TOWARDS AN OPTIMAL SYNTHETIC INDEX FOR MEASURING GENDER INEQUALITY

Irene Riobóo
Rey Juan Carlos University
Department of Applied Economics I
Paseo Artilleros s/n. 28032 Madrid, Spain
E-mail: irene.rioboo@urjc.es

José María Riobóo
University of Santiago de Compostela
Department of Quantitative Methods
El Burgo s/n. 15782 Santiago, Spain
E-mail: josem.rioboo@usc.es

KEYWORDS
Synthetic indicators, Gender, Inequality, Modelling.

ABSTRACT

Gender equality is a key aspect of human rights in order to achieve sustainability and development focused on people. Bearing in mind how important equality has become in decision making, indicators including the gender approach are useful for quantifying the differences between the two genders and how they have evolved over time.

Taking gender into account brings a diverse and complex reality into play which, in order to analyse accurately, entails distinguishing between various approaches and technical alternatives. This paper presents modelling, understood as creating a simplified recreation of a complex reality, as the most suitable framework.

A thorough analysis of recent literature reveals serious limitations, thus leading to the need to further research in areas such as the adaptation of empirical indicators to the theoretical framework, the expression applied when calculating simple indicators, the process used when rescaling them for normalisation purposes, the average used to obtain a synthetic index and the methods used to assign weights to integrated simple indicators. We have focused our research on these aspects, analysing and contrasting different alternatives, particularly focusing on ascertaining the most accurate methodology to attain a synthetic gender indicator, as a way of analysing these phenomena.

INTRODUCTION

Analysing gender equality is a very relevant subject in human development. In this sense, the Platform for Action of the IV World Conference on Women, which took place in Beijing in 1995, assigns a strategic objective to the need to gather and disseminate information broken down by gender and specific to the realities of women. In addition, it is worth highlighting that the third of the Millennium Development Goals (MDG), an international agreement signed by 189 countries in September 2000 and seconded by the United Nations, is "to promote gender equality and empower women". Consequently, tools to measure gender are essential, implying indicators that focus particularly on analysing the relative position of men and women.

This is the specific function of gender indicators as well as to detect the social changes both genders experience at different moments in time. More specifically, the use of synthetic indices involves aggregating several partial indices in order to carry out more aggregated analyses. In the process of obtaining them, there are very many occasions in which the researcher has to adopt resolutions, implying choosing from several suitable alternatives. In this sense, the simulation and development of alternative scenarios will allow us to obtain optimal results.

Having specified the target, we need to accurately establish the variables to be considered in accordance with their importance and data availability. The fact that various techniques and models can be used to select variables allows us to determine alternative scenarios, which at the same time are open to analysis, in order to establish a final set of variables.

In a second stage, criteria are defined to group the selected variables in order to determine the strata on which to obtain the synthetic indicators. The variety of options available when grouping demands the application of several techniques and models in order to determine the most suitable strata in terms of number and composition.

Prior to obtaining these indices, different alternatives are required for the variables under consideration and they must be measured in equivalent terms. This in turn requires research into various rescaling methods and their implications.

Finally, the paper discusses the application of the extensive and diverse Theory of Index Numbers to the topic this paper is concerned with. Aggregation, averaging and weighting methods provide us with a complex scenario for which various research alternatives are available and where the optimality sought after will depend on the properties previously deemed desirable for the indices to be obtained.

Proceedings 23rd European Conference on Modelling and Simulation ©ECMS Javier Otamendi, Andrzej Bargiela, José Luis Montes, Luis Miguel Doncel Pedrera (Editors)
ISBN: 978-0-9553018-8-9 / ISBN: 978-0-9553018-9-6 (CD)

This paper is structured as follows: Firstly, we present the aims of this research. Next, we make reference to the statistical sources that we have used and list them by topic. The next section details the methodology used to obtain simple gender indicators for all the variables considered. We then analyse several rescaling alternatives for these indices, following which we compare the different methods to obtain the synthetic indices of gender equality through the aggregation of partial indices that integrate the areas considered. In the next section results are presented along with the most significant conclusions drawn. The last section includes a series of bibliographical references in the field of gender indicators.

