

SUMMARY

FIRST WORKSHOP ON COMPOSITE INDICATORS OF COUNTRY PERFORMANCE

12 May 2003, Ispra, Italy

Background and overview

Composite indicators of country performance are synthetic indices of individual indicators which compare and rank countries in different performance areas, *e.g.* competitiveness, innovation, environmental sustainability. Composite indicators are valuable as communication tools and as inputs into policy-making. By reducing complex, multi-dimensional economic or social phenomena to a single index, composites can better focus and direct policy debates than can large sets of individual indicators. However, they confront many methodological difficulties, can be easily manipulated, and could lead to misleading information about countries and policies.

The Joint Research Centre (JRC) of the European Commission and the Organisation for Economic Co-operation and Development (OECD) held a workshop on 12 May 2003 to convene experts and practitioners to discuss the pros and cons of composite indicators of country performance. The following are the main conclusions of the Co-Chairs, Enrico Giovaninni (OECD) and Andrea Saltelli (JRC):

- The results of this workshop should be published and/or placed on the Internet. Presentations and background material are already available at: www.jrc.cec.eu.int/uasa/evt-oecd-jrc.asp
- Composite indicators are potentially useful tools for policy-making and can eventually be a standard statistical product. However, at present, they suffer from a dearth of theory, lack of operational rules, and little awareness of related risks and opportunities. There is as yet no basic statistical framework for practitioners nor readily available information for end-users.
- Yet composite indicators are increasingly being used by national and international organisations, such as the OECD, the United Nations and the EC, for cross-country comparisons.
- A number of guidelines, or good practices, were proposed in the presentations given at the workshop. These address theoretical, technical and procedural aspects needed for constructing composite indicators.
- An international network of experts should be created to develop quality guidelines for the construction of composite indicators. An electronic discussion group (EDG) could be formed for discussing the draft guidelines.
- Two additional workshops could be held: 1) to discuss the draft guidelines for composite indicators and 2) to discuss how policy lessons can be extracted from composites.

Quality guidelines for composite indicators

The following are examples of areas where guidelines or best practices are needed for the construction of composite indicators:

- *Theoretical framework* – A theoretical framework should be presented as providing the basis for the selection and combination of variables into a meaningful composite indicator. This analytical

underpinning will determine how sub-components and variables are weighted and should relate to a relevant policy process.

- *Data selection* – Variables should be selected on the basis of their analytical soundness, measurability, country coverage, relevance to the phenomenon being measured, and relationship to each other. Issues to be addressed include dealing with missing values, the reliability of “soft data” from surveys and other sources, problems of over-aggregation of data and double-counting of phenomena, whether to include both static values and growth rates, and difficulties in using countries as the unit of measure.
- *Correlation analysis of data* – A preliminary analysis of data could be done through the application of Principal Components Analysis and Cluster Analysis, with a view to gain insights into the relationships between the variables and an intuitive understanding of the phenomenon to be measured.
- *Standardisation methods* – Variables in a composite indicator should be standardised or normalised to render them comparable. Variables come in a variety of statistical units and sets with different ranges or scales which must be put on a common basis. The technique selected for standardisation – standard deviation, categorical scale, minimum-maximum, etc. – should be based on the theoretical framework and the data set in question.
- *Weighting approaches* – Variables in a composite indicator should be weighted according to an underlying theoretical framework or conceptual rationale. Greater weight should be given to components which are considered to be more significant in the context of the particular composite indicator. Weights may be assigned through expert opinion, techniques such as principal components analysis or factor analysis, or through correlations with dependent variables such as economic growth rates.
- *Country groupings* – Composite indicators which compare country performance should avoid comparing disparate countries, particularly in terms of development levels. Countries should first be divided into like groups or peer groupings so as to be compared or ranked within their relevant reference groups.
- *Sensitivity tests* – The robustness of composite indicators should be assessed in order to ensure their credibility and relevance to policy processes. Sensitivity tests should be conducted to assess the impact of including or excluding variables, changing weights, using different standardisation techniques and selecting alternative base years, etc. Composite indicators should be easily decomposed or disaggregated in order to conduct such tests.
- *Transparency/accessibility* – Composite indicators should be accompanied by detailed explanations of the underlying data sets, choice of standardisation techniques, selection of weighting methods, and assessment of robustness of alternative approaches. To the extent possible, the components of composites should be available electronically to allow users to change variables, weights, etc. and to replicate sensitivity tests.
- *Visualisation* – The presentation of the results of composite indicators should acknowledge their limitations, show the results of sensitivity tests, and include confidence intervals for country rankings. Composite indicators should be acknowledged as simplistic presentations and comparisons of country performance in given areas to be used as starting points for further analysis.

