, that is all *ad hoc* parameters should be avoided.
3. A theoretical guarantee that weights are used with the meaning of "symmetrical importance" must exist. As a consequence, complete compensability should be avoided. This entails that variables have to be used with an ordinal meaning. This is not a problem since no loss of information is implied (Arrow and Raynaud, 1986). Moreover, given that often the measurement of variables is imprecise (see OECD, 2003, p.7), it seems even desirable to use indicator scores with an ordinal meaning.
4. Desirable ranking procedures using ordinal information are always of a Condorcet type (Arrow and Raynaud, 1986, Moulin, 1988). A problem inherent to this family of algorithm is the presence of cycles, i.e. cases where aPb , bPc and cPa . This problem has been widely studied among

others by Fishburn, 1973; Fishburn et al., 1979; Kemeny, 1959; Moulin, 1985; Truchon, 1995; Young and Levenglick, 1978, Vidu, 2002; Weber, 2002. The probability $\pi(N, M)$ of obtain a cycle with N countries and M individual indicators increases with N as well as with the number of indicators. With many countries and individual indicators, cycles occur with an extremely high frequency. Therefore, the ranking procedure used has to deal with the cycle issue properly.

5. Arrow's impossibility theorem (Arrow, 1963) clearly shows that no perfect aggregation convention can exist. Then, it is essential to check not only which properties are respected by a given ranking procedure, but also if any essential property for the problem tackled is lost.

2.3 The proposed composite indicator

The mathematical aggregation convention we are proposing can be divided into two main steps:

1. Pair-wise comparison of countries according to the whole set of individual indicators used.
2. Ranking of countries in a complete pre-order.

For carrying out the pair-wise comparison of countries the following axiomatic system is needed (adapted from Arrow and Raynaud, 1986, p. 81-82).

Axiom 1: Diversity. Each individual indicator is a total order on the finite set A of countries to be ranked, and there is no restriction on the individual indicators; they can be any total order on A .

Axiom 2: Symmetry. Since individual indicators have incommensurable scales, the only preference information they provide is the ordinal pair-wise preferences they contain⁴.

Axiom 3: Positive Responsiveness. The degree of preference between two countries a and b is a strictly increasing function of the number and weights of individual indicators that rank a before b ⁵.

⁴ In our case, this axiom is needed since the intensity of preference (i.e. the difference between 2 variables measured on an interval or ratio scale) of individual indicators is not considered useful preference information given that compensability has to be avoided and weights have to be symmetrical importance coefficients. Moreover, thanks to this axiom, a normalisation step is not needed. This causes a further reduction of the sources of uncertainty and imprecise assessment.

Thanks to these three axioms a $N \times N$ matrix, E , called *outranking matrix* (Arrow and Raynaud, 1986, Roy, 1996) can be built. Any generic element of the matrix E , e_{jk} , $j \neq k$ is the result of the pair-wise comparison, according to all the M individual indicators, between countries j and k . Such a global pair-wise comparison is obtained by means of equation (2).

$$e_{jk} = \sum_{m=1}^M \left(w_m(P_{jk}) + \frac{1}{2} w_m(I_{jk}) \right) \quad (3)$$

where $w_m(P_{jk})$ and $w_m(I_{jk})$ are the weights of individual indicators presenting a preference and an indifference relation respectively. It clearly holds

$$e_{jk} + e_{kj} = 1. \quad (4)$$

Property (4), although obvious, is very important since it allows us to consider the outranking matrix E as a *voting matrix* i.e., a matrix where instead of using individual indicators, alternatives are compared by means of voters' preferences (with the principle one agent one vote). This analogy between a multi-criterion problem and a social choice one, as noted by Arrow and Raynaud (1986), is very useful for tackling the step of ranking the N countries in a consistent axiomatic framework.

The issue is now to exploit the information contained in the outranking matrix in order to rank all countries in a complete pre-order. A problem connected with the use of Condorcet consistent rules is that of cycles. A cycle breaking rule normally needs some arbitrary choices such as to delete the cycle with the lowest support. Now the question is: Is it possible to tackle the cycle issue in a more general way?

⁵ In social choice terms then the *anonymity* property (i.e. equal treatment of all individual indicators) is broken. Indeed, given that full decisiveness yields to dictatorship, Arrow's impossibility theorem forces us to make a trade-off between *decisiveness* (an alternative has to be chosen or a ranking has to be made) and anonymity. In our case the loss of anonymity in favour of decisiveness is even a positive property. In general, it is essential that no individual indicator weight is more than 50% of the total weight; otherwise the aggregation procedure would become lexicographic in nature, and the indicator would become a dictator in Arrow's term.

