Exploring measures at the aggregate level

2. How about composite indicators?

Andrea Saltelli and Ernesto Villalba
(Directorate-General Joint Research Center,
European Commission)

Abstract

Composite indicators can be useful to measure complex, multidimensional concepts such as competitiveness, good governance, environmental stewardship, as well as creativity. This chapter outlines some of the methodological issues related to the creation of composite indicators and the necessity of using sensitivity and robustness analysis to assess the quality of the measure obtained.

Introduction

Knowledge-based policies are made possible by the availability of high quality statistical information and analytic tools to draw inference-based data. In the increasingly complex policy landscape of the 21st century with conflicting constituencies and stakeholders at several levels of government, new measures are needed. An example of this process was the creation of the Open Method of Coordination in the European Union.

As Pisani-Ferry and Sapir (2006) have pointed out: '(…) civil societies learn from the experience of others. Such policy learning can be enhanced by initiatives that facilitate cross-country comparison and benchmarking'.

Composite indicators can serve this purpose particularly well. An indicator is a 'quantitative or a qualitative measure derived from a series of observed facts that can reveal relative positions (e.g. of a country) in an area' (Nardo et al., 2005, 13). A composite indicator is formed when individual indicators are compiled into a single index, on the basis of an underlying model of the multidimensional latent concept that is being measured (ibid.). Composite indicators encapsulate a series of different dimen-

sions to form a single measure that represents complex phenomena. It can provide a picture of a complex phenomena, possibly understandable for the public and non-technical people.

Composite indicators and the media

It is not surprising, thus, that media usually gives a lot of attention to this sort of indicator and uses them regularly. The journal *The Economist*, for example, uses different rankings — mostly created by *The Economist's* own intelligence unit — to inform their readers about the housing market, country risk, education opportunities, and so on, including the celebrated Big Mac Index[1]. Composite indicators, such as the Global Competitive Index[2] or the Human Development Index[3], tend to create great media attention, often forcing the political establishment to react.

The following quote taken from the *The Economist*, 16 October 2008, shows well the use of diverse range of statistical products in the media:

> The 1.3m people of Mauritius love to prove famous people wrong. On independence from Britain in 1968, pundits such as a Nobel Prize-winning economist, James Meade, and a novelist, V. S. Naipaul, did not give much of a chance to this tiny, isolated Indian Ocean island *1 800 km (1 100 miles)* off the coast of east Africa. Its people depended on a sugar economy and enjoyed a *GDP per person of only USD 200*. Yet the island now boasts a *GDP per person of USD 7 000*, and very few of its people live in absolute poverty. It once again ranks first in the latest annual *Mo Ibrahim index*, which measures governance in Africa. And it bagged 24th spot in the *World Bank's* global ranking for *ease of doing business* — the only African country in the top 30, ahead of countries such as Germany and France. How does it pull it off? (*The Economist*, October 2008, emphasis added).

It is interesting to see how official demographic, geographic and national accounts (GDP) data are mixed, in the same breath, with composite indices created by international organisations. The two indices mentioned (Ease of Doing Business and Mo Ibrahim index of African governance), are composite indicators, and provide league tables that rank countries, regions or cities in a specific issue, providing a tool for 'naming and shaming' on specific issues, thus holding politicians accountable for the outcomes of policy. As discussed in the Pisani-Ferry and Sapir paper just mentioned 'peer pressure and benchmarking should be integral parts of the political proc-

[1] See The Economist web pages under 'Market and Data', 'Rankings'.
[2] http://www.weforum.org/en/initiatives/gcp/index.htm
[3] http://hdr.undp.org/en/humandev/

ess'; further the use of these indices benefit the democratic process, 'as it empowers national electorates to review the performance of their own governments and it helps focus the debate on key areas of underperformance'.

Analysis via composite indicators

In recent years, the use of composite indicators has increased tremendously. In 2005, a simple search in Google Scholar for 'composite indicators' would return little less than 1 000 entries. The same entry in June 2009 will give more than 4 000 entries. This provides an indication of the increasing interest, not only for policymaking, but also for research and analysis purposes.

Because of the capacity of composite indicators to summarise complex constructs in a single number, they allow for country performance comparison with other variables. This might provide insight into the relationship between variables. The OECD, for example, has shown that differences in product market regulation (measured via a composite indicator by the same name) can explain divergent productivity trends (Cotis and Duval, 2007). The relationship between progress and fertility is compared in the journal Nature using the Human Development Index (Myrskylä et al., 2009). In a similar fashion Richard Florida discussed the importance of the creative class for economic growth of regions using composite indices.