Workshop presentations

The following are highlights of the principal workshop presentations, for which Internet links are also provided:

Benchmarking European Country Performance (Stefano Tarantola, IPSC, EC Joint Research Centre) – The EC Communication on Structural Indicators (COM(2002)551) stresses the need for a clear analytical framework for each composite indicator and for more standardised proposals for developing composites. Among the steps to be followed in constructing composites are: 1) deciding on the phenomenon being measured; 2) selecting sub-indicators; 3) checking data availability; 4) investigating statistical properties of sub-indicators; 5) normalising and weighting variables; and 6) visualising the composite. As an example, the EC *Summary Innovation Index* ranks countries according to both static values and growth rates in order to see progress on performance variables. There was some discussion of the value of comparing only EU countries and whether a wider reference group of countries is needed to obtain more insights into performance.

Developing Composite Indicators of OECD Country Performance for the Growth Follow-Up Project (Michael Freudenberg, OECD) – The OECD is developing composite indicators of country performance in the areas of entrepreneurship, innovation, ICT use, and human capital as part of a project to develop policy benchmarks on key micro-drivers of growth. Composites have been developed based on the analytical framework of the Growth Project, using largely OECD data, standardised through the “minimum-maximum approach”, and weighted by giving equal value to sub-components and to all variables within sub-components. A series of sensitivity tests, including of different standardisation approaches and with random weights, found no large differences in ranking outcomes. There was discussion of the need to validate composite results against other variables such as GDP growth, whether all variables could be measuring like phenomena, and the relationship between composites and their underlying policy processes.

Constructing the Global Competitiveness Index (Maria Oliva, World Economic Forum, Geneva) – The World Economic Forum publishes a composite indicator comparing 63 countries on competitiveness performance, based on 1) innovation and technological capabilities, 2) quality of public institutions, and 3) efficiency of the financial sector. Countries are divided into like groups based on the degree to which their economies are innovation-driven as measured by patent statistics. Composite variables for core innovators and non-innovators are weighted differently. Variables include soft data from an Executive Opinion Survey of CEOs within given countries. A revised index will be developed in 2004 which combines methodologies developed by Michael Porter and Jeffrey Sachs. Discussion focused on the reliability of survey data, the need for more extensive sensitivity tests, and the comparability of the composite findings over different years.

Constructing the Environmental Sustainability Index (Daniel Esty, Yale Center for Environmental Law and Policy) – The ESI measures overall progress towards sustainability for 142 countries on the basis of 20 components (and 68 underlying data sets) divided into five groupings: environmental systems (air quality, water quantity), reducing environmental stresses, reducing human vulnerability, social and institutional capacity, and global stewardship. The 20 core components are weighted equally, although anyone can enter the ESI website and experiment with different weights. In the future, the ESI will be narrowed down to focus on environmental performance and will include only static values (rather than past values, present values and future projections). Discussion included the value of the ESI in capturing the attention of policy-makers and making environmental policy-making more data-based, the danger of over-aggregation of indicators, the problems of missing data, and the need to divide countries according to development levels for comparison.

Constructing the Human Development Index (Claes Johansson, United Nations Development Program) – The HDI is intended as a counterpoint to simple GDP measures in comparing 175 countries on GDP, health care (life expectancy, etc) and educational attainment (literacy rates, school enrolment). Proposed additions to HDI have proved either too hard to measure (e.g. political freedom) or to overlap with existing components. The HDI faces severe data constraints for many countries and thus includes survey results. The three core components, which tend to be highly correlated, are weighted equally. The HDI is supplemented by the *Human Poverty Index* (HPI), the *Gender Development Index* (GDI), and the *Technology Achievement Index* (TAI). There was discussion of how certain measures were affected by country characteristics (e.g. electricity use as influenced by climate) and the relationship of HDI to the 42 indicators included in the *Millennium Development Goals* (MDG).