OBJECTIVE

The general objective of this research is to obtain indicators that facilitate the incorporation of the gender approach in the production of statistics as an important aspect of including gender equality issues. We therefore aim to further the research into determining a suitable methodology for obtaining a system of gender indicators that provides us with a real picture of each gender's roles, structures and needs, as well as verifying the compliance and effects of the agreements signed and activities undertaken respectively on a regional, national and international scale to foster equality.

INFORMATION SOURCES AND VARIABLES

The methodological contributions this paper makes are also applied in an empirical analysis of data referring to Spain. In order to do so, various official statistical sources are used that allowed us to periodically access reliable information broken down by gender. This information is essential for establishing the proportion of men and women in each of the different socioeconomic categories. The main shortcoming in this sense is that this breakdown usually appears for macro type data, which implies that we are missing information as we go deeper into more specific levels or when data gets mixed up with other variables.

This research takes into account annual data on 36 variables broken down by gender between 2000-2007. In order to obtain objective gender indicators, this selection, following criteria based on data availability, reliability and significance, only includes variables that are clear "positive" or "negative" in terms of gender equality. Hence, the most confusing, such as demography, nationality, religion, etc, have been excluded. However, it must be said that for comparative purposes, this set of variables is expressed in relative terms.

Once selected, the variables are grouped by dimensions of social concern, determined according to classifications of the situation of men and women used by international organisations, such as the United Nations (Gender Statistics Base) and the Organisation for Economic Co-operation and Development (Gender, Institutions and Development Database): Power and decision making, Social assistance, Education, Activity and Healthcare.

PARTIAL GENDER INDICATORS

The choice of partial indicators depends on the pre-established goal of constructing a synthetic index, and implies multiple criteria based on a general consensus regarding the suitability of a set of desirable characteristics that these indicators should have. These characteristics cover aspects as: *suitability,* that is, the degree to which the indicators fit the theoretical reality they are representing: *accessibility*, how readily available they are; *comprehensibility* or how easy they are to interpret; and *comparability* in terms of both time and space.

Indicators of gender equality, which are based on the existence of a gender role, aim to establish the relative situation of men and women and the changes that have occurred at different moments in time. The idea is to measure aspects linked to equality and inequality of gender, where such aspects can be measured, quantified and systemised. Therefore, gender indicators are tools used to understand social processes and provide informed and equitable options to maximise the positive impact of an intervention in this direction.

.

As there is such a wide variety of indicators, researchers must choose the most relevant and those which can be measured, always bearing in mind their ability to be compared to indicators elaborated in other fields and over time, whereby chronological series make it possible to study variable trends over time. Likewise, researchers should choose the indicators that are most common and easily transferrable from one type of action to another.

However, prior to this we must choose the type of variable to be used, that is, whether to employ levels, variations or growth rates, to name a few, as well as the type of data, that is, the original series, the cycle trend or the seasonally adjusted series, etc.

This paper aims to compare the indicators most commonly used to analyse gender equality, from which we have chosen the two below due to they are relative indicators, omitting absolute indicators which limit our analysis due to not providing points of reference:

$$G1_i = \frac{W_i}{M_i} * 100 \qquad (1)$$

$$G2_i = \frac{W_i}{T_i} * 100 \qquad (2)$$

where we calculate, for the *ith* observation, the percentage that represents the variable of our study referring to women (W_t) over men (M_t) and over the total of both genders (T_t), respectively. We should mention here that the *G2* index is also known as the feminisation index.

We explain both partial indices in the same way. In both cases gender equality would occur when the indicator scores 100, but if the value is below this figure, women would be in inferiority in relative terms, which could be either positive or negative depending on the connotations of variable in question.

NORMALISATION OF PARTIAL INDICATORS

The procedure used to build synthetic indicators for each of the dimensions considered comes from normalising simple indices obtained from the variables of each area. In this way we can eliminate the distorted effects of different scales (Mardia et al. 1979). There are several ways that enable us to normalise these indices (Freudenber 2003; Jacobs et al. 2004) when confronting the normalisation of scale values and the possibility of biases. The main methods to obtain these standard variables, $\widetilde{X}_{ij}$, that is to say, variables in a common scale, are as follows:

1. Standardisation:

$$\widetilde{X}_{ij} = \frac{X_{ij} - \overline{X}_j}{\sigma_j} \qquad (3)$$

where p is the number for simple indicators $I_1, I_2, \dots, I_p$ with values for a set of n observations; X_{ij} is the value of the *ith* observation for *jth* indicator, with $1 \leq i \leq n$ y $1 \leq j \leq p$; being $\overline{X}_j$ the arithmetic average for the *jth* indicator and σ_j its standard deviation.

