Unclassified **DSTI/DOC(2003)16**

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

12-Nov-2003

English text only

DIRECTORATE FOR SCIENCE, TECHNOLOGY AND INDUSTRY

DSTI/DOC(2003)16
Unclassified

COMPOSITE INDICATORS OF COUNTRY PERFORMANCE: A CRITICAL ASSESSMENT

STI WORKING PAPER 2003/16
INDUSTRY ISSUES

Michael Freudenberg

English text only

JT00153477

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

STI Working Paper Series

The Working Paper series of the OECD Directorate for Science, Technology and Industry is designed to make available to a wider readership selected studies prepared by staff in the Directorate or by outside consultants working on OECD projects. The papers included in the series cover a broad range of issues, of both a technical and policy-analytical nature, in the areas of work of the DSTI. The Working Papers are generally available only in their original language – English or French – with a summary in the other.

Comment on the papers is invited, and should be sent to the Directorate for Science, Technology and Industry, OECD, 2 rue André Pascal, 75775 Paris Cedex 16, France.

The opinions expressed in these papers are the sole responsibility of the author(s) and do not necessarily reflect those of the OECD or of the governments of its member countries.

http://www.oecd.org/sti/working-papers

COMPOSITE INDICATORS OF COUNTRY PERFORMANCE: A CRITICAL ASSESSMENT

Michael Freudenberg

Abstract

Composite indicators are synthetic indices of individual indicators and are increasingly being used to rank countries in various performance and policy areas. Using composites, countries have been compared with regard to their competitiveness, innovative abilities, degree of globalisation and environmental sustainability. Composite indicators are useful in their ability to integrate large amounts of information into easily understood formats and are valued as a communication and political tool. However, the construction of composites suffers from many methodological difficulties, with the result that they can be misleading and easily manipulated. This paper reviews the steps in constructing composite indicators and their inherent weaknesses. A detailed statistical example is given in a case study. The paper also offers suggestions on how to improve the transparency and use of composite indicators for analytical and policy purposes.

INDICATEURS COMPOSITES DE PERFORMANCES DES PAYS : EXAMEN CRITIQUE

Michael Freudenberg

Résumé

Les indicateurs composites, qui résultent de la fusion de plusieurs indicateurs, sont de plus en plus utilisés pour classer les pays dans des performances et des domaines d'action divers. A l'aide de ces indicateurs, des pays ont été comparés dans les domaines de la compétitivité, la capacité d'innovation, le niveau de mondialisation et la viabilité écologique. Les indicateurs composites sont appréciés en raison de leur capacité à intégrer de grandes quantités d'informations sous une forme aisément compréhensible, ce qui fait d'eux des outils de communication et des instruments d'action précieux. L'élaboration des indicateurs composites soulève néanmoins de nombreuses difficultés méthodologiques ; ils peuvent de ce fait être sources d'erreurs et sont aisés à manipuler. Le présent document décrit les différentes étapes d'élaboration des indicateurs composites et les lacunes qui leur sont inhérentes. L'annexe fournit des tableaux statistiques détaillés. Le document propose également des suggestions pour améliorer la transparence et l'utilisation des indicateurs composites à des fins d'analyse et d'orientation.

TABLE OF CONTENTS

INTRODUCTION

Composite indicators – which are synthetic indices of individual indicators – are being developed in a variety of economic performance and policy areas (**Table 1**). There are different types of composite indicators. The *OECD Composite of Leading Indicators*, which groups economic statistics by country in order to track and forecast business cycles, is one type which can be subject to empirical tests to see how well its forecasting ability matches reality. This paper, however, is concerned with composite indicators which compare and rank countries in areas such as industrial competitiveness, sustainable development, globalisation and innovation and which cannot be subject to an empirical test. 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. This paper reviews methods for constructing composite indicators and their inherent weaknesses. It is intended as a methodological companion piece to the construction of composite indicators in the OECD. The problems encountered in constructing an index of innovation performance are used for illustration.

There are a number of steps to be followed in constructing composite indicators, which are discussed in this paper:

- Developing a theoretical framework for the composite.
- Identifying and developing relevant variables.
- Standardising variables to allow comparisons.
- Weighting variables and groups of variables.
- Conducting sensitivity tests on the robustness of aggregated variables.

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. These summary indicators are valuable in that they limit the number of statistics to be presented and allow for quick comparisons of country performance. At a minimum, all composite indicators should be as transparent as possible and provide detailed information on methodology and data sources. They should always be accompanied by explanations of their components, construction, weaknesses and interpretation. Sensitivity tests should be carried out on standardisation, weighting and aggregation approaches. In general, composite indicators should be identified for what they are -- simplistic presentations and comparisons of country performance in given areas to be used as starting points for further analysis.

Table 1. Examples of composite indicators

Area / Name of Composite Indicator
Economy
Composite of Leading Indicators (OECD)
OECD International Regulation Database (OECD)
Economic Freedom of the World Index (Economic Freedom Network)
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)
Corruption Perceptions Index (Transparency International)
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)

Source: JRC (2002) and compilation by OECD.

CONSTRUCTION OF COMPOSITE INDICATORS

Developing a theoretical framework

Composite indicators are generally used to summarise a number of underlying individual indicators or variables. An indicator is a quantitative or qualitative measure derived from a series of observed facts that can reveal relative position in a given area and, when measured over time, can point out the direction of change. In the context of policy analysis at national and international levels, indicators are useful in identifying trends in performance and policies and drawing attention to particular issues. There are basically three levels of indicator groupings:

1) *Individual indicator sets* represent a menu of separate indicators or statistics. This can be seen as a first step in stockpiling existing quantitative information.

2) *Thematic indicators* are individual indicators which are grouped together around a specific area or theme. This approach requires identifying a core set of indicators that are linked or related in some way. They are generally presented individually rather than synthesised in a composite (*e.g.* OECD's *Measuring the Information Economy*).

3) *Composite indicators* are formed when thematic indicators are compiled into a synthetic index and presented as a single composite measure.

In the case of comparing the performance of countries on different dimensions, a typical composite indicator will take the form:

$$I = \sum_{i=1}^{n} w_i X_i ,$$ where:

I: Composite index,

X_i : Normalised variable,

w_i : Weight of the X_i , $\sum_{i=1}^{n} w_i = 1,$ and $0 \le w_i \le 1$

i: 1,…, n.

In practice, it is extremely difficult to integrate individual variables in a manner which accurately reflects economic reality. As a starting point, one needs an understanding and a definition of what it is that is being measured. A theoretical framework is needed to combine individual indicators into a meaningful composite and to provide a basis for the selection of components and weights in the formula above. Ideally, this framework will allow variables to be selected, combined and weighted in a manner which reflects the dimensions or structure of the phenomena being measured. The variables selected should carry relevant information about the core components and be based on a paradigm concerning the behaviour being analysed. It is this framework which indicates which variables to include and how to weight them to reflect their relative importance in the overall composite. But as yet, the theoretical underpinning of most composite indicators is very underdeveloped.

Composite indicators are increasingly being used to make cross-national comparisons of country performance in specified areas such as competitiveness, globalisation, innovation, etc. Rather than using a disaggregated menu of individual indicators, aggregated composites supposedly allow for analysis of interrelated performance or policy areas. They are popular in benchmarking exercises where countries wish to measure their performance relative to other countries and identify general areas where national performance is below expectations. Benchmarking with the aid of composites is often used to identify general trends, determine performance targets and set policy priorities.

Ideally, composites allow for investigation among different areas and provide for the understanding of economic performance beyond simple one-dimensional measures. For example, the levels of and changes in gross domestic product (GDP) per capita are a well-established measure of economic development and growth. However, as a stand-alone indicator, it may not be sufficient to measure cross-country disparities in development, which encompasses a variety of areas from health care to educational attainment to the provision of other basic needs. Taking gross output as a proxy measure can be misleading. The *Human Development Index* (HDI) developed by the United Nations Development Programme (UNDP) is one attempt to overcome such shortcomings in measuring economic development. The HDI integrates a range of indicators pertaining to life expectancy and educational attainment as well as GDP. However, the HDI, like most composite indicators, has been criticised for inconsistencies and methodological flaws.

In general, composite indicators by their nature may be incapable of reflecting the complexity of performance and policies or of capturing the intricate relationships between variables. A simple composite indicator, formulated as an average of individual indicators, implicitly assumes the substitutability of its components. For example, composite environmental indicators imply that clean air can compensate for water quality. In fact, the multidimensional nature of most performance areas argues for a set of individual indicators and against composites. The more comprehensive a composite, the weaker it may be in adequately reflecting actual country performance.

Selecting variables

A composite indicator is, above all, a sum of its parts. The strengths and weaknesses of a composite derive largely from the quality of the underlying variables. Ideally, variables should be selected on the basis of their analytical soundness, measurability, relevance to the phenomenon being measured, and relationship to each other. But by their nature, composite indicators can mask data problems rather than present statistical issues transparently. While poor quality indicators can only yield a weak composite, high quality indicators can yield either a weak or strong composite depending on other factors.

The greatest problem in constructing a composite indicator is the lack of relevant data. Statistics may be unavailable because a certain behaviour cannot be measured or no one has attempted to measure it. The data available may not be comparable across countries or exist only for a few countries. The indicators may be unreliable measures of the behaviour or not match the analytical concepts in question. The construction of composite indicators of country performance generally involves trade-offs between broad country coverage and lower quality data. Due to the expense and time involved in developing internationally comparable performance indicators, composites often rely on data sources of less than desirable quality. In the end, they may measure only the most obvious and easily accessible aspects of performance.

Because there is no single definitive set of indicators for any given purpose, the selection of data to incorporate in a composite can be quite subjective. Different indicators of varying quality could be chosen to monitor progress in the same performance or policy area. Due to a scarcity of full sets of comparable quantitative data, qualitative data from surveys or policy reviews are often used in composite indicators. The tendency to include "soft" qualitative data is another source of unreliability with regard to composites.

Measuring country performance rests on using the nation-state as the unit of measure, which has inherent limitations. A comparison of a small country and a large country, treated as equals in a composite indicator, can be misleading considering their relative size. Variables within a composite can misrepresent performance if they are related to GDP or other size-related factors. Small countries face different problems than large countries, most notably in scale economies and gaining critical mass in activities and production. Many indicators of performance (*e.g.* innovation, ICT-readiness) may be simple reflections of the level of development of a country or per capita income and highlight problems of causality. In addition, country-level composite indicators generally fail to capture the wide internal differences within large countries and do not pick up variations in regional performance variables.

Like most statistical series, composite indicators are plagued by problems of missing values. In many cases, data are only available for a limited number of countries or only for certain data components. Missing values can render the composite indicator less reliable for the countries for which only limited information is available and can distort the relative standing of all countries in the composite. There are a number of approaches for dealing with missing values, all of which have flaws:

- *Data deletion* -- omitting entire records (for variables or countries) when there is substantial missing data.
- *Mean substitution* -- substituting a variable's mean value computed from available cases to fill in missing values.
- *Regression* -- using regressions based on other variables to estimate the missing values.
- *Multiple imputation* -- using a large number of sequential regressions with indeterminate outcomes, which are run multiple times and averaged.
- *Nearest neighbour* -- identifying and substituting the most similar case for the one with a missing value.
- *Ignore them* -- take the average index of the remaining indicators within the component.

Changes in composite indicators over time are generally hard to interpret, which limits their value as a tool for identifying the determinants of country performance over time. One difficulty is obtaining data for points in time which are synchronised with measurements in other countries (*e.g.* selection of base years, mixing of years across indicators), which compounds the above-mentioned problems of missing values. Especially when the methodology and underlying data are not made public, it is virtually impossible for a reader to distinguish between the “real” performance (improvements or deteriorations in some or all areas) and the method and data coverage: differences in country rankings over time may be the result of data improvements, different weighting approaches or other changes to the composite indicator make-up or methodology rather than to any change in country performance. This is why composite indicators generally do not use time series data.