Constructing the European Innovation Index (Anthony Arundel, MERIT, the Netherlands) – The EII, which compares the innovative capacity of countries in the European Union, is based on traditional R&D and patent statistics. Sensitivity tests show that different approaches to constructing and weighting the EII produced identical results, indicating that all variables tended to measure R&D-based innovation. An alternative composite (which could predict innovative capabilities and growth performance) might be based on concepts of innovation as R&D, as technology diffusion and as buying in new technology, but resulted in the use of the same data set. In 2003, the EII will include indicators of structural conditions which affect innovation, different modes of innovation, and country peer group analysis. Discussion focused on limitations to improving the theoretical foundations of composites due to data gaps, how the use of more indicators may not necessarily lead to better results, and the problems in linking composites to growth performance (due to failure to capture all aspects of country performance and short timelines).

Operating the Dashboard of Sustainability (Jochen Jesinghaus, DG Development, European Commission) – The Dashboard (which can be downloaded at <http://esi.jrc.it/dc>) is an interactive computer tool which can analyse and compare country performance on four dimensions: social, environmental, economic and institutional – based on a set of 58 indicators. It is useful in tracking the progress of countries over time from less sustainable performance (red zone) to more sustainable (green zone). Composite indicators in different performance areas can be compared to establish linkages between behaviour patterns. Different indicator sets can also be correlated across countries and through time to identify links between policy actions and performance outcomes (e.g. level of skilled health staff with child mortality rates). Discussion focused on how to interpret spurious correlations, e.g. the positive correlation between CO2 emissions (an indicator of wealth) and life expectancy.

Mathematical Modelling of Composite Indicators and the Issue of Quality (Giuseppe Munda, Universitat Autònoma de Barcelona, Spain; Silvio Funtowicz, Knowledge Assessment Methodologies Group, JRC). This presentation focused first on the axiomatic system behind the mathematical modelling of standard composite indicators, finding that: 1) weights in linear aggregation rules imply trade-off ratios, and given that weights are used as important coefficients in composite indicators, a theoretical inconsistency exists; and 2) the assumption of preference independence is essential for the existence of a linear aggregation rule, but this is often not desirable in a composite indicator. Second, a new mathematical aggregation convention has been developed for composite indicators aimed at ranking countries. The main features of the proposed approach are: 1) the axiomatic system is made completely explicit, and 2) the sources of technical uncertainty and imprecise assessment are reduced to a minimum. Third, the overall quality of a composite indicator depends on the way this mathematical model is embedded in the social, political and technical structuring process.

Endogenous Weighting for Composite Performance Indicators (Tom Van Puyenbroeck, Europese Hogeschool, Brussels) – Weighting of variables is a continuing problem in composite indicators and one which calls into question the credibility of their results. In many composites, all variables are given common weights for reasons of simplicity. Greater weight can be given to variables assumed to have

greater significance in the composite according to a theoretical framework or according to expert opinion, or to compensate for missing data. In contrast to these *exogenous* weighting approaches, there is an *endogenous* approach where countries can be allowed to select their own weights for variables. This can promote greater political acceptance of composites by allowing countries to discount variables on which they are weak while showing their revealed preferences. There was mainly discussion of how accurate the endogenously-weighted composite would be in comparing country performance.

Explorative Data Analysis of Structural Indicators: Preliminary Results (Alice Zoppè, Eurostat)
Correlation analysis, principal components analysis and cluster analysis have been applied to the Eurostat Structural Indicators set, so that “the ... selected structural indicators are not seen in isolation but as different elements of the same picture...”. Currently the Structural Indicator dataset contains more than 30,000 figures and covers a timeframe from 1948 to 2003. Principal components analysis and cluster analysis have been used for graphical representations of individuals (country/year) and for identifying patterns and groups.