2. Ratio or percentage differences from the mean:

$$\widetilde{X}_{ij} = \frac{X_{ij}}{\overline{X}_j} \qquad (4)$$

3. Ratio or percentage differences over periods:

$$\widetilde{X}_{ij} = \frac{X_{ij}^{t} - X_{ij}^{t-h}}{X_{ij}^{t-h}} \qquad (5)$$

4. Lineal rescaling techniques, highlighted among them:

a) Minimal rescaling:

$$\widetilde{X}_{ij} = \frac{X_{ij} - X_{\bullet j}}{X_j^{\bullet} - X_{\bullet j}} * 10 \qquad (6)$$

where $X_{\bullet j}$ is the worst value according to the *jth* indicator and $X_j^{\bullet}$ is the best one.

b) Maximal rescaling:

$$\widetilde{X}_{ij} = \frac{X_j^{\bullet} - X_{ij}}{X_j^{\bullet} - X_{\bullet j}} * 10 \qquad (7)$$

Our work focuses on comparative results obtained from applying the standardisation and minimal rescaling methods, as they are the predominant alternatives at present.

When this point is reached, it is important to clarify the best and worst values of the indicator, references necessary to carry out minimal rescaling (Drewnowski 1972). The definition of these values is based on the definition of the synthetic index we are aiming to calculate. As this research aims to find an indicator to measure how the discrimination of women has evolved, the positive or negative connotation of each variable must be taken into account in order to discern the values that represent a better situation for women. Therefore, if the variable has positive connotations, the higher the value of the indicator the better and the less discrimination, whereas if the variable has negative connotations, the higher the indicator the higher the level of discrimination of women is.

Matters could be made easier by considering only the variables that associate high simple index values to more positive circumstances for women, that is, less discrimination, rather than those that are associated to lower values, but this is not the case in this research. Another alternative to simplify the method above would be to change the sign of simple indices values belong to variables with negative connotations. As per Hair et al. (1999), this inverse scoring method is more suitable when a factorial analysis is to be undertaken as it avoids the possible effects of variables offsetting one another when we really should consider that they both point in the same direction when it comes to the final objective being measured. This paper adopts this procedure and, as a result we do not need to identify the highest and lowest values scored by women for each and every variable. In all cases, the best score will be the highest value of the index and vice-versa.

Bearing these considerations in mind, the normalised index will make it possible to analyse the evolution of gender equality over time, assigning a value of 10 when the situation for women was at its best and 0 should the opposite apply.

If we opt for standardisation as a method to calculate normalised variables we turn simple indicators into a standard scale with a zero average and standard deviation one, which also allow us to add them up properly (World Economic Forum 1996). In this case, the interpretation of standardised simple indices would be the same as that for minimal rescaling, the higher the indicator the better the circumstances are for women.

The greatest drawback of this method would be in the cases where data distribution is not centred around its mean, because the presence of outliers will affect results substantially (McGranahan et al. 1972). Nevertheless, this drawback may be overcome by applying detection outlier techniques or also by applying winsorized and trimmed means.

It is worth highlighting that unlike simple indicators, which captured differences between genders, both normalised and standardised indicators and also the synthetic indicators calculated using them as a basis, will only provide information about the evolution of the discrimination of women.

Although the normalisation methods mentioned above are the usual practice, we believe future research should test alternative methods of short-term stable normalisation to overcome the shortcoming of periodical redefinition.

SYNTHETIC INDICATORS

Once we have established a partial indicator system, that is to say, the set of the integrated and systematic indices related to gender equality that allow us to quantify the variety of aspects this topic involves, a synthetic index must be obtained for each dimension along with a final synthetic index. In order to do so, we need to make decisions over the two of the most relevant aspects: how to determine the weights and averages we are going to use.

Determining weights

In the process used to add up simple indices, we need to address the subject of weighting, due to the fact that one of them might have different significance in the aggregated index. The systems used the most to assign weights could be divided into four groups:

-- Factorial Methods, Principal Component Analysis being the most popular. Used by several authors to set up indicators (Ram 1982 and Slotje et al. 1991, amongst others), this method aims to explain most of the total variability in a set of variables using the least possible number of components, defining these as a new category of independent variables that are lineal functions of the original ones. When applying this method we are able to consider the normalised partial indices through the factorial charges in the first principal component.

