THE ANALYSIS OF METHODS FOR CONSTRUCTING COMPOSITE INDICATORS

OLEKSANDR Chernyak[1], MARINA Shumayeva[2]

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Abstract

The relevance of research caused by global informatisation and active formation of open information society, through which are created entirely new conditions for the functioning of the global economy. Article summaries different approaches and algorithm for constructing composite indicators that are used to characterize the technological state of information and telecommunication infrastructure in the country.

Keywords: informational society, information and telecommunication technologies, composite indicators, sub-indicators.

JEL classification: B41, C02

1. Introduction

The revolution in information and communication technologies became the reason of paying more attention to reliable and statistical data for evaluation of new dimensions and directions of the development of information society. Special attention is paid to the unequal distribution of information-communication technologies (ICT) between developed countries and in the society. In order to make an objective assessment of country`s status in the information society special systems of indicators are used that characterize the technological state of information and telecommunication infrastructure in the country.

[1] Professor / Ph.D., Faculty of Economics/Department of Economic Cybernetics, chernyak@univ.kiev.ua

[2] Ph.D. student, Faculty of Economics/Department of Economic Cybernetics, marina.netesina@gmail.com

To implement the study of the Information Society World Summit on the Information Society was held in 2003 in Geneva and in 2005 in Tunis. There were allocated main organizations for providing ICT static data and developing an effective methodology for measuring informatisation of countries. The organizations that deal with statistical studies about the information society are: International Telecommunication Union, Eurostat, United Nations Conference on Trade and Development, UNESCO Institute for Statistics, World Bank, etc. This article deals with the concept and algorithm for constructing composite indicators, methods and techniques of sub-indicators systematization, normalization and aggregation. Article summaries different approaches and algorithm for constructing composite indicator and gives advises in which case which method should be used in order not to have false result.

2. Method for constructing composite indicator

A mathematical combination (aggregation) of some set of indicators is called the index or consolidated (composite indicator). Composite indicators are used to compare the performance of countries in various fields such as competitiveness, globalization, innovation. Instead of using a sample of individual indicators, composite indicators are aggregated to optimize the analysis (Bahovec, 2008). There are different types of composite indicators. The most popular are the list:

- Digital Opportunity Index, (DOI);
- Network Readiness Index, (NRI);
- Informational Society Index, (ISI);
- Digital Access index, (DAI);
- Digital Divide Index, (DDI);
- ICI Diffusion Index, (ICTDI).

Composite indicators make it possible to integrate large amounts of information in an easily understandable format for further evaluation. Different scientific organizations have made various models and approaches for constructing composite indicators and each country tries to have its own national mythology of composite indicators building with all the specification of country development.

Using of wrong mythology in building indicators could lead to the wrong systematization of sub-indicators and choosing the type of

normalization. The general scheme of the steps to build an appropriate composite indicator:

2.1. The development theoretical foundation

Composite indicators are usually used to summarize a number of basic individual indicators. A separate indicator – is a quantitative or qualitative assessment of certain factors obtained through series of observations. For example, the proportion of Internet users, the level of IT use in specific areas, etc.

To form a composite indicator the set sub-indicators is determined which characterize the system with highest possible completeness (Freudenberg, 2003). Basically there are three levels of hierarchy to form a composite indicator:

- individual set of indicators, which is a sample of some statistics;
- thematic set of indicators (sub-indicators) – this individual indicators grouped together under the same feature for the most complete description of the scope of the study;
- composite indicators are formed by drawing up sub-indexes a single synthetic index.