Composite indicator methodology

As discussed above a composite indicator can be used (and is in general used) for both analysis and advocacy (Saltelli, 2007). Thus it is not infrequent for these measures to be the subject of heated controversies (Saisana et al, 2005, 2007). The creator of one such measure is thus better advised to make sure that her work is of a sufficient quality to withstand scrutiny. A series of steps for the creation of a composite indicator have been proposed in an handbook by the OECD and the JRC (Nardo et al., 2008). This handbook is the result of five years of preparation, two round of consultation with OECD high level statistical committee and was finally endorsed March 2008.

The 10 steps involve:

- Step 1: Developing a theoretical framework: it is necessary to create a sound theoretical framework and definition of the phenomena that is to be measured. In the creation of the framework, it is recommended that all stakeholders interested in the results of the measurements have an active participation. This also involves the decision on the structure and subgroups of indicators, that is to say, the pillars of the index.

- Step 2: Selecting indicators: once the definition is agreed and established, with a framework and a clear structure, it is necessary to find the adequate variables (or indicators) that will compose the index. Availability of the data is necessarily a crucial aspect for selection. It will also be necessary to know the source of the data and its strengths and weaknesses. Also in this process, it will be desirable to have input from the relevant stakeholders, in order to establish a list variables which are relevant to them.
- Step 3: Imputation of missing data: it is likely that not all the data will be available for all the countries. There are different techniques of inputting missing values, but it is important to establish the reliability of the imputed missing data.
- Step 4: Multivariate analysis: once there is a full list of indicators, with no missing data, then it is necessary to establish a multivariate analysis of the indicators chosen to see how the different indicators relate to each other.
- Step 5: Normalisation of data: before it is possible to aggregate the data into a single measure, it is necessary to put each indicator in a similar metric. There are a range of different techniques for normalisation, each with its specific characteristics. These include: ranking, z scores, and min-max normalisation among others.
- Step 6: Weighting and aggregation: unless the weights are determined using statistical arguments, they are established by the proponents of the index, and imply value judgements. Through weights one determines the importance of each of the indicators and pillars that create the composite indicator. Weights and aggregation procedure together determine the 'model' on which the index is built.
- Step 7: Robustness and sensitivity: each of the previous steps required a series of decisions and techniques. Each of these choices is a source of uncertainty.
- Step 8: Back to the details: it is important that once a holistic picture is obtained, the analysis turns back to the individual components to understand the results. In other words, what are the indicators that have more importance for each of the units of analysis, how do they affect the overall index?
- Step 9: Association with other variables: this can be used to test the explanatory power of a composite. Simple cross-plots are often very informative in showing relationship between variables, especially at country level.
- Step 10: Presentation and dissemination: it is important that the composite indicator is presented adequately to the end-users, policymakers, media, and

other stakeholders. An organisation without the resources, the authority and the legitimacy to do this may have difficulty in developing a successful composite index.

The case of creativity

In the case of creativity, the conference addresses mainly steps 1 and 2. There is a need to build consensus on what creativity is and establish a framework that can allow for the choice of indicators. The conference and this publication, as the introduction showed, present a variety of indicators that could be considered for the creation of a composite indicator on creativity. However, as discussed during the conference, further debate is needed on the importance of the different aspects of creativity and their relationship. Further, a composite indicator on creativity would have to prove that it can contribute to make policies and complement existing indicators in science and technology or innovation.

Richard Florida proposed a creativity index (2002, 2004). As it is often the case with composite indicators, the index has attracted a large amount of attention. The index has inspired politicians to change urban policies. Media have also been sensitive to the issue, especially because of the inclusion of the controversial 'gay index' to measure aspects of tolerance. Finally, the index has created a wide debate in academic circles, both in economics and urban development (Glaeser 2004, Peck, 2005, Hoyman and Faricy, 2009). The structure and composition of Florida's Creativity index is given in Table 1.

Florida underlined the importance of numbers to provide advocacy for his arguments — although he also maintained that '(…) in retrospect, I probably could have written this book using no statistics at all'. Numbers, he maintains, are important because they show relative magnitude to the phenomena in focus and can confirm or disprove the assumptions put forward. In addition, numbers can 'point towards connections you did not see before' (Florida, 2004, 327). He was also aware of the need for a sound methodology: 'None of these benefits accrue, however, unless the numbers themselves are sound and carefully derived'.

Florida's Creativity Index is a 'composite measure that is based on four indices for the most current year available: the Innovation Index, High-Tech Index, the Gay Index and the Creative Class' (Florida, 2004, 334). Table 1 presents the different indexes used for Florida to create the Creativity Index.