Condorcet himself was aware of the problem of cycles in his approach; he built examples to explain it and he was even close to find a consistent rule able to rank any number of alternatives when cycles are present. However, attempts to fully understand this part of Condorcet's voting theory have arrived at conclusions like "... *the general rules for the case of any number of candidates as given by Condorcet are stated so briefly as to be hardly intelligible ... and as no examples are given it is quite hopeless to find out what Condorcet meant*" (E.J. Nanson as quoted in Black, 1958, p. 175). Or "*The obscurity and self-contradiction are without any parallel, so far as our experience of mathematical works extends ... no amount of examples can convey an adequate impression of the evils*" (Todhunter, 1949, p. 352 as cited by Young, 1988, p. 1234).

Attempts of clarifying, fully understanding and axiomatizing Condorcet's approach for solving cycles have been mainly done by Kemeny (1959) who made the first intelligible description of the Condorcet approach, and by Young and Levenglick (1978) who made its clearest exposition and complete axiomatization.

In the version presented by Young and Levenglick (1978), the main methodological foundation is the maximum likelihood concept. In fact, "*Condorcet's argument proceeds along the following lines. People differ in their opinions because they are imperfect judges of which decision really is best. If on balance each voter is more often right than wrong, however, then the majority view is very likely to identify the decision that is objectively best.*" (Young, 1988, p. 1232).

The maximum likelihood principle selects as a final ranking the one with the maximum pair-wise support. This selected ranking is the one which involves the minimum number of pair-wise inversions. Since Kemeny (1959) proposes the number of pair-wise inversions as a distance to be minimized between the selected ranking and the other individual profiles, the two approaches are perfectly equivalent. A formal proof of this equivalence can be found in Truchon (1988b, pp. 6-10). The selected ranking is also a median ranking for those composing the profile (in multi-criteria terminology it is the "compromise ranking" among the various conflicting points of view), for this reason the corresponding ranking procedure is often known as the Kemeny median order.

Arrow and Raynaud (1986, p. 77) arrive at the conclusion that the highest feasible ambition for an aggregation algorithm building a multi-criterion ranking

is to be Condorcet. These authors discard the Kemeny median order, on the grounds that preference reversal phenomena may occur inside this approach (Arrow and Raynaud, 1986, p. 96). However, although the so-called Arrow-Raynaud's method does not present rank reversal, it is not applicable if cycles exist. Since in the context where composite indicators are built, cycles are very probable to occur, here the only solution is to follow Kemeny rule (or the equivalent maximum likelihood ranking procedure), thus accepting that rank reversals might appear⁶.

The acceptance of rank reversals phenomena implies that the famous axiom of independence of irrelevant alternatives of Arrow's theorem is not respected. Anyway, Young (1988, p. 1241) claims that the maximum likelihood ranking procedure is the "*only plausible ranking procedure that is locally stable*". Where *local stability* means that the ranking of alternatives does not change if only an interval of the full ranking is considered.

The adaptation of the maximum likelihood ranking procedure to the ranking problem we are dealing with is reasonably simple (Munda, 2005). The maximum likelihood ranking of countries is the ranking supported by the maximum number of individual indicators for each pair-wise comparison, summed over all pairs of countries considered. More formally, all the $N(N-1)$ pair-wise comparisons compose the outranking matrix E , where $e_{jk} + e_{kj} = 1$, with $j \neq k$. Call R the set of all $N!$ possible complete rankings of alternatives, $R = \{r_s\}$, $s = 1, 2, \dots, N!$. For each r_s , compute the corresponding score φ_s as the summation of e_{jk} over all the $\binom{N}{2}$ pairs j, k of alternatives, i.e.

$$\varphi_s = \sum e_{jk} \quad . \quad (5)$$

where $j \neq k$, $s = 1, 2, \dots, N!$ and $e_{jk} \in r_s$

The final ranking (r^*) is the one which maximises equation (6), which is:

⁶ Anyway a Condorcet consistent rule always presents smaller probabilities of the occurrence of a rank reversal in comparison with any Borda consistent rule. This is again a strong argument in favour of a Condorcet's approach in our framework.

$$r^* \Leftrightarrow \varphi_* = \max \sum e_{jk} \quad \text{where } e_{jk} \in R. \quad (6)$$

Other properties of this ranking procedure are the following (Young and Levenglick, 1978).