-- Equal or unitary weights. Using standardised variation rates as partial indices allow for symmetrical treatment as a result of the equal variance inherent to standardisation. This method can be considered suitable in the case of working with dimensions of equal size. The primary shortcoming of this method is that the choice of partial indicators must be highly filtered due to the sensitivity of results to any deviation.

-- Use of quantifiable reference variables. In this case, weights can be determined on the basis of the relevancy of the indices according to each magnitude. We can distinguish three different alternatives: a) based on the involvement of each partial index in the aggregate of reference, b) based on the simple or partial correlation between partial indices and reference variable, c) based on the results of a model of lineal programming, such as DEA (Data Envelopment Analysis) initially proposed by Charnes, Cooper and Rhodes (1978) and applied by Zhu (2001) to establish weights when finding synthetic indices.

-- Qualitative weights. In this scenario weights are assigned according to subjective criteria resulting from the knowledge of the phenomenon we are considering. The most commonly used subjective criterion assigns weights according to the degree in which they comply with the most desirable theoretical properties for the different indices mentioned above.

This classification does not include all the methods available for assigning weights, as specific and sometimes complementary alternatives exist, such as the use of “Budget Allocation” (Winterfeld 1986) and the “Hierarchy Process” (Saaty 1987).

Given that in general terms the choice of weighting system has an important impact on the value of the indicator, in our analysis diverse alternative models have been tested, focusing mainly on the first two criteria mentioned: Principal Component Analysis and equal weighting.

Aggregating partial indicators

Once the measurement units of the elementary indices are homogenised, we proceed to aggregate them in order to obtain the synthetic index. In order to do so, two groups of techniques are often used: additive and multiplicative techniques, depending on whether the index is obtained from operations derived from aggregating the different elementary indices (ex.: arithmetic average) or if it is obtained from operations derived from their product (ex.: geometric average).

When determining the most suitable aggregation procedure, we must bear in mind the characteristics of both the elementary indices to be grouped and also their units of measurement, such as the average and weights to be used. Any error in this sense could lead synthetic indices to be biased, in that they would not correctly reflect the basic information. It has been demonstrated that the so called “scale effect”, “average effect” and “weighting effect” interact when determining the value of the synthetic index, being able to go so far as to be compensating or to impel the index value in the wrong direction.

Under a scales criterion, we group the aggregation methods in three key blocks: 1. When we work with ordinal scales the aggregation can be obtained simply by summing up the ranges. 2. If the considered scale is an ordinal one, we can obtain the aggregation as the difference between the number of indicators that are over and under the average or an arbitrary deviation according to it. 3. If we have a ratio scale, as is often the case in this kind of studies, aggregation implies that we have to determine the average we are going to use, as well as the weights, bearing in mind all the implications derived from the previous normalisation process.

For the average of relative magnitudes (percentages, rates, indices, etc.) there is sufficient justification from a theoretical point of view that the geometric average is the most suitable. However, in practice the arithmetic average is the most commonly used, in spite of the fact that it provides higher values. Nevertheless, apart from its theoretical suitability, the average to be used will be conditioned by the values of the variables considered and by the rescaling or standardisation carried out, given that in order to use the geometric average correctly, only positive values can be used.

PRESENTATION OF THE RESULTS

Following what was said before, the procedure used to make a synthetic gender indicator starts with the obtaining of the individual indices, G1 and G2, for each of the variables that represent the areas of interest considered in this work. Subsequently we proceed to normalise the above mentioned partial indicators and later aggregate them by area. As a result, each will have a synthetic gender index. The same procedure is applied to obtain the general synthetic indicator, with the difference that the values to be added up are the five synthetic indicators calculated for the five areas taken into account.

Among all the alternatives handled we will comment on those that in the current circumstances have clearly proved to be better at reflecting the characteristics of the phenomenon analysed. Thus, the modelling carried out can be summarised by the scheme shown in Figure 1.