In general, the quality and accuracy of composite indicators should evolve in parallel with improvements in data collection and indicator development. From a statistical point of view, the construction of composite indicators can help identify priority indicators for development and weaknesses in existing data. The current trend towards constructing composite indicators of country performance in a range of policy fields may provide an impetus to improving data collection, identifying new data sources and enhancing the international comparability of statistics.

Standardising variables

Variables need to be standardised or normalised before they are aggregated into composite indicators. Variables come in a variety of statistical units and different variable sets have different ranges or scales. Variables need to be put on a common basis to avoid problems in mixing measurement units (*e.g.* firms, people, money). They must be adjusted on dimensions such as size/population/income and smoothed through time against cyclical variability. Variables are normalised to avoid having extreme values dominate and also to partially correct for data quality problems. There is reason to believe that values extremely far from the average or normal range are more likely to reflect poor underlying data. If certain variables have highly-skewed distributions, they can be levelled through logarithmic transformations and the data can be truncated if there are extreme outliers.

Several techniques can be used to standardise or normalise variables. Commonly used methods include the following (of which the first four are illustrated in **Table 2** using business enterprise R&D (BERD) as a percentage of GDP):

1. ***Standard deviation from the mean***, which imposes a standard normal distribution (*i.e.* a mean of 0 and a standard deviation of 1). Thus, positive (negative) values for a given country indicate above (below)-average performance: $\left(\frac{\text{actual value - mean value}}{\text{standard deviation}}\right)$

2. ***Distance from the group leader***, which assigns *100* to the leading country and other countries are ranked as percentage points away from the leader: $100\left(\frac{\text{actual value}}{\text{maximum value}}\right)$

3. ***Distance from the mean***, where the (weighted or unweighted) mean value is given 100, and countries receive scores depending on their distance from the mean. Values higher than 100 indicate above-average performance: $100\left(\frac{\text{actual value}}{\text{mean value}}\right)$

4. ***Distance from the best and worst performers***, where positioning is in relation to the global maximum and minimum and the index takes values between 0 (laggard) and 100 (leader): $100\left(\frac{\text{actual value - minimum value}}{\text{maximum value - minimum value}}\right)$

5. ***Categorical scale****,* where each variable is assigned a score (either numerical such as between [1…k], k>1, or qualitative -- high, medium, low) depending on whether its value is above or below a given threshold.

Table 2. An example of four different normalisation methods

	Initial data	Normalisation methods			
	Business R&D expenditures as % of GDP	Standard deviation from the mean	Distance from the leader (Leader=100)	Distance from the best and worst performers (Leader=100, laggard=0)	Distance from the mean (Average=100)
Sweden	2.81	2.59	100.0	100.0	248.3
Japan	2.04	1.48	72.8	71.9	180.8
Switzerland	1.93	1.32	68.8	67.8	170.8
United States	1.93	1.31	68.6	67.6	170.4
Finland	1.90	1.27	67.7	66.7	168.2
Korea	1.85	1.20	66.0	64.8	163.7
Germany	1.58	0.80	56.1	54.7	139.4
France	1.38	0.52	49.3	47.6	122.4
Belgium	1.35	0.48	48.2	46.5	119.7
Denmark	1.24	0.31	44.1	42.3	109.5
United Kingdom	1.23	0.30	43.9	42.0	108.9
Austria	1.14	0.17	40.6	38.7	100.9
Netherlands	1.10	0.10	39.0	37.0	96.9
Canada	1.03	0.01	36.7	34.6	91.2
Ireland	1.01	-0.02	36.0	33.9	89.4
Norway	0.95	-0.12	33.7	31.5	83.6
Iceland	0.86	-0.24	30.7	28.4	76.1
Czech Republic	0.74	-0.42	26.2	23.8	65.1
Australia	0.72	-0.45	25.6	23.1	63.5
Slovak Republic	0.59	-0.64	21.0	18.4	52.2
Italy	0.53	-0.72	18.9	16.2	46.9
Spain	0.43	-0.87	15.4	12.6	38.3
New Zealand	0.32	-1.04	11.2	8.3	27.9
Poland	0.30	-1.07	10.5	7.6	26.1
Hungary	0.28	-1.09	10.0	7.0	24.8
Turkey	0.17	-1.25	6.1	2.9	15.0
Portugal	0.16	-1.27	5.5	2.4	13.7
Greece	0.15	-1.28	5.2	2.1	13.0
Mexico	0.09	-1.37	3.2	0.0	8.0
Unweighted average	1.03				
Standard deviation	0.69				

Note: No data for Luxembourg.

Each method has its advantages and disadvantages. The standard deviation approach is most commonly used because it has desirable characteristics when it comes to aggregation. It converts all variables to a common scale and assumes a "normal" distribution; it has an average of zero, meaning that it avoids introducing aggregation distortions stemming from differences in variable means. In the other approaches, the scaling factor is the range of the distribution, rather than the standard deviation, which means that extreme values can have a large effect on the composite index. For example, the "distance from the best and worst performers" is based on the extreme values which may be unreliable outliers. Distance from the median (rather than from the mean) may be less subject to distortion from outliers or variables that experience a large variance. Categorical scales have a high degree of subjectivity as the scale and the thresholds are by and large determined arbitrarily. Such ranking approaches also omit a great deal of information on the amount of variance between countries.

Weighting variables

Variables which are aggregated in a composite indicator have first to be weighted -- all variables may be given equal weights or they may be given differing weights which reflect the significance, reliability or other characteristics of the underlying data. The weights given to different variables heavily influence the outcomes of the composite indicator. The rank of a country on a given scale can easily change with alternative weighting systems. For this reason, weights ideally should be selected according to an underlying theoretical framework or conceptual rationale for the composite indicator. A stated methodology should be used for determining weights and should be explained transparently. It is also useful to illustrate how general findings and country rankings change with the use of alternative weighting systems.

In many composite indicators, all variables are given common weights largely for reasons of simplicity. Theoretical frameworks for deriving coherent weighting approaches are difficult to construct. This implies, however, that all indicators in the composite have equal importance, which may not be the case. Another approach is to give equal weights to all sub-indices or sub-components (which may comprise varying numbers of indicators), which again implies that each grouping of indicators has the same impact on the performance being measured.

With the equal weighting approach, there is the risk that certain performance aspects will be double weighted. This is because two or more indicators may be measuring the same behaviour. For example, in composites of ICT readiness, indicators relating to Internet access, Internet websites and Internet use overlap and, when used together, will tend to heavily weight one aspect of the composite. As a remedy, indicators could be tested for statistical correlations, and lower weights could be given to variables strongly correlated with each other. On the other hand, correlations may merely show that country performance on these indicators is similar.

Greater weight should be given to components which are considered to be more significant in the context of the particular composite indicator. The relative economic impact of variables could be determined by economic theory or through empirical analysis, particularly by methods based on correlations among the sub-indicators. These include regression analysis, principal components analysis and factor analysis. Such techniques can be used to examine the interrelationship among the base indicators, but it is not certain that the correlations will correspond to the real-world links between the phenomena being measured (JRC, 2002). Weights can also be set based on correlation coefficients between indicators and a dependent variable such as economic growth: the higher the correlation, the more important the indicator could be in the composite and weighted accordingly.

Indicators could also be weighted in co-operation with experts who understand the strengths, weaknesses and nuances of the data within a given theoretical context. In the United Kingdom, international e-commerce benchmarking is based on the judgement of independent experts on the influence of the factors being measured on e-commerce adoption and impacts (ONS, 2002). Another example of a differential weighting system is the World Markets Research Centre Globalisation Index which assigns different weights to "*new economy*" and "*old economy*" factors according to various conceptions of the functioning of the economy (WMRC, 2001). However, opinion-based weighting approaches can also be quite subjective and introduce new distortions to a composite indicator.

Another approach is to give less weight to variables that suffer most from missing values in the attempt to partially correct for data problems. The reliability of a composite indicator can be improved by giving more weight to the components with higher quality and availability. However, this can skew the results of composite rankings towards simpler, more readily available or identifiable factors. In the case of country rankings, more developed nations with more comprehensive data collection and statistics may show up better than those with less developed and less complete indicator sets.

Testing for robustness

As seen above, a variety of difficulties can arise with regard to selecting, weighting, standardising and aggregating variables into a composite indicator. Outcomes and country rankings may depend largely on the approach selected. For this reason, sensitivity tests should be conducted to analyse the impact of including or excluding various variables, changing weights, using different standardisation techniques and selecting alternative base years, etc., on the results of the composite indicator.

Various statistical tests can help ensure that the composite is robust and not heavily dependent on the choice of standardisation or weighting approaches or on the levels of aggregation of sub-components. The bivariate correlation of selected variables -- within both a sub-component and in the overall dataset -- can be used to check whether a given set of variables is empirically-related and to examine the grouping of individual variables. This will show whether selected variables are (positively) correlated within and across the core components, implying that they could be aggregated in different configurations without introducing bias to the composite indicator. If variables are strongly correlated (both with indicators from the same component and with those from other components), this renders the overall results less sensitive to the problem of missing values and the weights attached to the indicators. On the other hand, it may also imply that the indicators are double-counting the same behaviour.

Factor or cluster analysis can also be used to test the robustness of a composite indicator. These approaches can facilitate comparisons across countries with like profiles and help identify similarities among countries across multiple heterogeneous dimensions. Through factor analysis, interrelationships or common threads among variables are identified and coefficients or "factor loadings" are calculated that link variables and are used for weighting or standardising variables in the composite. During the past few years, factor analysis has gained more respect among researchers. This is partly due to the development of maximum-likelihood estimation, which facilitates verification of the number of factors or common trends driving the original variables. Before maximum-likelihood estimation was used in the area of factor analysis, the choice of the number of factors was often very arbitrary.

Standardisation approaches can also be tested and compared for their usefulness or accuracy. There are questions as to whether indicators should be aggregated using an arithmetic or geometric mean, or other approach. The correlation of different standardisation techniques (*e.g.* standard deviation, distance from the leader, distance from the mean) can show whether the results of the composite indicator are heavily influenced by the choice of approach. For example, when the rank correlation coefficient between the standard deviation method and other approaches is close to one, this would imply that the rankings of the majority of countries remain unchanged when different standardisation methods are applied.

The results or rankings from a composite indicator can be compared with other measures to test findings against economic theory or empirical evidence. However, simple correlations cannot provide information on cause and effect. A statistically significant correlation between the composite and another indicator can be interpreted as either: *i*) there may be a cause and effect relationship, although the direction is not known without other information, *ii*) at a minimum, the two factors do not interfere with each other, or *iii*) it may just be a spurious relationship, *i.e.* both indicators are driven by a third one. For example, a strong correlation between innovation performance and GDP per person employed would not indicate that a high innovative capability increases productivity. Rather, it suggests that there might be a relationship between the two, that they are both correlated with a third variable which is missing, or that the factors increasing the composite do not interfere with improved productivity.

CASE STUDY: DEVELOPING AN INDEX OF INNOVATION PERFORMANCE

The pitfalls of constructing composite indicators can be illustrated through attempts to formulate an index of performance relating to innovation across OECD countries. The intent is to measure innovation performance in a set of countries, which will then be the subject of more in-depth studies with regard to innovation policies. The starting point is the analysis and recommendations of the OECD Growth Project, which identified innovation and technology diffusion as one of the key micro-drivers of productivity and growth in OECD countries in the 1990s (OECD, 2001).

Theoretical framework

Innovation can be defined as the development, deployment and economic utilisation of new products, processes and services, and is an increasingly important contributor to sustained and sustainable economic growth, both at micro-economic and macro-economic levels. It enables firms to respond to more sophisticated consumer demand and stay ahead of their competitors, both domestically and internationally, and contributes to the growth of multifactor productivity. Beyond its contribution to economic growth and efficiency, innovation facilitates the fulfilment of other societal needs, such as improved health and environmental protection.