- *Neutrality*: it does not depend on the name of any country, all countries are equally treated.
- *Unanimity* (sometimes called *Pareto Optimality*): if all individual indicators prefer country *a* to country *b* then *b* should not be chosen.
- *Monotonicity*: if country *a* is chosen in any pair-wise comparison and only the individual indicator scores (i.e. the variables) of *a* are improved, then *a* should be still the winning country.
- *Reinforcement*: if the set *A* of countries is ranked by 2 subsets *G*₁ and *G*₂ of the individual indicator set *G*, such that the ranking is the same for both *G*₁ and *G*₂, then *G*₁ ∪ *G*₂ = *G* should still supply the same ranking. This general consistency requirement is very important in the framework of composite indicators, since one may wish to apply the individual indicators belonging to each single dimension first and then pool them in the general model. It has to be noted that the maximum likelihood ranking procedure is the only Condorcet consistent rule which holds the reinforcement property and as noted by Arrow and Raynaud, reinforcement “... has definite ethical content and is therefore relevant to welfare economics and political science.” (Arrow and Raynaud, 1986, p. 96).

2.4 The computational problem

Moulin (1988, p. 312) clearly states that the Kemeny method (or equivalently the maximum likelihood approach) is “the correct method” for ranking alternatives, and that the “only drawback of this aggregation method is the difficulty in computing it when the number of candidates grows”. In fact the number of permutations can easily become unmanageable; for example when 10 countries are present, it is 10!=3,628,800. Indeed this computational drawback is very serious since the Kemeny median order is NP-hard to compute. This NP-hardness has discouraged the development of algorithms searching for exact

solutions, thus the majority of the algorithms which have been proposed in the literature; are mainly heuristics based on artificial intelligence, branch and bound approaches and multi-stage techniques (see e.g., Barthelemy et al., 1989; Charon et al., 1997; Cohen et al., 1999; Davenport and Kalagnam, 2004; Dwork et al., 2001; Truchon, 1998b). Recently, a new numerical algorithm aimed at solving the computational problem connected to linear median orders by looking for exact solutions has been developed too (Munda, 2006).

Thanks to the existence of all these efficient computational algorithms, the maximum likelihood (or Kemeny) ranking procedure can always be applied in the context of composite indicators, where a high number of countries to be ranked is the normal state of affairs.

2.5 A sensitive issue: is information on intensity of preference complete lost in a Condorcet framework?

Given that the preference structure is based on equation (2), one might wonder if information on intensity of preference (when variables are measured on an interval or ratio scale) is completely lost in a Condorcet framework (since small and big intensities are treated equally). The problem is indeed very old. Its origins may be found in the famous bold paradox in Greek philosophy: how many hairs one has to cut off to transform a person with hairs to a bold one? Luce (1956) was the first one to discuss this issue formally in the framework of preference modelling. He introduced the idea of the existence of a sensibility threshold below which an agent either does not sense the difference between two elements, or refuses to declare a preference for one or the other. Mathematical characterisations of preference modelling with thresholds can be found in Roubens and Vincke (1985).

By introducing a positive indifference threshold q the resulting preference model is the so-called *threshold model*:

$$\left\{ \begin{array}{l} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) + q \\ a_j I a_k \Leftrightarrow |g_m(a_j) - g_m(a_k)| \leq q \end{array} \right\} \quad (7)$$

where a_j and a_k belong to the set A of countries and g_m to the set G of individual indicators. If one wishes to take into account the possible uncertainty around the value of the threshold q , sensitivity analysis and robustness analyses can be used (Saltelli et al., 2004), another possibility is the use of mathematical sophisticated concept such as the one of fuzzy preference modelling (Munda, 1995).

Finally, one should note that the intensity of preference can easily be used in the benchmarking step, where is not the ranking but the distance from a reference point what matters. For the majority of indicators used in assessment exercises no clear reference point is available, for instance, when GDP is used nobody knows the ideal value of a Country GDP, thus it is quite common to compare with other Countries GDP, e.g. the USA one. A first very simple benchmarking procedure can be the application of a normalisation rule known as “*distance from the group leader*”, which assigns 100 to the leading country in that particular individual indicator and other countries are ranked as percentage points away from the leader (OECD, 2003). More elegant approaches can be based on the so-called ideal point approaches, which is a well established technique in multi-criteria evaluation literature (see e.g. Yu, 1985; Zeleny, 1982). In this framework, to get a set of reference values, an “ideal point” can be defined by choosing the best values reached in any single indicator, and then mathematical distances are used.

3.6 An Illustrative numerical example

Let’s take into consideration a simple hypothetical example where there are three countries (A, B, C) to be ranked according to a composite sustainability indicator. Let’s assume that three dimensions have to be considered, i.e. the economic, the social and the environmental ones, and that each dimension should have the same weight, that is 0.3333.

The following individual indicators are used

Economic dimension

Indicator: GDP per capita. Weight: 0.165. Objective: maximisation of economic growth. Variable: US dollar per year.