Table 1: Richard Florida's Creativity Index

Technology	Innovation Index:		Patented innovation per capita (version 2002), Average annual patent growth from 1990 to 1999 (version 2004)
	High-Tech Index	Developed by DeVol et al. (1999)	Metropolitan high-tech industrial output as a percentage of total US high-tech industrial output
			Percentage of region's own total economic output that comes from high-tech industries compared to national percentage
Tolerance	Gay Index	Developed by Graves et al. (2000)	Fraction of all US gay people who live in a given metropolitan area divided by the fraction of the total US population who live in that area
	Bohemian Index		Fraction of all artistically creative people (includes authors, designers, musicians, composers, actors, directors, painters, sculptors, artist, printmakers, photographers, dancers, artists, and performers) who live in a given metropolitan area divided by the fraction of the total US population who live in that area
	Racial Integration Index	(version 2004 only)	Census track ethnicity composition in relation to the composition of the whole MSA (metropolitan statistical area)
Talent	Creative Class Index		Percentage of creative occupations of total employed

Florida does not explicitly present his normalisation technique, nor how exactly the composite was formed. Thus, it is not completely clear what method was used for normalisation in order to combine different indices. An analysis of his work would seem to imply that a 'ranking' normalisation techniques is used (see Nardo et al., 2008). Florida gives the best performing region the maximum scored for each of the indexes: innovation, technology, tolerance and talent; and then adds the different scores using equal weighting. Regions are then ranked according to their total scored. In the 2004 edition, the Creativity Index is expressed as a percentage of the maximum possible score of the groups being compared (in the 2004 edition, 268 regions in the US). That is to say, a perfect score in all three Ts (a region being first in all indexes of the 268 regions) would have a score of 1.

Florida tries to present the indexes and its composition in a clear, transparent manner, short of technicalities. Unfortunately, this makes it very complicated to understand how the composites where formed. Further, he failed to evaluate the consequences of any of the assumptions he made. For example, a different normalisation technique could have given a different result, because of a different treatment of outliers. In a similar way, different aggregation methods could have led to different re-

sults. In other words, without analysing the sources of uncertainness it is not possible to conclude how robust his index is, and how useful it might be for policy purposes. The fact that a region ranks high in the three indexes would have to be tested against the assumptions made when constructing the different indexes.

Despite these technical concerns, the index is and has been used to support policy, especially urban policies (see e.g. Peck, 2005). Without a clear understanding of how robust the indexes developed are, most of the conclusions drawn from the indexes cannot be properly supported, and decisions might be taken in an inadequate direction, or with unexpected consequences.

A European measure of creativity should, without doubt, follow the steps described in the handbook discussed above and use, as much as possible, consensus for its construction and determination of components.

It's another story as to whether this index is useful for Europe. As argued elsewhere in this workshop different existing measures of human and social capital might already capture relevant measures of creativity.

References

Hoskins, B., Villalba, E., Van Nijlen, D. and Barber, C. (2008). *Measuring civic competence in Europe: A composite Indicator based on IEA Civic Education Study 1999 for 14 years old in School.* JRC Scientific and Technical Reports 23210 EN. Luxembourg: OPOCE.

Glaeser, G. L. (2004). Review of Richard Florida's The rise of the creative class. Retrieved last August 2008, from: http://www.creativeclass.org.

Hoyman, M. and Faricy, C. (2009). It Takes a Village: A Test of the Creative Class, Social Capital and Human Capital Theories, *Urban Affairs Review*, 44 (3), 311-333.

Cotis, J. P. and Duval, R. (2007). Competitiveness, Innovation and economic growth, *Second OECD World Forum on "Statistics, Knowledge and Policy". Istanbul, Turkey, 27-30 June 2007*. Retrieved last, August 2009 from: http://www.oecd.org/dataoecd/20/52/38859413.pdf

Myrskylä, M., Kohler, H.-P., and Billari, F.C. (2009). Advances in development reverse fertility declines, *Nature*, 460(6), 741-743

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

Aghion, P. , Dewatripont, M., Hoxby, C. and Sapir, A. (2008). Higher aspirations: An agenda for reforming European universities, *Bruegel Blueprint Series N.5*, 2008. Brussels: Bruegel.

Peck, J. (2005). Struggling with the creative class, *International Journal of Urban and Regional Research*, 29 (4), 740-770.

Pisani Ferry, J. and Sapir, A. (2006). Last exit to Lisbon. Bruegel, Retrived last August 2009 from: http://www.bruegel.org/

Saisana, M., D'Hombres, B., Saltelli, A. (Forthcoming). Rickety Numbers: Volatility of university rankings and policy implications. Submitted to *Research Policy*.

Saltelli, A. (2007). Composite indicators between analysis and advocacy, *Social Indicators Research*, 81(1), 65-77.

Wilkinson, R. and Pickett, K. (2009). *The Spirit Level, Why More Equal Societies Almost Always Do Better*. London: Allen Lane Publisher.