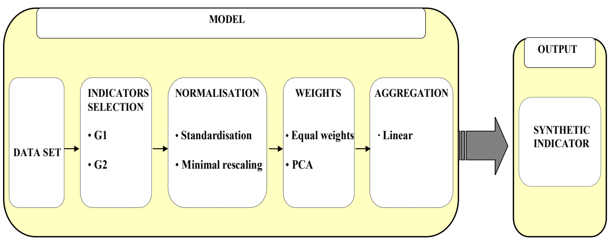

Figure 1: Description of the model

We will begin by discussing that which assigns the same weight to all partial indices. In this case, the composite indicator, both those referring to each area and also the general indicator, is obtained as a simple arithmetic average of the partial indices. As summary we show the graphs below (Figures 2-5), that represent the synthetic indices obtained as detailed previously, where not only is it possible to extract information about the general trend of every aggregate, but also about the dimensions that have the greatest influence on the evolution the overall index (G.I.).

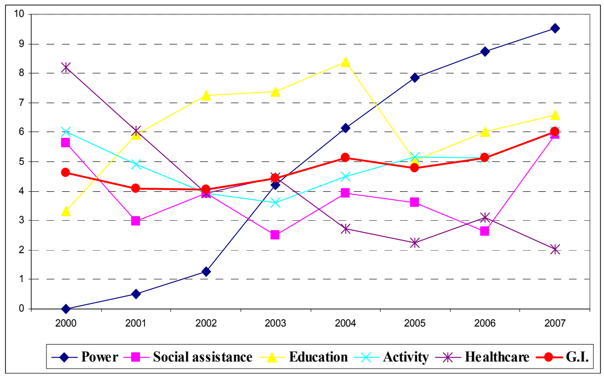

Figure 2: G1 normalised - unitary weights

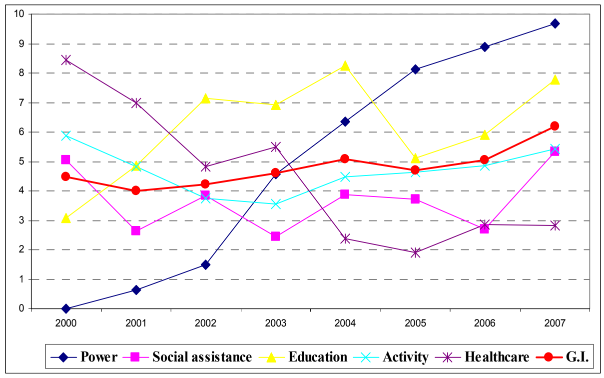

Figure 3: G2 normalised - unitary weights

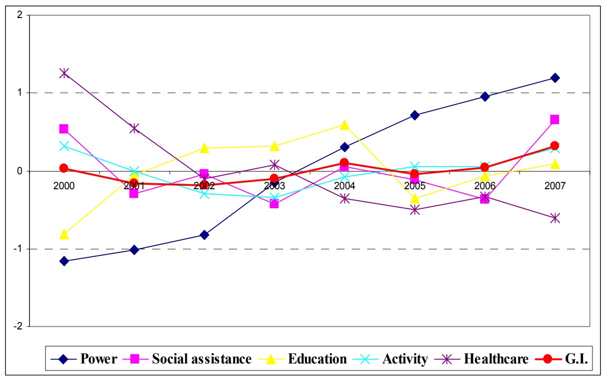

Figure 4: G1 standardised - unitary weights

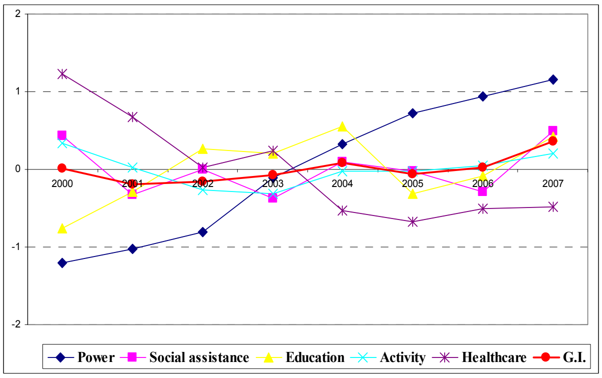

Figure 5: G2 standardised - unitary weights

Another alternative used to obtain the composite indicators has been to assign weights derived by Principal Component Analysis (PCA) to the various data series this indicator comprises. This method allows us to detect existing correlations between the normalised partial indices gathered in one dimension and the proper dimension in itself. These correlations are summed up in factors that provide a common variance and facilitate the weight of the involvement that each partial index has in the aggregate.