Indicator: Unemployment rate. Weight: 0.165. Objective: minimisation of unemployed people. Variable: percentage of population.

Environmental dimension

Manuscript revised for Applied Economics

Indicator: Solid waste generated per capita. Weight: 0.333. Objective: minimisation of environmental impact. Variable: tons per year.

Social dimension

Indicator: Income disparity. Weight: 0.165. Objective: minimisation of distributional inequity. Variable: Q5/Q1.

Indicator: Crime rate. Weight: 0.165. Objective: minimisation of criminality. Variable: robberies per 1000 inhabitants.

The impact matrix described in Table 1 can then be constructed.

	Indicators	GDP	Unemp. rate	Solid waste	Inc. dispar.	Crime rate
Countries						
A		22,000	0.17	0.4	10.5	40
B		45,000	0.09	0.45	11.0	45
C		20,000	0.08	0.35	5.3	80

Table 1. Impact matrix of the illustrative numerical example

Considering axioms 1, 2 and 3, and equation (3), the following pair-wise comparisons hold:

$$AB=0.333+0.165+0.165=0.666^7$$

$$BA=0.165+0.165=0.333$$

$$AC=0.165+0.165=0.333$$

$$CA=0.165+0.333+0.165=0.666$$

$$BC=0.165+0.165=0.333$$

$$CB=0.165+0.333+0.165=0.666$$

These results can be synthesised in the outranking matrix:

$$E = \begin{bmatrix} & A & B & C \\ A & 0 & 0.666 & 0.333 \\ B & 0.333 & 0 & 0.333 \\ C & 0.666 & 0.666 & 0 \end{bmatrix}$$

⁷ Comparing A with B it turns out that A stands out according to the indicators Solid wastes (weight 0.333), Income disparity (weight 0.165) and Crime rate (weight 0.165).

By applying the Kemeny rule to the 3! possible rankings it is:

ABC $\varphi_1 = 0.666 + 0.333 + 0.333 = 1.333$

BCA $\varphi_2 = 0.333 + 0.333 + 0.666 = 1.333$

CAB $\varphi_3 = 0.666 + 0.666 + 0.666 = 2$

ACB $\varphi_4 = 0.333 + 0.666 + 0.666 = 1.666$

BAC $\varphi_5 = 0.333 + 0.333 + 0.333 = 1$

CBA $\varphi_6 = 0.666 + 0.666 + 0.333 = 1.666$

The final ranking r^* is then CAB.

Notice that using one of the standard ways to produce a composite indicator the result would be different. If for each country the composite indicator is calculated as difference from the group leader (which assigns 100 to the leading country and ranks the others as percentage points away from the leader), the impact matrix becomes:

	Indicators	GDP	Unemp. rate	Solid waste	Inc. dispar.	Crime rate
Countries						
A		48.9	47.05	87.5	50.5	100
B		100	88.9	77.8	48.2	88.9
C		44.4	100	100	100	50

Table 2. Impact matrix: distance from the leader

The index will be calculated averaging each indicator (with the same weights as in the multi-criterion matrix) obtaining $I_A=69.8$, $I_B=79.7$, and $I_C=81.9$. The ranking would be CBA, different from what found with our non-linear/non-compensatory proposal.

It is interesting to note that the only sources of uncertainty and imprecision assessment left in the mathematical modelling of the composite indicator proposed are:

1. The scale adjustment technique used (neither normalisation nor a common measurement unit are needed).

2. The values of weights attached to dimensions and individual indicators.

As a consequence, these sources of uncertainty and imprecision assessment are reduced at the minimum possible level and a sensitivity analysis of the results is much easier to carry out than any other composite indicator.

3. A Real-World Application: the Environmental Sustainability Index

The "Environmental Sustainability Index" (ESI) for 2005 is produced by a team of environmental researchers from the Yale and Columbia Universities, in co-operation with the World Economic Forum and the EU's Joint Research Centre. The aim of the ESI is that of benchmarking the ability of 146 nations to protect the environment over the next decades, by integrating 76 data sets into 21 individual indicators of environmental sustainability (see Esty et al., 2005). The data base used to construct the ESI covers a wide range of aspects of environmental sustainability ranging from variables measuring the physical state and stress of the environmental systems (like natural resource depletion, pollution, ecosystem destruction), to the more general social and institutional capacity to respond to environmental challenges. Poverty, short-term thinking and lack of investment in capacity and infrastructure committed to pollution control and ecosystem protection thus compete in determining the measure of a country's sustainability.

Although the official ESI ranking is based upon the linear aggregation of 21 equally weighted individual indicators, in the methodological appendix, an attempt has been made to apply the non-compensatory approach presented here, in order to tackle the issue of weights as "importance measure" and the compensability of different and crucial dimension of environmental sustainability (see the Methodological Appendix in Esty et al., 2005).