In constructing an index of innovation performance, the policy recommendations of the OECD Growth Project (OECD, 2001) were used as a framework for selecting and placing indicators in three performance areas deemed most relevant for innovation in OECD countries: *i*) the conduct of basic research and production of new knowledge, *ii*) the existence of links between public and private research, and *iii*) high levels of industrial innovation. The three areas parallel the main policy recommendations of the OECD Growth Project regarding innovation: *i*) ensuring generation of new knowledge and making government funding more effective; *ii*) fostering science-industry links and enhancing knowledge diffusion; and *iii*) creating incentives for private sector innovation.

This framework, however, gives a somewhat narrow interpretation of innovation performance. In its overall conception, the framework is a linear presentation of the innovation process where knowledge proceeds from government to industry through defined channels and then is transformed into technology and innovative outputs. In reality, new knowledge can arise in any sector, and innovation results from cumulative interactions among economic actors in the public and private sectors. The innovation process is also international, with countries adopting and adapting knowledge and technology from abroad rather than innovating within the confines of national borders. However, the multi-faceted innovation process is difficult to capture in available indicators.

The framework is also limited in indicating how to measure the innovation performance of a country. It points to countries with high levels of innovative inputs and/or outputs as the best performers. However, indicators do not necessarily reflect the quality or efficiency of countries, *e.g.* high R&D intensity does not means that R&D inputs are efficiently used. One could seek to identify countries that are the most efficient innovators, *i.e.* those that have high levels of innovation and/or economic payoffs once inputs and structural factors are taken into account. Alternatively, one could look at the dynamics of innovation and identify those countries that have made significant improvements in innovative capabilities over time. For conceptual and statistical reasons, the more simplistic Growth framework is used here to develop a composite of innovation performance.

Selecting variables

On this basis, the index consists of three core components that combine between three to five underlying variables, mostly derived from OECD databases (**Table 3**). The first core component (*generation of new knowledge*) aggregates variables such as basic research as a percentage of GDP and non-business researchers as a share of the labour force. The second core component (*industry/science linkages*) looks at public/private links through data relating to R&D, the scientific content of patents, and publications. In the third component (*industrial innovation*), data on business research, patents and the introduction of new products and processes are used to measure private sector innovative performance.

These variables were selected because internationally comparable data were available for a large number of OECD countries. However, there are many other relevant indicators of innovation performance which were not selected. They were either too imperfect, not obtainable for all OECD countries or not internationally comparable. In addition, certain aspects of the innovation process cannot be represented through such metrics. As a result, the index compares countries only in areas where suitable data exist and omits more intangible aspects of innovation performance as well as areas where the data are incomplete. This is an important limitation to the usefulness of the Index.

Some important areas were purposely omitted. Because the selection of indicators is intended to reflect domestic performance with regard to innovation and technology diffusion, indicators of international flows of technology, patents and licences, etc. are absent. If they had been included, countries such as Ireland would rank more highly on the innovation composite due to their large imports of technology. Other indicators were rejected because of limited coverage. For example, a relevant indicator for industry-science linkages would be survey data on the "*share of firms with co-operation arrangements with government or higher education institutes*", but such information from the Community Innovation Survey (CIS) (managed by Eurostat) is available for only 16 countries and not always comparable. Similarly, data for broad measures of "*expenditures on innovation*" (*e.g.* expenditure related to the scientific, technological, commercial, financial and organisational steps that lead to the implementation of technologically new or improved products and processes) are available only for a limited number of (mostly European) countries.

In some cases, it is not clear whether a given indicator is a measure of input, output or some other feature of the innovation process. Judgements need to be made which influence both the selection and weighting of indicators. For example, *PhD graduation rates in science, engineering and health* can be seen as an output of the education system, but are here used as an input for the innovation system. The *share of firms introducing new or improved products or processes* is a true output indicator, but a preferred measure would be the share of such products or processes in total turnover. The limited number of true output indicators is one reason why input indicators (*e.g.* business enterprise R&D as a percentage of GDP) are included in the Index. However, the large number of input indicators raises the question of potential biases in the measures.

In addition, the variables selected have inherent weaknesses. The *ratio of science, engineering and health PhDs per population* uses measures of educational attainment as a proxy for skills and is limited to the age group 25 to 34 years. Similarly, the *number of scientific and technical articles per million population* is used as a proxy for scientific output.

Indicators can be strongly influenced by a few industries or large enterprises, and such structural differences across countries may contribute to differences in overall innovation performance. The extent of industrial innovation and science/industry links will depend partly on the level of industrialisation and the sectoral specialisation of countries. For example, with regard to the number of *non-business researchers* and *business researchers,* OECD countries differ widely concerning the share of business researchers in the national total for structural reasons not wholly related to their innovation performance. The indicator of *share of firms having introduced at least one new or improved product or process on the market* is weighted by number of employees in order not to underestimate the weight of large firms.

Table 3. Indicators used to measure innovation performance

Category / indicators (weights)	Number of countries with available data
Generation of new knowledge (33 1/3%)	
R&D performed by the non-business sector as a percentage of GDP	29
Non-business researchers per 10 000 labour force	29
Basic research as a percentage of GDP	21
PhD graduation rates in science, engineering and health	24
Scientific and technical articles per million population	28
Industry-science linkages / technology diffusion (33 1/3%)	
Business-financed R&D performed by government or higher education as a percentage of GDP	28
Scientific papers cited in US-issued patents	23
Publications in the 19 most industry-relevant scientific disciplines per million population	23
Industrial innovation (33 1/3%)	
BERD as a percentage of GDP	29
Business researchers per 10 000 labour force	29
Patents in triadic patent families per million population	30
Share of firms introducing new or technologically improved products or processes on the market	21

The **generation of new knowledge** is an important element for future innovative capacity and is strongly influenced by research and development (R&D) activities and the supply of skilled human resources. According to the OECD *Frascati Manual* definition, R&D comprises "*creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications*". Five indicators have been used to compare performance across OECD countries with the following results (see **Table 4** which presents both the initial data and the “minimum-maximum normalisation method”):

- *R&D performed by the non-business sector as a percentage of GDP* is a proxy for a country’s relative efforts to create new knowledge, though it should be noted that new knowledge can also originate in firms or in partnership with firms. Public R&D activities also disseminate new knowledge and exploit existing knowledge bases in the public sector. The ratio for 1995-99 is highest in Iceland, the Netherlands, Finland and Sweden.

- The number of *non-business researchers per 10 000 labour force* in 1999 is highest in Iceland, Finland and Australia, followed by Sweden, Norway, New Zealand, Denmark, Japan and France. Overall, approximately 3.2 million researchers are engaged in R&D activities in the OECD area. However, OECD countries differ widely concerning the share of business researchers in the national total. For example, four out of five researchers in the United States work in the business sector, whereas this is the case for only one out of two researchers in the European Union.

- *Expenditures on basic research as a percentage of GDP* during the 1990s were highest in Switzerland, Sweden, France, Germany, Denmark, the United States and Australia. Basic research is experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundation of phenomena and observable facts, without any particular application or use in view. When there is a significant time lapse before the "results" of basic research can be applied, this is considered long-term research whose results are sometimes utilised at a much later date and to ends not foreseen by the initial researcher. There is evidence that innovation efforts draw increasingly on basic research, owing to greater possibilities of commercialisation of its results.

- The *ratio of science, engineering and health PhDs per population* aged 25 to 34 years in 1999 is highest in Switzerland, Sweden, Germany, Finland, the United Kingdom and France. This age group was chosen because it is the only one for which there are internationally comparable data. Human capital is heterogeneous, and no single type of attribute can adequately represent the many human characteristics that bear on the economy and society. While the level of individuals' skills, knowledge and competencies can be taken to represent the "stock" of human capital at any one time, these various attributes cannot be easily quantified. Measures of educational attainment consider qualification levels as a proxy for skills, despite their imperfections. The skill level is reflected by the highest level of education completed by each adult as defined by the ISCED (International Standard Classification of Education).

- The *number of scientific and technical articles per million population* -- a proxy for the scientific output per country -- in 1997 was highest in Switzerland, Sweden, Denmark, Finland and the Netherlands. Article counts of scientific research are based on scientific and engineering articles published in approximately 5 000 of the world's leading scientific and technical journals. Article counts are based on fractional assignments; for example, an article with two authors from different countries is counted as one-half article to each country.

Industry-science linkages are also an important element for future innovative capacity. Innovation depends not only on the creation of knowledge but also on flows of knowledge within and between economies. Greater formal and informal knowledge-sharing among R&D-conducting firms and between these firms and public research is critical for boosting innovative effectiveness. Collaboration between business and non-business entities is becoming more common. Businesses are eager to exploit research undertaken by the higher education and government sectors, the higher education sector is interested in obtaining funding for current and future research activities by commercialising its research efforts, and governments look to alliances that ensure that the economy benefits from public research. It should be noted that owing to problems in measuring linkages and to missing values, this is the least robust of the three performance categories (although several relevant indicators are now under development). For example, a low level of business-financed R&D performed by the public sector (as is the case in *e.g.* the United States) is not necessarily an indication of low industry-science linkages. Industry-science linkages are measured here by three variables with the following results (**Table 5**):

- *Business-financed R&D performed by government or higher education as a percentage of GDP* in 1999 is highest in Iceland, the Netherlands, New Zealand, Finland and the United Kingdom. R&D expenditure financed by industry but performed by public research institutions or universities is a primary indicator of links between the public and private sectors. The financial resources are directed towards industry needs, and tend to be focused on applied and developmental research with more direct economic objectives than public research.

- The *number of scientific papers cited in US-issued patents per million population* is highest in Denmark, the United States, Ireland, Hungary, the United Kingdom, Canada and Australia. This indicator is based on US patent data and may favour English-speaking countries. Publications are the major output of scientific research. With the increase in scientific activity and the strong incentive for researchers to publish (publications are used to evaluate researchers in many countries), the number of journals and articles is growing steadily. By 1997, the scientific output of the OECD area amounted to around half a million articles annually.

- Whereas the number of publications indicates quantity, quality is more closely associated with the number of highly cited papers or by their industry relevance. The ratio of *publications in the 19 most industry-relevant scientific disciplines per million population* between 1980 and 1995 is highest in Sweden, the United Kingdom, Switzerland, Finland, the Netherlands, New Zealand, Denmark and Australia.

Industrial innovation depends strongly on private sector investments in R&D and patents and their ability to develop and exploit new technology and ideas are important measures of the extent to which business is innovating. While the government and the higher education sectors carry out research, industrial R&D is most closely linked to the creation of new products and production techniques, as well as to a country's innovation efforts. Industrial innovation is measured here by four variables with the following results (**Table 6**):

- *Business enterprise R&D (BERD) as a percentage of GDP* between 1996 and 1999 was highest in Sweden (2.8% of GDP), followed by Japan, Switzerland, the United States, Finland and Korea. This covers R&D activities carried out in the business sector regardless of the origin of funding. It has to be mentioned that R&D data are often underestimated, especially in small and medium-sized enterprises (SMEs) and in service industries. They also tend to measure inputs (and not outputs) and may give little indication of the efficiency or quality of innovative activities.

- The number of *business researchers per 10 000 labour force* in 1999 is highest in the United States, Japan, Sweden, Norway and Finland. The indicator of R&D personnel is limited to researchers, who are viewed as the central element of the R&D system. Researchers are defined as professionals engaged in the conception and creation of new knowledge, products, processes, methods and systems and are directly involved in the management of projects.