A very simple change of scale of the coefficients of the component matrix, making its sum equal to one, provides the weights used to calculate the aggregated indicators, obtained in this case as a weighted arithmetic average of the partial indicators. The synthetic indices for each dimension will be added up in the same way, giving each of them a weight provided again by PCA, obtaining thus the overall index. Here we introduce (Table 1) the different weights obtained for G1 and G2, as well as for each of the normalisation methods applied (minimal rescaling, N, and standardisation, S):

Table 1: Weightings obtained through PCA

	G1-N	G2-N	G1-S	G2-S
Power	0,27	0,27	0,19	0,20
Social assistance	0,18	0,13	0,16	0,14
Education	0,16	0,26	0,20	0,21
Activity	0,11	0,06	0,23	0,23
Healthcare	0,28	0,28	0,21	0,22

As we can see, there are significant differences here, mainly concerning the type of normalisation undertaken; hence the results obtained have been quite different, as Figures 6-9 show.

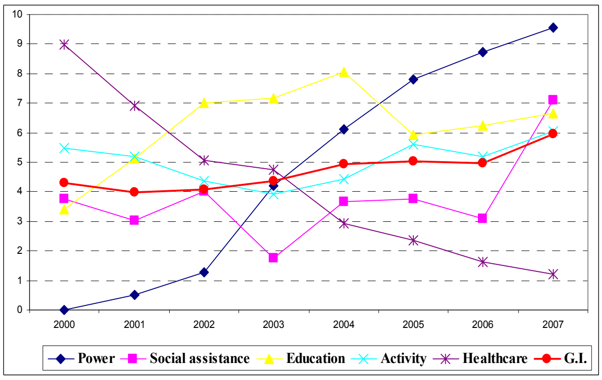

Figure 6: G1 normalised-PCA weights

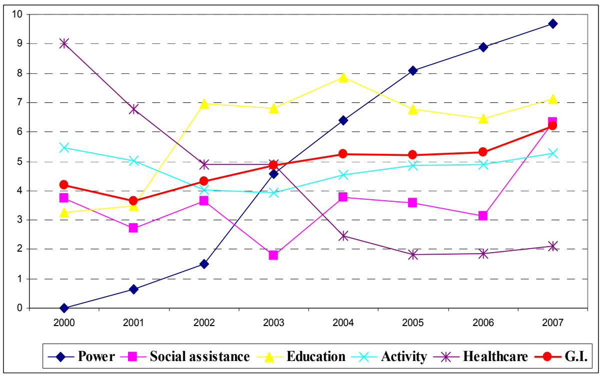

Figure 7: G2 normalised-PCA weights

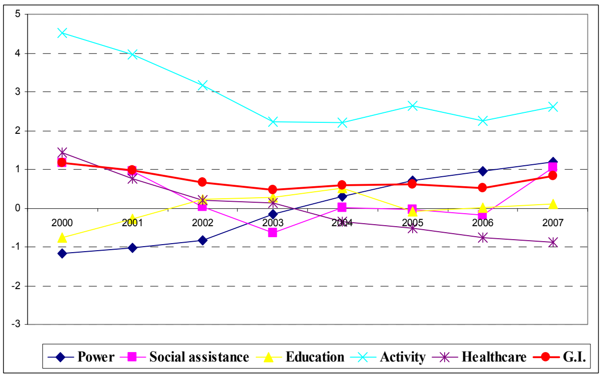

Figure 8: G1 standardised - PCA weights

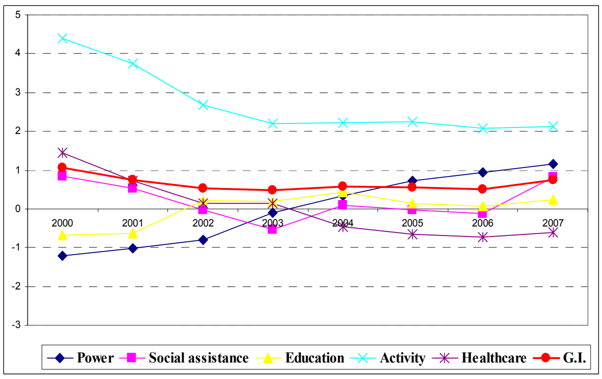

Figure 9: G2 standardised - PCA weights

From the figures included in this epigraph we can immediately and easily interpret the results obtained. In summary, we end up by saying that according to the choice of indices G1 and G2, results do not register significant differences. Nevertheless, the different procedures of normalisation and assignment of weights applied will lead us to very disparate results, both in the contribution of each dimension to the synthetic indices and also in the proper trend of these indicators. As regards minimal rescaling, significant differences are not observed between the weighting systems used, whereas in case of the standardisation these differences are quite marked. In this sense, we could say that the method based on minimal rescaling is more robust when dealing with outliers than the method based on standardisation, and it is also less dependent on the weightings of each indicator.