Figure 1 compares the ranking obtained by means of the non-linear/non-compensatory aggregation rule with the one of the ESI2005. In both cases all 21 individual indicators are equally weighted. From this figure it is clear that the aggregation method used affects mainly the middle-of-the-road and, to a lesser extent, the leader and the laggard countries. Overall for the set of 146 countries,

the assumption on the aggregation scheme has an average impact of 8 ranks and a rank-order correlation coefficient of 0.962. In particular, while the top 50 countries only move on average by 5 positions, the following 50 countries move on average by 12 positions and the latter 46 countries by 8 positions.

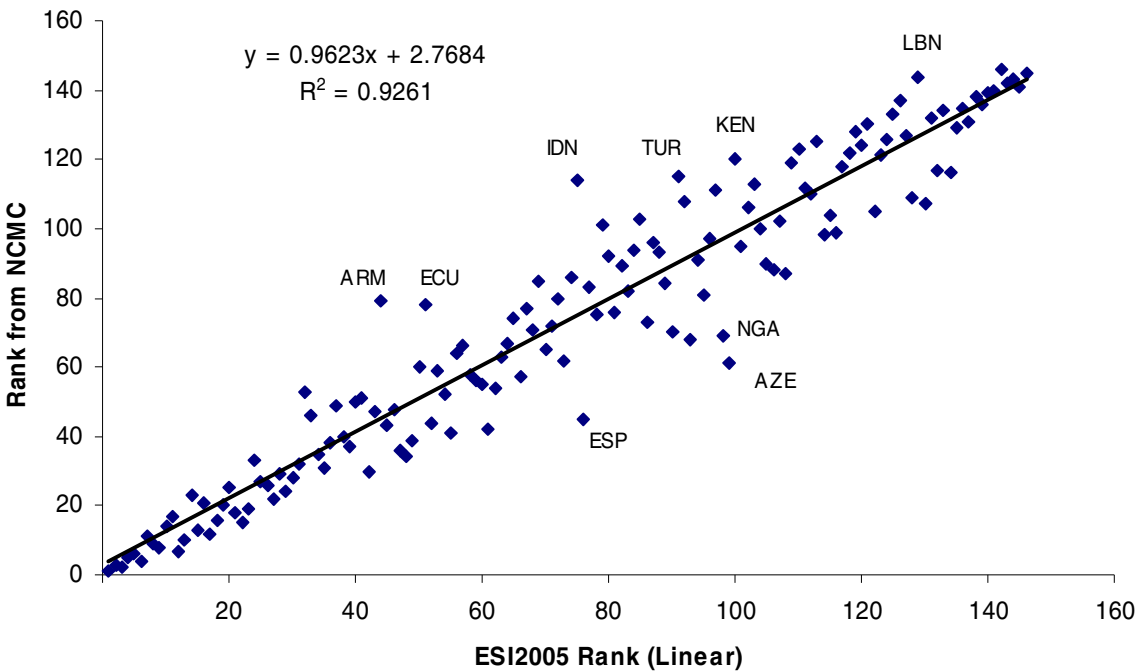


Figure 1. Comparison of rankings obtained by the linear aggregation (named ESI2005 in the x-axis) and the non-linear/non-compensatory (named NCMA in the y-axis) rules

It is important to underline that although both aggregation rules seem to produce consistent rankings (the R^2 is 0.92), those rankings do not nevertheless coincide. Using the non-linear/non-compensatory approach, 43 out of 146 countries display a change in rank higher than 10 positions (none before the 30th ESI2005 rank). When compensability among indicators is not allowed, countries having very poor performance in some indicators, such as Indonesia or Armenia worsen their rank with respect to the linear yardstick, whereas countries that have less extreme values improve their situation, such as Azerbaijan or Spain. Table 3 shows the countries displaying the largest variation in their ranks.

	Aggregation	ESI rank with LIN	rank with NCMC	Change in Rank
Improvement	Azerbaijan	99	61	38
	Spain	76	45	31
	Nigeria	98	69	29
	South Africa	93	68	25
	Burundi	130	107	23
Deterioration	Indonesia	75	114	39
	Armenia	44	79	35
	Ecuador	51	78	27
	Turkey	91	115	24
	Sri Lanka	79	101	22
Average change over 146 countries				8

Table 3. ESI rankings obtained by linear aggregation (LIN) and non-linear/non-compensatory (NCMC) rules: countries that largely improve or worsen their rank position

4. Conclusion

This article starts from the assumption that, in some real-world applications, the linear aggregation rule is not the correct one for the building of relevant composite indicators. For this reasons a Condorcet consistent non-linear/non-compensatory mathematical aggregation rule for the construction of composite indicators aimed at ranking countries has been developed here. This mathematical aggregation rule can be divided into two main steps:

- pair-wise comparison of countries to be ranked,
- ranking of countries in a complete pre-order.