- *The number of patents in "triadic" patent families per million population* in 1995 was highest in Switzerland, Sweden and Japan, followed by Germany, Finland, the Netherlands and the United States. An inventor seeking protection files a first application (priority) generally in his/her country of residence. Then, the inventor has a 12-month legal delay for applying or not for protection of the original invention in other countries. Patent families, as opposed to patents, are a set of patents taken in various countries for protecting a single invention. They are provided with the intention of improving international comparability (the "home advantage" is suppressed, the values of the patents are more homogeneous). A patent is a member of the patent families if and only if it is filed at the European Patent Office (EPO), the Japanese Patent Office (JPO) and is granted by the US Patent & Trademark Office (USPTO).

- *The share of firms having introduced at least one new or improved product or process on the market* over a given period of time is an indicator of the output of innovative activities. This indicator is taken from the Community Innovation Survey (CIS2) managed by Eurostat. It is weighted here by number of employees in order not to underestimate the weight of large firms (unweighted results would give an unduly large weight to the mass of small firms). According to the CIS, innovative firms represent between 70% and 80% of all firms in several countries (*e.g.* Germany, the Netherlands) which raises serious questions about the reliability of the survey. In addition, data are available only for a limited number of countries (21 OECD countries). Such surveys are relatively new and comparability of data across countries may still be limited. In particular, the coverage of services is partial in some countries.

Table 4: Underlying indicators for "Generation of New Knowledge"

	R&D performed by the non-business sector as a percentage of GDP	Non-business researchers per 10 000 labour force	Basic research as a percentage of GDP	PhD graduation rates in science, engineering and health	Scientific and technical articles per million population	Simple average of indicators
Australia	62.5 (0.84)	87.1 (51.9)	49.3 (0.41)	38.0 (0.70)	51.8 (735)	57.8
Austria	41.7 (0.65)	18.5 (15.3)	31.9 (0.29)	30.9 (0.57)	39.7 (570)	32.5
Belgium	27.9 (0.53)	40.2 (26.8)	..	23.6 (0.44)	44.9 (641)	34.2
Canada	51.4 (0.74)	39.8 (26.6)	..	23.6 (0.44)	55.5 (786)	42.6
Czech Republic	15.8 (0.42)	17.7 (14.9)	20.3 (0.21)	19.1 (0.36)	17.6 (267)	18.1
Denmark	52.5 (0.75)	54.8 (34.7)	53.6 (0.44)	25.5 (0.47)	73.2 (1028)	51.9
Finland	72.3 (0.92)	98.7 (58.1)	..	51.0 (0.93)	66.6 (938)	72.2
France	64.2 (0.85)	51.5 (32.9)	63.2 (0.51)	42.7 (0.78)	38.7 (557)	52.1
Germany	53.2 (0.75)	38.6 (26.0)	55.8 (0.46)	72.7 (1.32)	38.1 (548)	51.7
Greece	14.3 (0.41)	30.7 (21.8)	..	..	16.9 (258)	20.6
Hungary	14.2 (0.41)	33.0 (23.0)	18.1 (0.20)	22.1 (0.41)	15.6 (240)	20.6
Iceland	100.0 (1.17)	100.0 (58.7)	38.9 (0.34)	0.5 (0.02)	..	59.8
Ireland	11.3 (0.38)	23.4 (17.9)	2.9 (0.09)	30.7 (0.57)	27.4 (401)	19.1
Italy	21.6 (0.47)	29.3 (21.0)	24.6 (0.24)	12.5 (0.24)	23.9 (354)	22.4
Japan	58.6 (0.80)	51.8 (33.0)	..	23.0 (0.43)	25.9 (381)	39.8
Korea	48.9 (0.71)	19.4 (15.7)	39.4 (0.34)	22.1 (0.41)	6.7 (118)	27.3
Mexico	0.0 (0.28)	0.0 (5.4)	0.0 (0.07)	..	0.0 (27)	0.0
Netherlands	72.6 (0.92)	39.9 (26.7)	26.7 (0.25)	33.4 (0.62)	62.3 (879)	47.0
New Zealand	54.1 (0.76)	55.4 (34.9)	..	0.0 (0.02)	51.2 (728)	40.2
Norway	51.5 (0.74)	57.8 (36.3)	24.9 (0.24)	17.6 (0.33)	52.7 (748)	40.9
Poland	16.5 (0.43)	40.7 (27.1)	19.8 (0.21)	..	8.5 (143)	21.4
Portugal	25.5 (0.51)	36.0 (24.6)	11.3 (0.15)	..	9.1 (152)	20.5
Slovak Republic	6.5 (0.34)	39.0 (26.2)	17.9 (0.19)	..	16.1 (246)	19.9
Spain	16.5 (0.43)	42.1 (27.9)	11.0 (0.15)	14.9 (0.28)	23.5 (348)	21.6
Sweden	71.7 (0.92)	62.8 (38.9)	63.8 (0.51)	99.6 (1.81)	85.0 (1190)	76.6
Switzerland	58.6 (0.80)	33.9 (23.5)	100.0 (0.76)	100.0 (1.81)	100.0 (1395)	78.5
Turkey	6.3 (0.34)	2.1 (6.6)	..	5.1 (0.11)	0.8 (37)	3.6
United Kingdom	40.8 (0.64)	33.1 (23.1)	..	45.4 (0.83)	54.1 (767)	43.4
United States	44.8 (0.68)	16.7 (14.3)	52.3 (0.43)	32.2 (0.59)	46.1 (657)	38.4

Note: The figures represent the values according to the "minimum-maximum method" and those in parentheses the initial values. No data available for Luxembourg.

Table 5: Underlying indicators for "Industry-science linkages / technology diffusion"

	Business-financed R&D performed by government or higher education as a percentage of GDP	Scientific papers cited in US-issued patents	Publications in the 19 most industry-relevant scientific disciplines per million population	Simple average of indicators
Australia	20.4 (0.04)	55.9 (3.0)	53.2 (5.5)	43.2
Austria	0.7 (0.01)	10.3 (1.45)	34.0 (3.7)	15.0
Belgium	24.7 (0.05)	41.2 (2.5)	41.5 (4.4)	35.8
Canada	27.7 (0.06)	55.9 (3.0)	46.8 (4.9)	43.5
Czech Republic	14.3 (0.03)	..	..	..
Denmark	9.1 (0.03)	100.0 (4.5)	56.4 (5.8)	55.2
Finland	42.9 (0.08)	17.6 (1.7)	71.3 (7.2)	44.0
France	27.1 (0.06)	26.5 (2.0)	39.4 (4.2)	31.0
Germany	24.6 (0.05)	14.7 (1.6)	35.1 (3.8)	24.8
Greece	4.2 (0.02)	..	..	..
Hungary	13.3 (0.03)	61.8 (3.2)	30.9 (3.4)	35.3
Iceland	100.0 (0.18)	..	..	..
Ireland	13.3 (0.03)	73.5 (3.6)	31.9 (3.5)	39.6
Italy	3.5 (0.02)	14.7 (1.6)	27.7 (3.1)	15.3
Japan	2.2 (0.02)	11.8 (1.5)	30.9 (3.4)	14.9
Korea	27.5 (0.06)	0.0 (1.1)	16.0 (2.0)	14.5
Mexico	..	52.9 (2.9)	0.0 (0.5)	26.5
Netherlands	51.1 (0.10)	38.2 (2.4)	61.7 (6.3)	50.3
New Zealand	44.7 (0.08)	47.1 (2.7)	60.6 (6.2)	50.8
Norway	6.5 (0.02)	11.8 (1.5)	35.1 (3.8)	17.8
Poland	30.0 (0.06)	..	..	..
Portugal	0.0 (0.01)	8.8 (1.4)	10.6 (1.5)	6.5
Slovak Republic	6.3 (0.02)	..	..	..
Spain	12.1 (0.03)	44.1 (2.6)	36.2 (3.9)	30.8
Sweden	15.1 (0.04)	20.6 (1.8)	100.0 (9.9)	45.2
Switzerland	18.0 (0.04)	50.0 (2.8)	72.3 (7.3)	46.8
Turkey	33.8 (0.07)	..	..	..
United Kingdom	34.9 (0.07)	58.8 (3.1)	75.5 (7.6)	56.4
United States	7.4 (0.02)	91.2 (4.2)	35.1 (3.8)	44.5

Note: The figures represent the values according to the "minimum-maximum method" and those in parentheses the initial values.

Table 6: Underlying indicators for "Industrial innovation"

	BERD as a percentage of GDP	Business researchers per 10 000 labour force	Number of patents in "triadic" patent families per million population	Share of firms with new or technologically improved products or processes	Simple average of indicators
Australia	23.1 (0.72)	22.0 (15.12)	8.3 (8.2)	64.5 (59.9)	29.5
Austria	38.7 (1.14)	27.5 (18.73)	24.4 (24.1)	76.3 (66.2)	41.7
Belgium	46.5 (1.35)	40.2 (27.15)	32.0 (31.5)	22.3 (37.4)	35.2
Canada	34.6 (1.03)	46.6 (31.37)	11.9 (11.7)	78.4 (67.4)	42.9
Czech Republic	23.8 (0.74)	16.0 (11.14)	0.2 (0.3)	..	13.3
Denmark	42.3 (1.24)	39.0 (26.34)	31.2 (30.7)	84.1 (70.4)	49.1
Finland	66.7 (1.90)	61.1 (40.94)	50.4 (49.6)	53.8 (54.2)	58.0
France	47.6 (1.38)	41.7 (28.11)	30.4 (29.9)	47.2 (50.7)	41.7
Germany	54.7 (1.58)	50.6 (34.00)	53.1 (52.2)	100.0 (78.9)	64.6
Greece	2.1 (0.15)	5.5 (4.23)	0.1 (0.1)	..	2.6
Hungary	7.0 (0.28)	11.2 (7.96)	1.4 (1.4)	..	6.5
Iceland	28.4 (0.86)	51.0 (34.25)	21.9 (21.5)	..	33.7
Ireland	33.9 (1.01)	49.3 (33.13)	5.5 (5.4)	89.7 (73.4)	44.6
Italy	16.2 (0.53)	17.2 (11.95)	9.9 (9.7)	32.2 (42.7)	18.9
Japan	71.9 (2.04)	96.0 (63.99)	69.6 (68.5)	..	79.2
Korea	64.8 (1.85)	44.9 (30.26)	7.0 (6.9)	..	38.9
Luxembourg	..	..	27.3 (26.8)	95.8 (76.7)	61.5
Mexico	0.0 (0.09)	0.0 (0.58)	0.1 (0.1)	38.0 (45.8)	9.5
Netherlands	37.0 (1.10)	34.4 (23.30)	47.3 (46.5)	93.4 (75.4)	53.0
New Zealand	8.3 (0.32)	12.8 (9.05)	3.7 (3.6)	..	8.3
Norway	31.5 (0.95)	62.3 (41.74)	18.5 (18.2)	52.3 (53.4)	41.1
Poland	7.6 (0.30)	8.0 (5.89)	0.0 (0.1)	30.0 (41.5)	11.4
Portugal	2.4 (0.16)	2.8 (2.40)	0.2 (0.2)	0.0 (25.4)	1.3
Slovak Republic	18.4 (0.59)	14.0 (9.80)	0.4 (0.4)	..	10.9
Spain	12.6 (0.43)	13.0 (9.14)	2.2 (2.2)	17.5 (34.8)	11.3
Sweden	100.0 (2.81)	77.9 (52.08)	74.8 (73.6)	66.3 (60.9)	79.8
Switzerland	67.8 (1.93)	46.8 (31.51)	100.0 (98.4)	89.6 (73.4)	76.1
Turkey	2.9 (0.17)	1.3 (1.44)	0.0 (0.0)	14.7 (33.3)	4.7
United Kingdom	42.0 (1.23)	47.5 (31.94)	22.6 (22.2)	64.7 (60.1)	44.2
United States	67.6 (1.93)	100.0 (66.66)	43.1 (42.4)	..	70.2

Note: The figures represent the values according to the "minimum-maximum method" and those in parentheses the initial values.