CONCLUSIONS AND FURTHER RESEARCHS

Gender indicators are important, not only to promote policies focused on overcoming the lack of parity and to review their impact, but also because gender equality plays a vital role in socioeconomic development.

In the model proposed to obtain these indicators, various dimensions measured on different scales must be combined, which means choosing from the various methods available to normalise, weight and aggregate data. Therefore, in light of the uncertainty associated to the model and the decisions taken, which are to a certain extent subjective, we believe it is worth carrying out a sensitivity and uncertainty analysis to capture extant plurality.

After identifying the main factors that affect the variability of the synthetic index, it is necessary to further the research that takes us to the quantification and classification of the above mentioned sources of change. A combined analysis of simulation, uncertainty and sensitivity, would allow us to propose a certain number of groups of acceptable alternative weightings along with a measurement of their permissible variation. All of this would contribute positively to the necessary consensus among professionals in this issue, in order to carry out homogeneous research that facilitates comparability.

REFERENCES

Drewnowski, J. 1972. “Social indicators and welfare measurement: remarks on methodology.” *Journal of Development Studies*, No. 8, 77–90.

EC Joint Research Centre. 2002. “State-of-the-Art Report on Current Methodologies and Practices for Composite Indicator Development.” Institute for the Protection and Security of the Citizen, Italy.

Frederik B. 2002. “An overview and evaluation of composite indices of development.” *Social Indicators Research*, No. 59, 115–151.

Freudenberg, M. 2003. *Composite indicators of country performance: a critical assessment*. OECD, Paris.

Hair, J.F.; R.E. Anderson; R.L. Tatham and W.C. Black. 1999. *Análisis Multivariante*. Prentice Hall Iberia, Madrid.

Jacobs, R.P.; P. Smith and M. Goddard. 2004. "Measuring performance: an examination of composite performance indicators.” Centre for Health Economics, Technical Paper Series 29.

Mardia, K.W.; J.T. Kent and J.M. Bibby. 1979. Multivariate *Analysis*. Academic Press, New York.

McGranahan, D.V.; C. Richard-Proust; N.V. Sovani and M. Subramanian. 1972. “Contents and Measurement of Socioeconomic Development.” UNRISD, New York.

OECD. 2005. *Handbook on constructing composite indicators: Methodology and user guide*.

Ram, R. 1982. “Composite indices of physical quality of life, basic needs fulfilment and income: a principal component representation.” *Journal of Development Economics*, No. 11-2, 227-247.

Saaty, R.W. 1987. “The Analytic Hierarchy Process-What it is and how it is used.” *Mathematical Modelling*, No. 9, 171-176.

Slottje, D.; G. Scully; J.G. Hirschberge and Hayes, K.J. 1991. “Measuring the quality of life across countries.” *Review of Economics & Statistics*, No. 73, 684-93.

Winterfeld, D. and W. Edwards. 1986. *Decision Analysis and Behavioural Research*. Cambridge University Press.

World Economic Forum. 1996. “Global Competitiveness Report.” WEF, Geneva.

AUTHOR BIOGRAPHIES

IRENE RIOBÓO went to the University of Santiago de Compostela (Spain), where she studied Economics and obtained her degree in 2000. She worked for the University of Castile-La Mancha for eight years where she received a Ph.D. in 2007. She is currently a professor of Statistics in the Department of Applied Economics at the Rey Juan Carlos University in Madrid. Her research interests include economic indicators, deflators and gender analysis.

JOSÉ MARÍA RIOBÓO went to the University of Granada (Spain), where he studied Political, Economic and Commercial Science and obtained his degree in 1971. He is currently a professor of Economic Statistics at the University of Santiago de Compostela where he started to work in 1972 and received a Ph.D. in 1984. He has participated in several research projects, his main research interests being economic indicators, regional accounts and input-output analysis.