Weights are never combined with intensities of preference, as a consequence the theoretical guarantee they are importance coefficients is assured. Since intensities of preference are not used the degree of compensability connected with the aggregation model is at the minimum possible level. Given that the summation of weights is equal to one, the pair-wise comparisons can be synthesised in an outranking matrix, which can be interpreted as a voting matrix.

The information contained in the voting matrix is exploited to rank all alternatives in a complete pre-order by using a Condorcet consistent rule. A problem connected with the use of Condorcet consistent rules is the one of cycles. A cycle breaking rule normally needs some arbitrary choices such as to delete the cycle with the lowest support, and so on. A non-arbitrary cycle breaking rule is the so-called Kemeny median order which coincides with the maximum likelihood ranking proposed by Young and Levenglick.

Of course the approach we are proposing is not the “first best” possible approach for the construction of composite indicators. It is a “second best” approach based on theoretical and empirical grounds, which makes assumptions explicit and thus easier to be evaluated in relation with a particular use. This increases both the coherence and the transparency of the ranking obtained. One has to remember, however, that Arrow’s theorem implies that some useful and desirable properties in an aggregation convention are always lost. Here the properties lost are anonymity (which does not constitute a problem since weights are explicitly allowed and decisiveness is gained) and independence of irrelevant alternatives (which on the contrary is a serious loss since rank reversals may appear). Furthermore, information on intensity of preference of variables is never used. This loss of information is a necessary price to pay for guaranteeing that compensability is reduced and that weights can be considered as symmetrical importance coefficients. However, information on the intensity of preference can be partly used if indifference thresholds are used. Moreover, ranking procedures can be complemented with benchmarking exercises which are fully based on the use of intensity of preference (of course when variables are measured on an interval or ratio scale).

Finally, one has to note that the overall quality of any composite indicator depends on the following elements:

1. *information available* (often data to measure variables are unreliable or simply missing; thus *garbage in, garbage out* phenomena may easily occur (Funtowicz and Ravetz, 1990));
2. *individual indicators and variable chosen* (i.e. which representation of reality is used). This, to a large extent, determines *interpretability* of an index.
3. *direction of each indicator* (i.e. the bigger the better or vice versa, sometimes this decision is not that easy);

4. *relative importance of these indicators* (weights attached which are one of the main sources of technical uncertainty of the results provided; methodological discussion and sensitivity analysis are necessary to tackle this uncertainty by making it explicit);
5. *mathematical aggregation convention used* (the present paper has dealt with this step).

The quality of the aggregation convention is, yet, an important ingredient to guarantee the consistency between the assumptions used and the ranking obtained. Indeed, the overall quality of a composite indicator depends crucially on the way this mathematical model is embedded in the social, political and technical structuring process (Munda, 2004). This is especially true for the choice of weights that remains the most important source of uncertainty and debate. We have the firm conviction that weights are and must be *context-dependent* since they reflect political, social and economic priorities and depend on the development path a country (or a group of countries) wants to pursue. Often we have seen policymakers using the umbrella of “science” to disguise lobbies, individual interests, or incompetence. Again, whatever aggregations procedure is used, we think that transparency must remain one of the main ingredients: without *transparency* the *interpretability* and *coherence* of indicators are difficult to achieve.

ACKNOWLEDGEMENTS: This research has been partially developed in the framework of the European Commission financially supported project “*Knowledge Economy Indicators: Development of Innovative and Reliable Indicator Systems (KEI)*” (KEI SCS8 2004-502529).