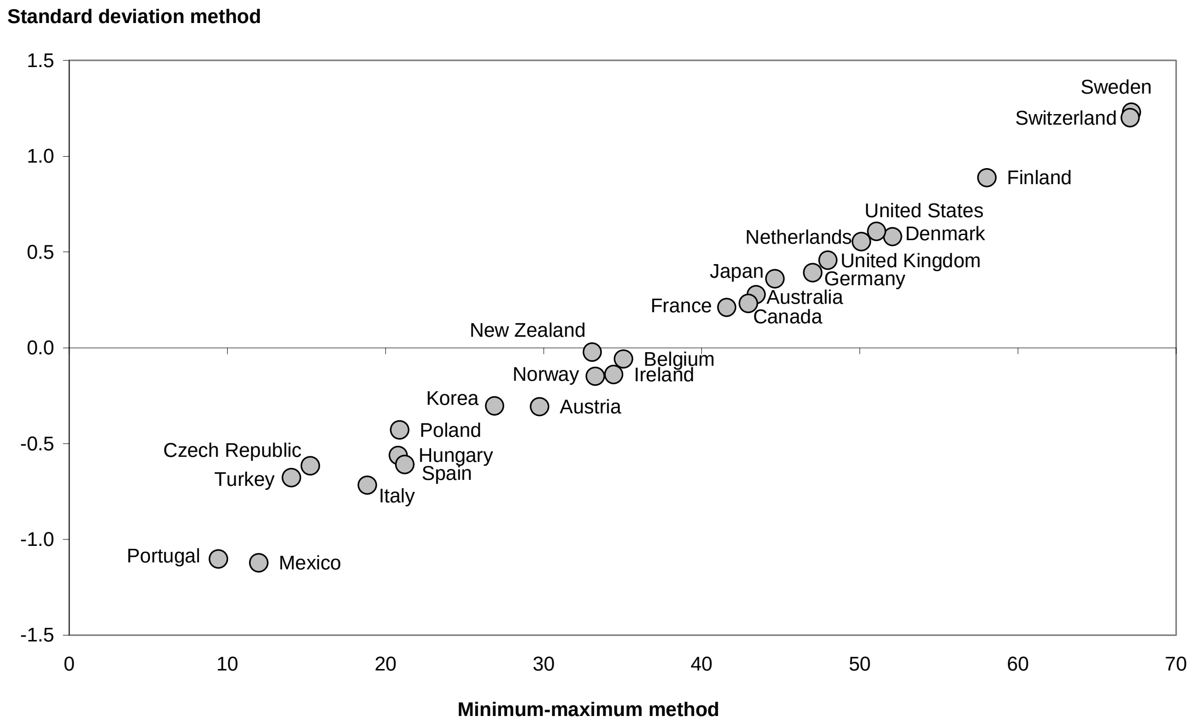
Standardising variables

The variables selected for the index are expressed in various units (*e.g.* R&D expenditures as a percentage of GDP or number of patents per million population) and have to be normalised to render them comparable. The "minimum-maximum" method is used here to standardise or normalise the variables. This method yields indices that range by construction between 0 (laggard with minimum value) and 100 (leader with maximum value). For a given country, the index expresses their distance from the overall best and the worst performing countries:

$$100\left(\frac{\text{actual value - minimum value}}{\text{maximum value - minimum value}}\right)$$

The normalisation method does not affect the country rankings for individual indicators (since any normalisation method is just a simple transformation of the initial values, see **Table 2**). In contrast, it can affect the overall findings of a composite indicator, since individual indicators are not only normalised, but also aggregated into a composite (see **Table A1**). The results of different standardisation approaches have been compared for robustness. The “minimum-maximum method” is compared here to the “standard deviation from the mean” method, where positive (negative) values for a given country indicate above (below)-average performance. Taking an example based on the simplest of the four methods of weights (each indicator within a component has the same weight to derive a sub-index; each of the three sub-indices has the same weight to derive the overall index), it can be seen that the two normalisation methods yield similar, although not identical, rankings of countries (**Figure 1**). Similar comparisons with other methods suggest that the composite indicator is not unduly influenced by the standardisation method (most correlation coefficients are in the .98-.99 range).

Figure 1. Comparison of two standardisation methods for an index of innovation performance



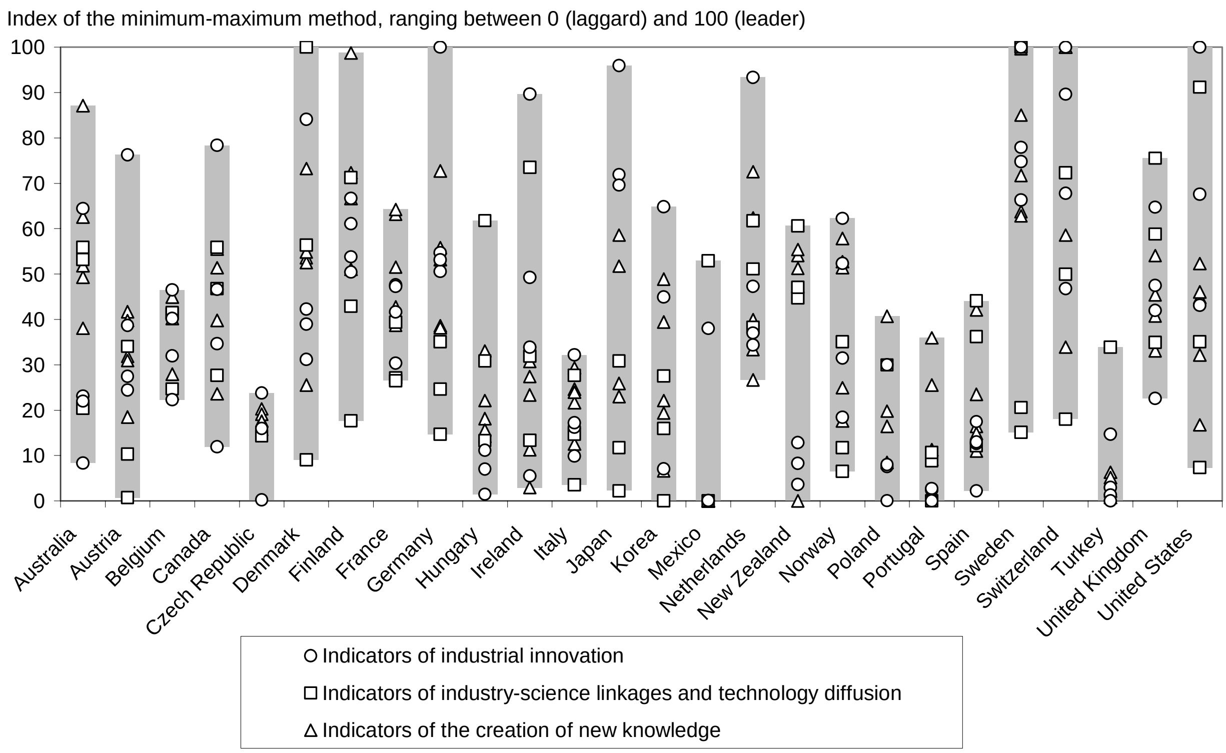
Weighting variables

Whereas the influence of the standardisation method on the results of composite indicators seems limited, the weights attached to individual indicators in contrast strongly influence the overall index. An index based on these 12 indicators can have very different values for some countries depending on the weights. The spread of the values according to the "minimum-maximum approach" for the underlying indicators shows the dispersion of the extreme values of the indicators, and gives an idea of the values a composite index could theoretically attain (**Figure 2**).

The indices for Belgium, for example, range between 22 and 47: this means that whatever weight the indicators are given, Belgium could never have an overall index that is lower than 22 or higher than 47. In contrast, Sweden for example has indices that range between 15 and 100. In practice, the composite indicator does not take any of the extreme values (this would only happen if zero weights are attached to all indicators other than those at a single extreme) but rather some value in-between. However, the results for countries with a polarised innovation profile (*i.e.* low performance in some areas and high performance in others) such as Japan, Sweden and the United States depend more strongly on the weights than others.

But how to assign the weights? *Derive them from a theoretical model? Estimate them empirically? Assign them arbitrarily? Or assign them randomly?* The first two are not feasible, and the latter two are not satisfying. This appears to be a major problem which makes the entire exercise hazardous. However, there is some scope for examining the empirical significance of this problem. The following sections will examine and compare the results stemming from arbitrarily and randomly assigned weights to assess the sensitivity of the composite to the choice of the weights.

Figure 2. The range of the indices for the 12 indicators of innovation performance



Two-step approach to weighting

An often-used approach to weighting, and one which requires some judgement as to how to assign weights, proceeds in a two-stage approach using sub-indices on the assumption that indicators that are related in some way can usefully be organised into “thematic”sub-groups. First, it assigns a weight to the indicators in a given component to derive a sub-index; and then it assigns a weight to the sub-indices to derive an overall index. This two-step approach is partly done to avoid overestimating (underestimating) those components for which more (fewer) indicators are available.

In the first step, the sub-index for each component is defined as the simple average of the indices of the underlying indicators that relate to a common theme. Equal weights are assigned to the underlying variables within a given component. Although some variables are weaker than others, it was decided not to give greater weight to those with more robust data. This is partly because, in the context of national innovation systems, many of these variables are interrelated and not easily distinguishable. If a value for an underlying variable is missing, the core component is expressed as the weighted average of the other indicators within the same component. The simple average reported in the last columns of Tables 4, 5 and 6 are illustrative of what thematic sub-indices might look like.

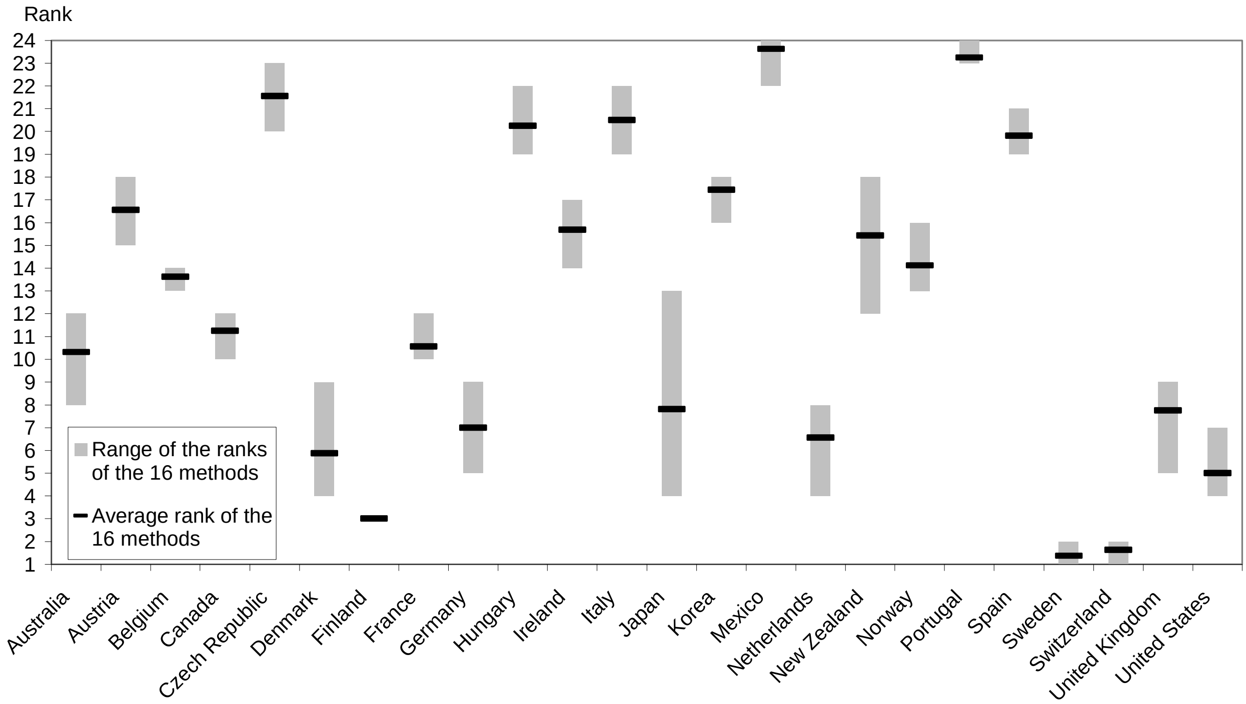
In the second step, an overall index is obtained by weighting the sub-indices of the three underlying components. The simplest approach would be to structure it such that the three core components have the same weight, which implicitly assumes them to be of equal importance in influencing a country's innovation performance, although this is obviously not necessarily the case. Although some components or indicators may play a more important role than others in determining a country's innovation performance, it is difficult to establish a hierarchy. The theoretical framework on innovation performance is not sufficiently developed to give variable weights to the components or indicators selected

While the robustness of results with respect to changes in weights may be a good sign, it can also signal potential problems with the specification of the components: if components are highly correlated, a composite will always have similar indices, no matter what the weights. Bivariate correlation coefficients are thus calculated to discern correlations among country values of the individual indicators and the three thematic components. **Table A2** shows that the indices for "g*eneration of new knowledge*" are strongly correlated with "*industrial innovation*" (correlation coefficient of 0.74) and moderately correlated with "*industry -science linkages*" (0.49). In contrast, the correlation between "*industry–science linkages*" and "*industrial innovation*" is relatively weak (0.26). At a more detailed level, the variables are reasonably well correlated, both with indicators from the same component and with those from other components.