A typical composite indicator can be expressed by the following equation:

$$I = \sum_{i-1}^{n} w_i X_i \ , \qquad (1)$$

where

I – composite indicator;

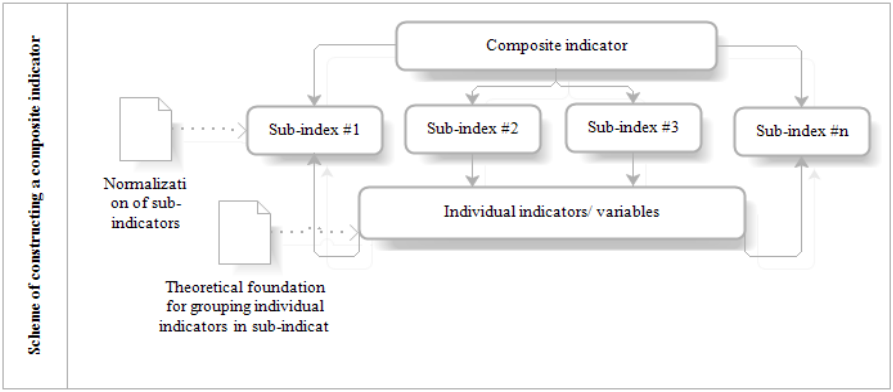
X_i – normalized sub-indicator;

w_i – weight of X_i;

$$\sum_{i-1}^{n} w_i = 1 \text{ and } 0 \le w_i \le 1 \qquad (2)$$

In practice it is extremely difficult to integrate separate variables so that they reflect the economic reality. The theoretical foundation is necessary to understand the logic of forming a set of individual indicators with consideration of individual weight of sub-indicators for composing an effective composite indicator.

Figure 1: Scheme of constricting a composite indicator



Source: compiled by the authors.

The theoretical basis is needed to combine individual indicators into a meaningful system and provide a foundation for the selection components and weights in the above formula. Ideally these data selected, combines and weighted in a appropriate way should reflect the size and structure of the phenomena that is measured.

2.2. The determination of individual indicators

Strengths and weaknesses of composite indicators largely derived from the quality of the basic variables. Ideally, variables should be selected based on their relevance, analytical validity, timeliness, accessibility, etc.

2.3. Identification and assessment of relevant variables

Variables come in a variety of statistical units have different ranges or weights. Variables should be standardized or normalized before they are aggregated to a composite index to avoid problems with mixed units. The variables are normalized to avoid domination of extreme values and problem with poor-quality database. Studies have shown that generally values that have a large gap from the average or normal range, most likely are the consequence of low quality database. If some variables have highly crossed distribution, they can be aligned through logarithmic transformation, and data can be reduced if there are an extremes. For each composite indicator special

method of normalization is applied. The most popular and effective methods of normalizations where chosen:

1) Ranking.

This is the simplest method of normalization. External factors in this method are not taken into account. The important factor is only the serial number of the sub-indicator. Such type of normalization leads to the considerable loss of information.

$$I_{qc}^{t} = Rank(x_{qc}^{t}) \tag{3}$$

2) Standardization.

This method converts values to a common scale with a mean value of "0" and a standard deviation of "1". This way indicators of extreme values have a greater effect on the composite indicator. This effect may be further corrected by the aggregation of indicators by including a sample of the best and worst values of individual indicators or by assigning differential weights (Nardo, 2005).

$$I_{qc}^{t} = \frac{x_{qc}^{t} - x_{qc=\bar{c}}^{t}}{\sigma_{qc=\bar{c}}^{t}} \tag{4}$$

where σ_{qc}^{t} – standard deviation.

3) Mini-max.

This method normalizes values to have identical range [0, 1] by subtracting the minimum value of the indicator and dividing it on the difference of maximum and minimum values. However, the disadvantage of this method is that extreme values can distort the transformed indicator. The advantage of this method is that the mini-max normalization can expand the range of indicators that lie within a small interval increasing influence on the composite indicator greater than in standardization.

$$I_{qc}^{t} = \frac{x_{qc}^{t} - \min_{c}(x_{q}^{t_0})}{\max_{c}(x_{q}^{t_0}) - \min_{c}(x_{q}^{t_0})} \tag{5}$$

4) Distance to reference.

The method normalizes the index by measuring the position of the indicator according to the reference point. This point can mean a goal that must be reached within a specified period of time. Point of reference can also be an external standard. For example the United States and Japan are often used as a point of reference for the normalization of composite indicators. In addition,

the reference point can be an average country within a group that is set to 1, while other countries will receive points based on their distance to the mean. In this way standardized indicators that are greater than "1" indicate the country to the level above average. The point of reference might be the country's leader with the group, but this approach is based on the extreme values that might be wrong.

$$I_{qc}^{t} = \frac{x_{qc}^{t}}{x_{qc=\bar{c}}^{t_0}} \tag{6}$$

or

$$I_{qc}^{t} = \frac{x_{qc}^{t} - x_{qc=\bar{c}}^{t_0}}{x_{qc=\bar{c}}^{t_0}} \tag{7}$$

5) Categorical scale.