References

- Arrow K.J.** (1963) - *Social choice and individual values*, 2d edition, Wiley, New York.
- Arrow K.J., Raynaud H.** (1986) - *Social choice and multicriterion decision making*, M.I.T. Press, Cambridge.
- Barthelemy J.P., Guenoche A., Hudry O.** (1989) – Median linear orders: heuristics and a branch and bound algorithm, *European Journal of Operational Research*, 42, pp. 313-325.
- Binner J. M., Fielding A. Mullineux A. W.** (1999) - Divisia money in a composite leading indicator of inflation, *Applied Economics*, Volume 31, Number 8 /, pp. 1021 – 1031.
- Binner J. M., Bissoondeal R. K., Mullineux A. W.** (2005) - A composite leading indicator of the inflation cycle for the Euro area, *Applied Economics*, Volume 37, Number 11 / 20, pp. 1257 – 1266.
- Black D.** (1958) – *The theory of committees and elections*, Cambridge University Press, Cambridge.
- Borda J.C. de** (1784) – Mémoire sur les élections au scrutin, in *Histoire de l' Académie Royale des Sciences*, Paris.
- Bouyssou D.** (1986) – Some remarks on the notion of compensation in MCDM, *European Journal of Operational Research*, 26 pp. 150-160.
- Bouyssou D., Vansnick J.C.** (1986) – Noncompensatory and generalized noncompensatory preference structures, *Theory and Decision*, 21, pp. 251-266.
- Charon I., Guenoche A., Hudry O., Woïrgard F.** (1997) – New results on the computation of median orders, *Discrete Mathematics*, 165/166, pp. 139-153.
- Cherchye L.** (2001) – Using data envelopment analysis to assess macroeconomic policy performance, *Applied Economics*, 33, pp. 407-416.
- Cribari-Neto F., Jensen M.J., Novo A.** (1999) - Research in econometric theory: quantitative and qualitative productivity rankings, *Econometric Theory*, Volume 15, Issue 05, pp 719-752.
- Cohen W., Schapire R., Singer Y.** (1999) – Learning to order things, *Journal of Artificial Intelligence Research*, 10, pp. 213-270.
- Condorcet, Marquis de** (1785) – *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la probabilité des voix*, De l' Imprimerie Royale, Paris.
- 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 Statistical Society*, 155(3), pp. 353-393.
- Davenport A., Kalagnanam J.** (2004) – *A computational study of the Kemeny rule for preference aggregation*, in McGuinness D.L., Ferguson G. – Proceedings of the Nineteenth National Conference on Artificial Intelligence, Sixteenth Conference on Innovative Applications of Artificial Intelligence, July 25-29, 2004, San Jose, California, AAAI Press/The MIT Press, USA.
- Diewert W.E.** (1976) - Exact and superlative index numbers, *Journal of Econometrics*, Volume 4, Issue 2 , Pages 115-145.
- Dwork C., Kumar R., Naor M., Sivakumar D.** (2001) – *Rank aggregation methods for the web*, in Proceedings 10th WWW, pp. 613-622.
- Ebert U., Welsch H.** (2004) – Meaningful environmental indices: a social choice approach, *Journal of Environmental Economics and Management*, 47, pp. 270-283.
- Esty, D.C., Levy M., Srebotnjak, T., de Sherbinin A.** (2005) – *Environmental Sustainability Index: Benchmarking National Environmental Stewardship*. New Haven, Conn.: Yale Center for Environmental Law and Policy.
- Färe R., Grosskopf S., Norris M., Zhang Z.** (1994) – Productivity growth, technical progress and efficiency change in industrialised countries, *American Economic Review*, 84(1), pp. 66-83.
- Fishburn P.C.** (1973) – *The theory of social choice*, Princeton University Press, Princeton.
- Fishburn P.C., Gehrlein W.V. and Maskin E.** (1979) – Condorcet's proportions and Kelly's conjecture, *Discrete Applied Mathematics*, 1, pp. 229-252.
- Forni M., Hallin M., Lippi M., Reichlin L.** (2001) Coincident and Leading Indicators for the Euro Area, *The Economic Journal*, Volume 111, Issue 471, Page 62.
- Funtowicz S.O., Ravetz J.R.** (1990) - *Uncertainty and quality in science for policy*, Kluwer Academic Publishers, Dordrecht.
- Granger W.J.** (2001) - Macroeconometrics – Past and future, *Journal of Econometrics*, Volume 100, Issue 1 , Pages 17-19.