The index can be analysed for its sensitivity to the weights chosen for the three components. The quality and reliability of the components differ strongly, and sub-index 2 (*industry-science linkages*) is thought to be much weaker than sub-index 1 (*generation of new knowledge*) and sub-index 3 (*industrial innovation*). Thus, giving a weight of one-third to each of the three components may assign unreasonable weight to "industry-science linkages". Calculations are thus carried out based on indices with four different weights of sub-indices (33%-33%-33%, 20-40-40, 40-20-40 and 40-40-20). Simultaneously, four different normalisation methods (minimum-maximum method, maximum method, mean method and the standard deviation from the mean) are considered to provide an assessment of the sensitivity of the ordering of countries to the issues so far considered here.

As can be seen for the 24 countries that have enough data coverage in all specifications, overall indices are sensitive to these choices (especially concerning the weights) and individual country rankings can thus be affected (**Figure 3**). Whereas the rankings for *e.g.* Finland and Belgium are hardly affected by the choice of the weights and the normalisation method, the weights do matter strongly for example for Japan. Depending on the specification, Japan, which has a polarised innovation profile (good performance in industrial innovation and a low performance in industry-science links), is in the 4th position in two specifications and in the 13th place in two specifications (which incidentally give a low weight to industrial innovation). The results seem much more robust when the "unit of analysis" is not the individual country, but when groups of countries are examined. For example, the 12 countries with the highest average ranks are the 12 top performing countries in any specification (with the exception of Japan), and vice versa for the six lagging countries (**Table A4**).

Figure 3. Range and average rankings of overall index of innovation performance with four different normalization methods and four different weights of sub-indices



Notes: The rankings are from 1 (best performer) to 24 (worst performer). Six countries are excluded because they do not have enough data coverage and may be missing from some (Iceland, Poland, Slovak Republic, Turkey) or all specifications (Greece and Luxembourg).

Randomly assigned weights

In a first test, the robustness of country rankings has been explored on the basis of arbitrarily defined weights. In a further sensitivity test, the relative weights for the 12 underlying indicators were chosen randomly. This was done 10 000 times, resulting in 10 000 different random composite indices (according to the "minimum-maximum method"). **Figure 4** shows the results of the dispersion of the index by country, with countries being ranked by the average random index. Areas in light grey indicate 5% of the highest and 5% of the lowest indices, and those in dark grey the spread of the remaining 90% of the results. The random composite indices for Sweden, for example, range between 52 and 87 in all cases, and between 62 and 77 in 90% of the cases. Whereas the average of these indices is very close to an index calculated on the basis of predefined weights (equal weights for indicators within a given component and equal weights for the three components), the range of the random composite indices is substantial for most countries. Countries with close average indices thus have ranges that strongly overlap (even at the 90% level), suggesting that country rankings substantially depend on the choice of weights.

Figure 4. Range and average of 10 000 random composite indices

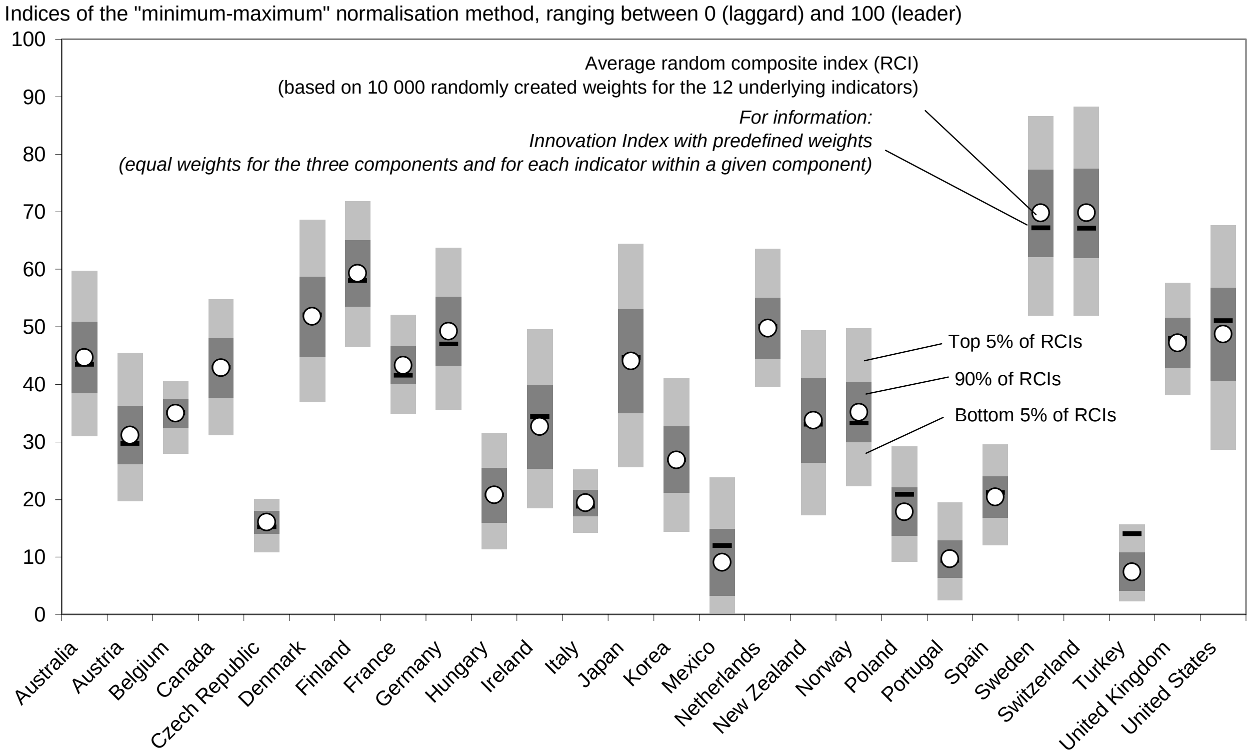
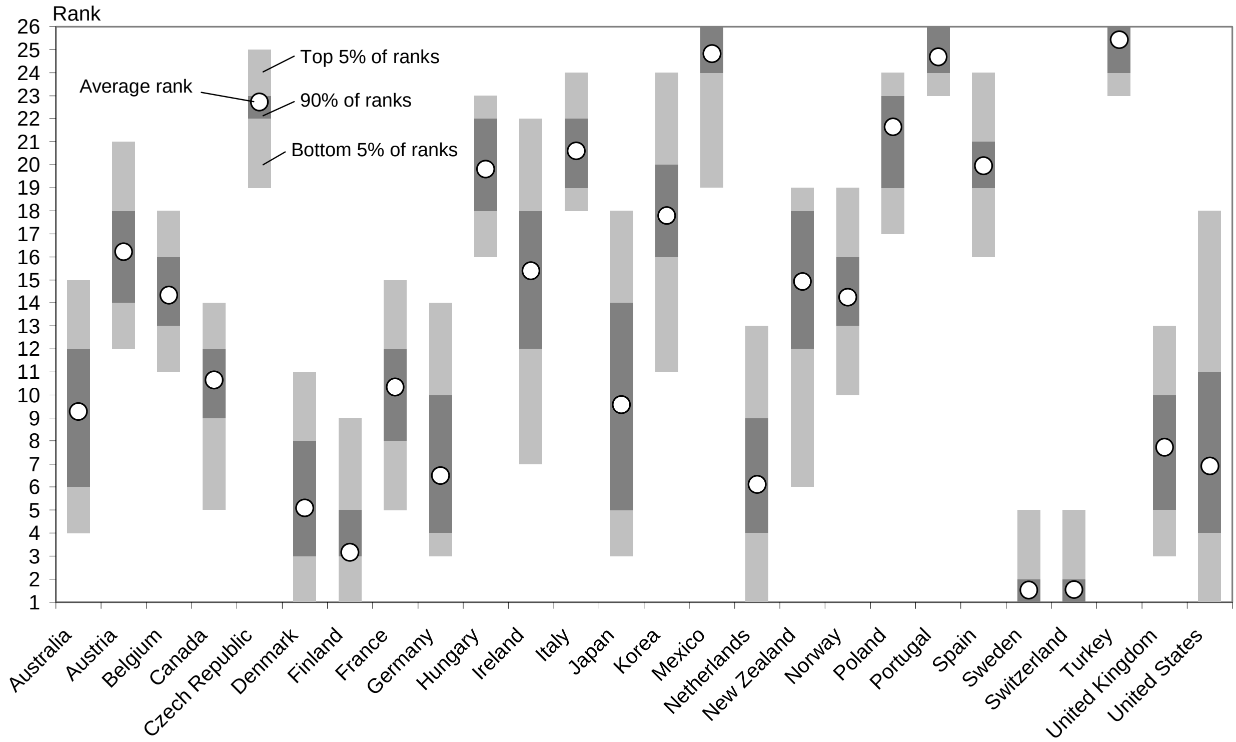


Figure 5 translates the 10 000 random composite indices directly into country rankings. The United States, for example, ranks first in at least one case and is on rank 18 in other extreme cases; and is between rank 4 and rank 11 in 90% of the cases. Again, the overlap in the range and thus the probability of changes in country rankings is very high.

Finally, **Table A3** shows the frequency (in %) that a given country is among the top 1 (2,3, .. 26) OECD countries, based on the 10 000 random composite indices. The results suggest that it is impossible to identify a single candidate for the top performing OECD country: in fact there are two candidates, as both Switzerland and Sweden rank in the 1st place in about 50% of the cases. In contrast, it is possible to identify these two countries as the top two performers (both are on rank 1 or 2 in more than 95% of the cases, whereas the best third country is in this group in only 4% of the cases). Likewise, there are three clear candidates to be among the top 3: the two former countries (with a frequency of nearly 100%) and Finland (86%), whereas the best fourth country is in this group in only 8% of the cases. It is in turn impossible to identify a clear group of countries that would be the top 4, 5, 6, 7, 8, 9, 10 and 11 performers. Only the top 12 countries form again a group that distinguishes it from the remaining countries.

Two main results emerge from these sensitivity tests. First, although country rankings may be justified in some cases, they can be meaningless and even misleading in other cases, especially when applied mechanically. Second, it is more robust to identify "groups" of best-performing countries, but even here, the group size may need to be determined through sensitivity tests.