Based on the categorical scale, quantitative or qualitative, to each indicator seatrain amount of scores is assigned. For example, given the best results, the first 5% receive 100 points, the data averages from 95 to 85 percent receive 70 points, etc.

$$I_{qc}^{t} = \begin{cases} 0, x_{qc}^{t} < P^{15} \\ 20, P^{15} \le x_{qc}^{t} < P^{25} \\ 40, P^{25} \le x_{qc}^{t} < P^{65} \\ 60, P^{65} \le x_{qc}^{t} < P^{85} \\ 80, P^{85} \le x_{qc}^{t} < P^{95} \\ 100, P^{95} \le x_{qc}^{t} \end{cases} \tag{8}$$

6) Indicators above or below average.

In this method all indicators are divided into tree parts. Indicators that are in the middle of the sample are assigned to "0", while all indicators above and below a certain threshold – are assigned to "1" and "-1".

This normalization method is simple and does not depend on emissions. However, the threshold level for different countries may be different, which leads to the omission of information.

$$I_{qc}^{t} = \begin{cases} 1, w > (1+p) \\ 0, (1-p) \le w \le (1+p) \\ -1, w < (1-p) \end{cases} \quad (9)$$

Marking for all types of normalization:

$\bar{c}$ – reference point;

x_{qc}^{t} – the value of sub-indicator q for country c in time t.

$P^{i} - i$ percent of distribution of sub-indicator;

σ_{qc}^{t} – standard deviation;

w_q – weight normalized sub indicator.

Each method has its advantages and disadvantages. The most commonly used the method of standardization, since all variables are converted to a common scale, which provides a normal distribution. In other approaches, for example, the distance to the reference, the technique is based on the extreme values that can be actually emissions. Method of categorical scales has a large degree of subjectivity because the thresholds are determined arbitrarily.

2.4. Weighting variables and group variables.

The choice of weighting coefficients is the most difficult stage of index construction. In many cases, using a simple average of the normalized sub- indicators can lead to unstable and unreliable index values. When choosing a method of aggregation is necessary to consider the amount sub-indicator, difference sets of values, and the correlation between them (Bahovec, 2008). To determine the weight coefficients the following methods or combination of methods should be used:

- Factor analysis;
- Analysis of coverage data;
- Analysis of hierarchical processes;
- Multiple linear regression;
- Expert opinion;
- The distance to the target;
- Neutralization of correlation effects.

2.5.Test for sensitivity and reliability of aggregated variables

Composite indicator results mostly depend on the choice of the method of weighing, standardization and aggregating variables. Test sensitivity makes it possible to analyze the appropriateness of inclusion or exclusion of different variables of the composite indicator, changes in weight or method of standardization and selection of alternative baseline years.

3. Conclusions

Analysis of building composite indexes proved that they have significant differences both in the conceptual model of the information society and in the choice of base indicators than are grouped to assess various areas of the information society. Moreover, the choice of priorities for the evaluation depends not only on the goals of the research object but also on the subjective definition of the role and place of an element of the information society. Therefore, it is clear that none of the established international indicators can not be directly transferred to assessments of development of information society in a particular country, for example of Ukraine. An individual model with specific methods should be build for analyzing trend of the IT structure of the investigated countries.

4. References

- Bahovec, O. B. (2008) *Preconditions of formation and information society*. Kyiv: Azimut.
- Freudenberg, M. (2003) *Composite indicators of country performance: a critical assessment.* Paris: OECD Publishing.
- JRC (2005) Handbook on constructing composite indicators: methodology and user guide, *Joint Research Centre, Belgium.*
- Nardo, M. (2005) *Tools for Composite Indicators Building*. Ispra.
- *Report on the World Summit on the Information Society* (2006). Kyiv: Zvyazok.