- Griliches Z.** (1990) – Patent statistics as economic indicators, *Journal of Economic Literature*, 28, pp. 1661-1707.
- Grupp H., Mogee M.E.** (2004) - Indicators for national science and technology policy: how robust are composite indicators?, *Research Policy*, vol. 33(9), pages 1373-1384.
- Guerard J.B.** (2001) – A note on the forecasting effectiveness of the US leading economic indicators, *Indian Economic Review*, 36(1), pp. 251-268.
- Huggins R.** (2003) – Creating a UK competitive index: regional and local benchmarking, *Regional Studies*, 37, pp.89-96.
- Journal of Economic and Social Measurement** (2002) - *Special Issue on Index Number Aggregation*, Volume 28, Numbers 1-2.
- Kemeny J.** (1959) – Mathematics without numbers, *Daedalus*, 88, pp. 571-591.
- Kleinknecht A., Van Montfort K., Brouwer E.** (2002) – The non trivial choice between innovation indicators, *Economic Innovation and New Technologies*, 11(2), pp. 109-121.
- 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, pp. 507-518.
- Luce R.D.** (1956) – Semiorders and a theory of utility discrimination, *Econometrica*, 24, pp.178-191.
- Moulin H.** (1985) – From social welfare orderings to acyclic aggregation of preferences, *Mathematical Social Sciences*, 9, pp. 1-17.
- Moulin H.** (1988) - *Axioms of co-operative decision making*, Econometric Society Monographs, Cambridge University Press, Cambridge.
- Munda G.** (1995) - *Multicriteria evaluation in a fuzzy environment*, Physica-Verlag, Contributions to Economics Series, Heidelberg.
- Munda G.** (2004) – “Social multi-criteria evaluation (SMCE)”: methodological foundations and operational consequences, *European Journal of Operational Research* Vol. 158, Issue 3: 662-677.
- Munda G.** (2005) – Multi-Criteria Decision Analysis and Sustainable Development, in J. Figueira, S. Greco and M. Ehrgott (eds.) – *Multiple-criteria decision analysis. State of the art surveys*, Springer International Series in Operations Research and Management Science, New York, pp. 953-986.
- Munda G.** (2006) –A Solution to the Computational Problem of linear median orders, unpublished manuscript.
- Munda G. and Nardo M.** (2006) - Constructing Consistent Composite Indicators: the Issue of Weights, manuscript submitted to *Journal of Applied Statistics*.
- Nardo M., Saisana M., Saltelli A., Tarantola S., Hoffman A., Giovannini E.** (2005) – *Handbook on constructing composite indicators: methodology and user guide*, OECD Statistics Working Paper, Paris.
- Nicoletti G., Scarpetta S. and Boylaud O.** (2000) – Summary indicators of product market regulation with an extension to employment protection legislation, *OECD Economics Department Working papers* No. 226.
- Öcal N.** (2006) - Nonlinear models, composite longer leading indicator and forecasts for UK real GDP, *Applied Economics*, Volume 38, Number 9 / 20, pp. 1049 – 1053.
- OECD** (2003) – *Composite indicators of country performance: a critical assessment*, DST/IND(2003)5, Paris.
- Podinovskii V.V.** (1994) -Criteria importance theory, *Mathematical Social Sciences*, 27, pp. 237 - 252.
- Roubens M., Vincke Ph.** (1985) - *Preference modelling*, Springer-Verlag, Heidelberg.
- Roy B.** (1996) - *Multicriteria methodology for decision analysis*, Kluwer, Dordrecht.
- Saisana M. and Tarantola S.** (2002) – *State-of-the-art report on current methodologies and practices for composite indicator development*, EUR 20408 EN Report, European Commission, JRC, Ispra, Italy.
- Saltelli A. Tarantola S., Campolongo, F. and Ratto, M.** (2004) - *Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models*, John Wiley & Sons publishers, New York.
- Todhunter I.** (1949) – *A history of mathematical theory of probability*, Chelsea, New York.
- Truchon M.** (1995) – Voting games and acyclic collective choice rules, *Mathematical Social Sciences*, 25, pp. 165-179.
- Truchon M.** (1998a) – *Figure skating and the theory of social choice*, Cahier 9814 du Centre de Recherché en Economie et Finance Appliqués (CREFA).

Truchon M. (1998b) – *An extension of the Condorcet criterion and Kemeny orders*, Cahier 9813 du Centre de Recherche en Economie et Finance Appliquées (CREFA).

Vansnick J. C. (1990) - Measurement theory and decision aid- in Bana e Costa C.A. (ed.)- *Readings in multiple criteria decision aid*, Springer-Verlag, Berlin, pp. 81-100.

Vidu L. (2002) – Majority cycles in a multi-dimensional setting, *Economic Theory*, 20, pp. 373-386.

Vincke Ph. (1992) - *Multicriteria decision aid*, Wiley, New York.

Weber J. (2002) – How many voters are needed for paradoxes?, *Economic Theory*, 20, pp. 341-355.

Wilson J.W., and Jones C.P. (2002) – An analysis of the S&P 500 index and Cowles’s extensions: price indexes and stock returns, 1870-1999, *Journal of Business*, 75, pp. 505-533.

Young H.P. and Levenglick A. (1978) – A consistent extension of Condorcet’s election principle, *SIAM Journal of Applied Mathematics*, 35, pp. 285-300.

Young H.P. (1988) – Condorcet’s theory of voting, *American Political Science Review*, Vol. 82, No. 4, pp. 1231-1244.

Yu P. L. (1985) - *Multi criteria decision making: concepts, techniques and extensions*, Plenum Press, New York.

Zeleny M. (1982) - *Multiple criteria decision making*, McGraw Hill, New York.