Figure 5. Range of country ranks in 10 000 random composite indices



The rankings are from 1 (best performer) to 26 (worst performer). Four countries do not have enough data coverage and are excluded (Greece, Iceland, Luxembourg, Slovak Republic).

Back to the details: the need to use individual indicators

Composite indicators provide a starting point for analysis and can be used to guide further analytical and policy-related work, but every analysis will need to revisit the original data. This has two main objectives: i) the de-composition of the results of composites, so the contribution of the components and individual indicators to the aggregate results can be identified; and ii) the profiling of national performance to identify major strengths and weaknesses at a disaggregated level.

The three components of the index contribute very differently to the aggregated index in the various countries. For example, a country like Japan is very strong in industrial innovation but weak in industry-science linkages, which results in an average overall performance. New Zealand has the opposite results, where very low industrial innovation is combined with good industry-science linkages. This decomposition of the composite into sub-components gives an idea of what leads to the overall innovation performance of different countries.

The performance in each of the sub-components can be further disaggregated and national strengths and weaknesses can be seen at the indicator level. Japan's low ranking in industry-science linkages can be explained by its low performance on two indicators: i) government or higher education R&D financed by business, and ii) scientific papers cited in US-issued patents. Japan's good performance in industrial innovation is confirmed by all the relevant indicators, but especially in the number of business researchers where Japan is the leading country (**Figure 6**). In the case of New Zealand, the low ranking on industrial innovation is explained by low performance on three of the indicators, while its high ranking on industry-science linkages is due the large share of R&D financed by business (**Figure 7**). Composites can in this way be deconstructed to explain the underlying reasons for overall country performance.

Figure 6. Profiling Japan's innovation performance

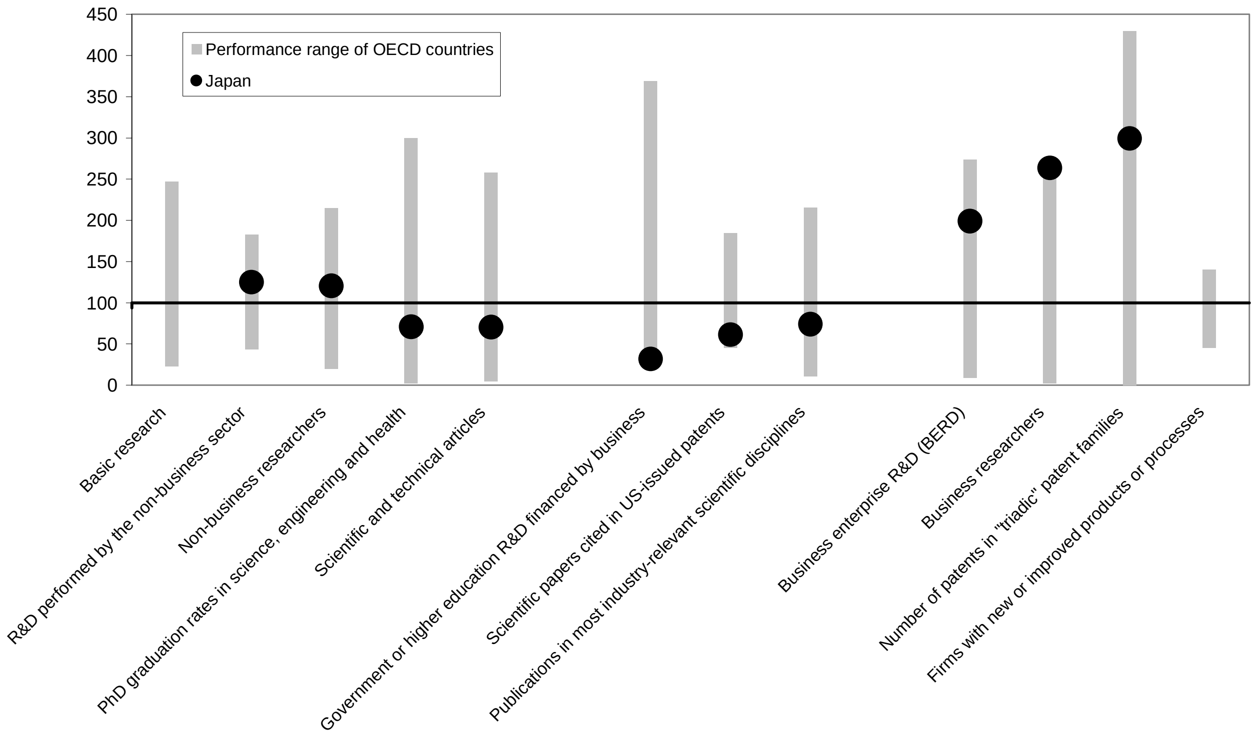
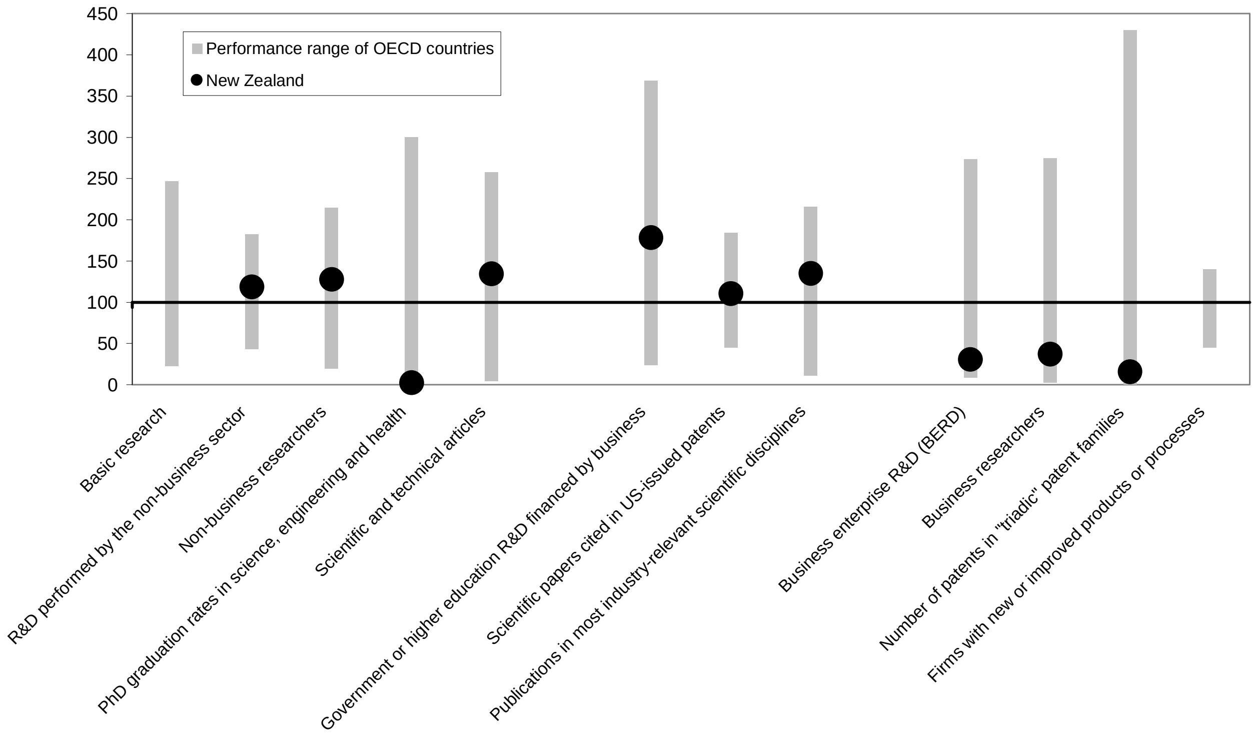


Figure 7. Profiling New Zealand's innovation performance



CONCLUSIONS

Due to their value as a communication and political tool, composite indicators of country performance will continue to proliferate. This is despite their many methodological shortcomings. The construction and availability of consistent and comparable indicators of country structure and performance are important to the analysis of a variety of economic policy areas. In using indicators, it is important to understand their inter-relationships and interactions as well as their relative importance to the economic performance and policies concerned. However, in the absence of evidence of their relative contribution and significance, the construction of single composite indicators of performance can be based on spurious degrees of quantification and lead to misleading country comparisons. These risks become larger as composite indicators extend their coverage over broader and more loosely defined conceptual areas.

For these reasons, composite indicators risk becoming exercises in measurement without a theoretical underpinning. Ranking countries (*e.g.* for competitiveness studies) is becoming increasingly popular and widespread. Even where the conceptual problems can be solved, there remain two major sets of problems. The first concerns the information that is *not* available. The construction of composites can only draw on the data that exist. There is no meaningful way to adjust a composite for information that should be reflected in a performance measure but is not reflected in available indicators.

The second concerns the use of what *is* available. In this regard, the most important problem may be the selection of weights for the different components. To be useful for policy, weights need to reflect the relative importance of individual indicators in determining performance outcomes. The choice of weights should vary according to the area being studied and either have a theoretical basis or be determined through econometric or statistical tests of relevant relationships. These techniques can also help overcome deficiencies concerning non-linearities in the underlying relationships, interaction among variables, risks of double-counting as well as accounting for the influence of missing variables. However, all too often, the choice of weighting scheme remains ad hoc and mostly arbitrary due to the lack of theoretical constructs and empirical foundation.

Notwithstanding the importance of these considerations, some pragmatism in approaching the use of composites is desirable. Empirical and policy analyses often benefit from the use of measures and indicators that are less the ideal and to rule out the use of composites altogether would be going too far. But their conceptual limits must be kept in mind and they should be supplemented as appropriate with other information and supporting analysis. Where they are used, they should be accompanied by an account of their methodological limits and include detailed explanations of the underlying data sets, choice of standardisation technique and selection of weighting method. In addition, they should include an assessment of robustness to alternative ways of treating the methodological issues that arise in their construction.

This paper has examined many of these issues as they would be encountered in the construction of an index to measure “innovation performance” in OECD countries. The conceptual obstacles to creating such an index are severe: available indicators largely measure inputs or reflect progress at various stages of the innovation process while relatively few reflect outputs or aspects of innovative performance. Furthermore, the relevance of the indicators that are available is highly variable, and in some cases (*e.g.* scientific papers cited in US-issued patents) the appropriateness for cross-country comparison purposes is questionable.

Such an index is also quite sensitive to variations in methods of standardising the data and in weighting the individual indicators. As a result, any effort to rank countries on the basis of such an index or to treat it as data for empirical purposes would be very hazardous even if the high or low rankings of a few countries seem reasonably robust. Any such index must be used very cautiously in any empirical or policy analytic work and be supplemented by other information and analysis.

In order to improve the transparency of composite indicators and to minimise their misuse, the OECD and the European Commission (EC) are undertaking joint work with a view to developing "quality guidelines" for the construction and application of composite indicators.

REFERENCES

Center For International Development (CID) (2002), *Networked Readiness Index* (NRI), Harvard University, Cambridge, US.

European Commission (EC) (2001), *European Innovation Scoreboard*, Brussels.

EC Joint Research Centre (JRC) (2002), "State-of-the-Art Report on Current Methodologies and Practices for Composite Indicator Development", Institute for the Protection and Security of the Citizen, Ispra, Italy.

Eurostat (2002), "European Innovation Scoreboard: Draft Methodology Report 2002", Luxembourg.

Gregoir, S. and F. MAUREL (2002), "Les indices de compétitivité des pays : interprétation et limites", INSEE Département des Etudes Economiques d'Ensemble, Paris.

National Science Foundation (NSF) (2000), *Science and Engineering Indicators*, Washington, D.C.

OECD (2001), *The New Economy: Beyond the Hype* - *The OECD Growth Project*.

OECD (2002), "Aggregated Environmental Indices; Review of Aggregation Methodologies in Use", [ENV/EPOC/SE(2001)2/FINAL].

Office For National Statistics (ONS) (2002), *Towards a Measurement Framework for International e-Commerce Benchmarking*, London.

World Markets Research Centre (WMRC) (2001), *World Markets Research Centre Globalisation Index*, London.

Yale Center for Environmental Law and Policy (2002), *Environmental Sustainability Index*.

ANNEX TABLES -- Table A1. Example of country rankings in a composite index based on two indicators and four standardisation methods

	Initial values and ranks				Distance from the mean (Average=100)				Distance from the leader (Leader=100)				Distance from the best and worst performers (Leader=100, laggard=0)				Standard deviation from the mean			
	A	Rank	B	Rank	A	B	A&B	Rank	A	B	A&B	Rank	A	B	A&B	Rank	A	B	A&B	Rank
Switzerland	1395	*(1)*	98.4	*(1)*	258	432	345	*(1)*	100	100	100	*(1)*	100	100	100	*(1)*	2.5	2.9	2.7	*(1)*
Sweden	1190	*(2)*	73.6	*(2)*	220	323	271	*(2)*	85	75	80	*(2)*	85	75	80	*(2)*	1.9	1.9	1.9	*(2)*
Finland	938	*(4)*	49.6	*(5)*	174	218	196	*(3)*	67	50	59	*(3)*	67	50	59	*(3)*	1.1	1.0	1.1	*(3)*
Netherlands	879	*(5)*	46.5	*(6)*	163	204	183	*(5)*	63	47	55	*(4)*	62	47	55	*(4)*	1.0	0.9	0.9	*(4)*
Denmark	1028	*(3)*	30.7	*(9)*	190	135	162	*(7)*	74	31	52	*(5)*	73	31	52	*(5)*	1.4	0.3	0.9	*(5)*
Japan	381	*(17)*	68.5	*(3)*	70	301	185	*(4)*	27	70	48	*(6)*	26	70	48	*(6)*	-0.5	1.8	0.6	*(6)*
Germany	548	*(15)*	52.2	*(4)*	101	229	165	*(6)*	39	53	46	*(7)*	38	53	46	*(7)*	0.0	1.1	0.6	*(7)*
United States	657	*(11)*	42.4	*(7)*	122	186	154	*(8)*	47	43	45	*(8)*	46	43	45	*(8)*	0.3	0.8	0.5	*(8)*
Belgium	641	*(12)*	31.5	*(8)*	119	138	128	*(9)*	46	32	39	*(9)*	45	32	38	*(9)*	0.3	0.3	0.3	*(10)*
United Kingdom	767	*(7)*	22.2	*(12)*	142	98	120	*(10)*	55	23	39	*(10)*	54	23	38	*(10)*	0.7	0.0	0.3	*(9)*
Norway	748	*(8)*	18.2	*(13)*	138	80	109	*(12)*	54	18	36	*(11)*	53	18	36	*(11)*	0.6	-0.2	0.2	*(11)*
France	557	*(14)*	29.9	*(10)*	103	131	117	*(11)*	40	30	35	*(12)*	39	30	35	*(12)*	0.0	0.3	0.2	*(12)*
Canada	786	*(6)*	11.7	*(14)*	145	52	98	*(14)*	56	12	34	*(13)*	55	12	34	*(13)*	0.7	-0.4	0.1	*(13)*
Austria	570	*(13)*	24.1	*(11)*	105	106	106	*(13)*	41	24	33	*(14)*	40	24	32	*(14)*	0.1	0.0	0.1	*(14)*
Australia	735	*(9)*	8.2	*(16)*	136	36	86	*(15)*	53	8	31	*(15)*	52	8	30	*(15)*	0.6	-0.6	0.0	*(15)*
New Zealand	728	*(10)*	3.6	*(19)*	135	16	75	*(16)*	52	4	28	*(16)*	51	4	27	*(16)*	0.5	-0.7	-0.1	*(16)*
Italy	354	*(18)*	9.7	*(15)*	65	43	54	*(17)*	25	10	18	*(17)*	24	10	17	*(17)*	-0.5	-0.5	-0.5	*(17)*
Ireland	401	*(16)*	5.4	*(18)*	74	24	49	*(18)*	29	6	17	*(18)*	27	5	16	*(18)*	-0.4	-0.7	-0.5	*(18)*
Spain	348	*(19)*	2.2	*(20)*	64	10	37	*(19)*	25	2	14	*(19)*	23	2	13	*(19)*	-0.6	-0.8	-0.7	*(19)*
Czech Republic	267	*(20)*	0.3	*(23)*	49	1	25	*(22)*	19	0	10	*(20)*	18	0	9	*(20)*	-0.8	-0.9	-0.8	*(20)*
Hungary	240	*(23)*	1.4	*(21)*	44	6	25	*(21)*	17	1	9	*(21)*	16	1	9	*(21)*	-0.9	-0.8	-0.8	*(22)*
Greece	258	*(21)*	0.1	*(25)*	48	1	24	*(23)*	19	0	9	*(22)*	17	0	9	*(22)*	-0.8	-0.9	-0.8	*(21)*
Slovak Republic	246	*(22)*	0.4	*(22)*	46	2	24	*(24)*	18	0	9	*(23)*	16	0	8	*(23)*	-0.8	-0.9	-0.9	*(23)*
Korea	118	*(26)*	6.9	*(17)*	22	30	26	*(20)*	8	7	8	*(24)*	7	7	7	*(24)*	-1.2	-0.6	-0.9	*(24)*
Portugal	152	*(24)*	0.2	*(24)*	28	1	14	*(25)*	11	0	6	*(25)*	9	0	5	*(25)*	-1.1	-0.9	-1.0	*(25)*
Poland	143	*(25)*	0.1	*(27)*	26	0	13	*(26)*	10	0	5	*(26)*	8	0	4	*(26)*	-1.1	-0.9	-1.0	*(26)*
Turkey	37	*(27)*	0.0	*(28)*	7	0	3	*(27)*	3	0	1	*(27)*	1	0	0	*(27)*	-1.5	-0.9	-1.2	*(27)*
Mexico	27	*(28)*	0.1	*(26)*	5	1	3	*(28)*	2	0	1	*(28)*	0	0	0	*(28)*	-1.5	-0.9	-1.2	*(28)*
Unweighted average	541		22.8		100	100	100		39	23	31		38	23	30		0.0	0.0	0.0	
Maximum value	1395		98.4		258	432	345		100	100	100		100	100	100		2.5	2.9	2.7	
Minimum value	26.6		0.0		5	0	3		2	0	1		0	0	0		-1.5	-0.9	-1.2	

Indicator A: Scientific and technical articles per million population.
Indicator B: Number of patents in "triadic" patent families.
Excluding Iceland and Luxembourg.

Table A2. Correlation coefficients of variables and indices

	Generation of New Knowledge*	R&D performed by the non-business sector as % of GDP	Basic research as % of GDP	Scientific and technical articles per million population	Non-business researchers per 10 000 labour force	PhD graduation rates in science, engineering and health	**Industry-science linkages***	Business-financed R&D performed by government or higher education as % of GDP	Scientific papers cited in US-issued patents	Publications in the 19 most industry-relevant scientific disciplines per million population	**Industrial innovation***	BERD as % of GDP	Business researchers per 10 000 labour force	Number of patents in "triadic" patent families	Share of firms introducing new or technologically improved products or processes
Generation of New Knowledge*	1.00	.87	.88	.91	.71	.69	.49	.33	-.04	.84	.74	.72	.61	.78	.57
R&D performed by the non-business sector as % of GDP	.87	1.00	.68	.77	.73	.31	.33	.56	-.21	.66	.66	.62	.61	.62	.57
Basic research as % of GDP	.88	.68	1.00	.78	.32	.74	.43	.12	.05	.66	.77	.76	.57	.85	.49
Scientific and technical articles per million population	.91	.77	.78	1.00	.55	.67	.68	.23	.23	.89	.71	.64	.54	.74	.65
Non-business researchers per 10 000 labour force	.71	.73	.32	.55	1.00	.09	.30	.49	-.13	.57	.26	.26	.31	.27	.16
PhD graduation rates in science, engineering and health	.69	.31	.74	.67	.09	1.00	.26	-.19	-.10	.62	.71	.72	.39	.78	.52
Industry-science linkages*	.49	.33	.43	.68	.30	.26	1.00	.58	.71	.74	.26	.19	.15	.25	.53
Business-financed R&D performed by government or higher education as % of GDP	.33	.56	.12	.23	.49	-.19	.58	1.00	.04	.48	.02	-.00	.04	.04	.18
Scientific papers cited in US-issued patents	-.04	-.21	.05	.23	-.13	-.10	.71	.04	1.00	.16	-.01	-.15	.01	-.10	.33
Publications in the 19 most industry-relevant scientific disciplines per million population	.84	.66	.66	.89	.57	.62	.74	.48	.16	1.00	.52	.53	.35	.57	.45
Industrial innovation*	.74	.66	.77	.71	.26	.71	.26	.02	-.01	.52	1.00	.93	.90	.88	.81
BERD as % of GDP	.72	.62	.76	.64	.26	.72	.19	-.00	-.15	.53	.93	1.00	.87	.83	.57
Business researchers per 10 000 labour force	.61	.61	.57	.54	.31	.39	.15	.04	.01	.35	.90	.87	1.00	.71	.61
Number of patents in "triadic" patent families	.78	.62	.85	.74	.27	.78	.25	.04	-.10	.57	.88	.83	.71	1.00	.54
Share of firms introducing new or technologically improved products or processes	.57	.57	.49	.65	.16	.52	.53	.18	.33	.45	.81	.57	.61	.54	1.00

*. Simple average of components listed below.

Table A3. Frequency (in %) that a country is among the top 1 (2,3, .. 26) in a ranking of a composite index based on 10 000 randomly assigned weights for the underlying indicators.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	Switzerland	50	96	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
2	Sweden	49	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
3	Finland	1	4	86	95	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
4	Denmark	0	0	8	46	68	82	91	97	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
5	Netherlands	0	0	0	17	42	63	80	92	97	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
6	Germany			1	18	35	54	70	85	92	97	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
7	United States	0	0	5	18	36	49	62	74	82	88	95	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100
8	United Kingdom			0	1	5	22	46	70	88	95	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
9	Australia				0	4	11	21	35	51	68	84	98	100	100	100	100	100	100	100	100	100	100	100	100	100	100
10	Japan			0	5	10	18	26	35	48	57	67	87	94	97	98	99	100	100	100	100	100	100	100	100	100	100
11	France					0	1	3	8	23	50	82	98	100	100	100	100	100	100	100	100	100	100	100	100	100	100
12	Canada					0	0	1	4	20	44	68	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100
13	Norway										0	0	1	31	60	85	97	100	100	100	100	100	100	100	100	100	100
14	Belgium											0	2	28	58	82	96	100	100	100	100	100	100	100	100	100	100
15	New Zealand						0	0	0	1	2	3	10	26	42	60	74	90	100	100	100	100	100	100	100	100	100
16	Ireland							0	0	0	0	2	5	19	31	46	70	88	100	100	100	100	100	100	100	100	100
17	Austria												0	3	10	23	52	92	99	100	100	100	100	100	100	100	100
18	Korea											0	0	0	2	5	12	30	86	91	96	99	100	100	100	100	100
19	Hungary																0	1	13	51	69	86	98	100	100	100	100
20	Spain																0	0	2	27	80	95	99	100	100	100	100
21	Italy																		0	24	38	79	99	100	100	100	100
22	Poland																	0	0	7	16	36	78	99	100	100	100
23	Czech Republic																			0	1	4	25	97	100	100	100
24	Portugal																							0	49	82	100
25	Mexico																			0	0	0	1	4	41	72	100
26	Turkey																							0	9	47	100

Note: Based on 26 countries. Greece, Iceland, Luxembourg, Slovak Republic do not have enough data coverage and are excluded.