Knowledge Economy Indicators

Workpackage 5

Applied Sensitivity Analysis of Composite Indicators with R

Deliverable 5.5

2008

List of contributors:

Michaela Saisana, Joint Research Centre, Ispra
Ralf Münnich, Jan Seger and Stefan Zins, University of Trier
Luis Huergo, University of Tübingen.

Main responsibility:

Michaela Saisana, Joint Research Centre, Ispra,
Luis Huergo, University of Tübingen.

CIS8–CT–2004–502529 KEI

The project is supported by European Commission by funding from the Sixth Framework Programme for Research.

http://europa.eu.int/comm/research/index_en.cfm

http://europa.eu.int/comm/research/fp6/ssp/kei_en.htm

http://www.cordis.lu/citizens/kick_off3.htm

http://kei.publicstatistics.net/

Preface

The aim of this deliverable is to demonstrate the sensitivity analysis module in action with an application to the Knowledge Economy dataset developed within the KEI project (cf. deliverables 2.3 and 2.5). The programmes in use are based on an R port of the original work from the ISPRA team (cf. deliverables D5.1 and 5.2) and on the multiple imputation module (cf. deliverable 3.2) of the KEI project.

The program developed in this deliverable executes a combined uncertainty and sensitivity analysis of the KEI composite indicator country scores and ranks that are built using the KEI framework (a list of about 116 indicators available between 2001 and 2004 for 24 European Union Member States, the USA and Japan). The analysis includes a plurality of scenarios in which different sources of uncertainty (related to the imputation, the normalization method, the exclusion of an indicator form the dataset and the weighting scheme) are activated simultaneously.

As a result of the uncertainty analysis, country scores for the KEI composite indicator are estimated in a Monte Carlo framework. Subsequently, a frequency matrix of the country ranks is calculated across the different simulations. Such a multi-modeling approach allows one to deal with the criticism, often made to composite indicators, that ranks are presented as if they were calculated under conditions of certainty while this is rarely the case due to the fact that the encoding process of building a composite indicator or a ranking system is fraught with uncertainties of different order (Saisana et al., 2005).

For the purposes of sensitivity analysis, the Sobol' method, which belongs to the class of the variance-based techniques, is used in order to obtain the most complete and general pattern of sensitivity of the country scores/ranks to the uncertainties in the development of the composite indicator. The program is based on an easy-to-code implementation offered in Saltelli et al. (2008), pp. 164-67, which provides all the pairs of first-order and total effect sensitivity measures. The first order sensitivity measures capture the direct impact of an input factor, whilst the total effect sensitivity measures capture the direct and indirect (due to interactions) impact of an input factor. The Sobol' method does not rely on any assumption about the linearity or the monotonic nature of the input-output mapping.

The deliverable is based on a presentation from Michaela Saisana and Luis Huergo given during the useR!2007 conference in Vienna. Luis Huergo was responsible for the R port. The Trier team contributed with a front-end and the final implementation of the sensitivity study within this deliverable. Michaela Saisana always took care of an adequate translation of the ISRPA work.

The authors would like to thank Arno Petri, University of Tübingen, for supporting the translation of the ISRPA programmes from Matlab to R.

Contents

Chapter 1

Introduction

The programme elaborated in this deliverable executes a sensitivity analysis on KEI data sets. The used data are explained in deliverables 2.3 and 2.5 (cf. Thees et al., 2008; Arundel and Hansen, 2008). In general sensitivity analysis is the study of how the variation in the output of a model can be apportioned to different sources of variation, and of how the given model depends upon the information fed into it. The traditional literature in the evaluation of models refers to probabilistic sensitivity analysis as the process of assigning probability distributions to uncertain factors, drawing from their distributions and subsequently analysing the resulting distribution of the output of interest. Among researchers on the field, this procedure is known as uncertainty analysis. Sensitivity analysis in a strong sense represents a step further towards an understanding of the causes of uncertainty by investigating how the output of a model can be apportioned to different sources of uncertainty in the model input, in this case the KEI data sets. A thorough overview of sensitivity analysis can be found in Saltelli et al. 2000 or Saltelli et al. 2008. The following details are based on Saisana et al. (2005), Nardo et al. (2005) and JRC/OECD (2008).

Some methods to portion the variability of the output to the different sources of variation are currently being used, for example the torpedo graph (cf. Vose, 2006). Although its easiness of implementation makes it attractive, it is only useful for linear models, since the correlation coefficient only works for linear relationships. The use of rank correlation coefficients to circumvent this problem helps only for monotonic models. It is thus desirable to have a method, which does not place such restrictions on a model, because it is sometimes difficult to know a priori whether a complex model behaves linearly or even monotonically.

The possibility to vary all factors at the same time within their range of variation is another desirable property of a sensitivity analysis method, which is done in the programme below. The proposed method, known as method of Sobol', relies on a high dimensional model representation of the model in question and the orthogonal decomposition of its variance, which gives rise to the Sobol' Sensitivity Measures (First Order and Total Effect). This method works also for non-monotonic models and can detect and deal with interactions among the factors. By mean of this method, the contribution to the variance of the output from every factor (alone and through interactions) can be calculated. Two factors are said to interact when their effect on the output cannot be expressed as a sum of their single effects. Interactions may imply, for instance, that extreme values of the output are uniquely associated with particular combinations of model inputs, in a way that is not described by the single effects. Interactions represent an important feature in composite indicators, and are more difficult to be detected and estimated than single effects. A possible drawback of the method arises from the assumption that the variance of a distribution is a correct measure of the uncertainty of an output. Although broadly accepted, this assumption is not free of controversy In this report, uncertainty analysis is used to estimate, under a plurality of methodological scenarios, the country scores and ranks in the Knowledge Economy Index for EU Member States, the USA and Japan. Sensitivity analysis is also used in order to decompose the variance of the country scores into the uncertainties in four main assumptions made during the development of the Knowledge Economy Index.

The main programme consists of several subprogrammes. Every single subprogramme has a clearly defined task to manage.

The sensitivity analysis conducted in the given context is based on four cornerstones:

1. `Sample Generation`
2. `Uncertain Inputs`
3. `Composite Indicator`
4. `Sensitivity Indices`

The aim of the following chapters is to explain the cornerstones themselves and also the corresponding interactions.

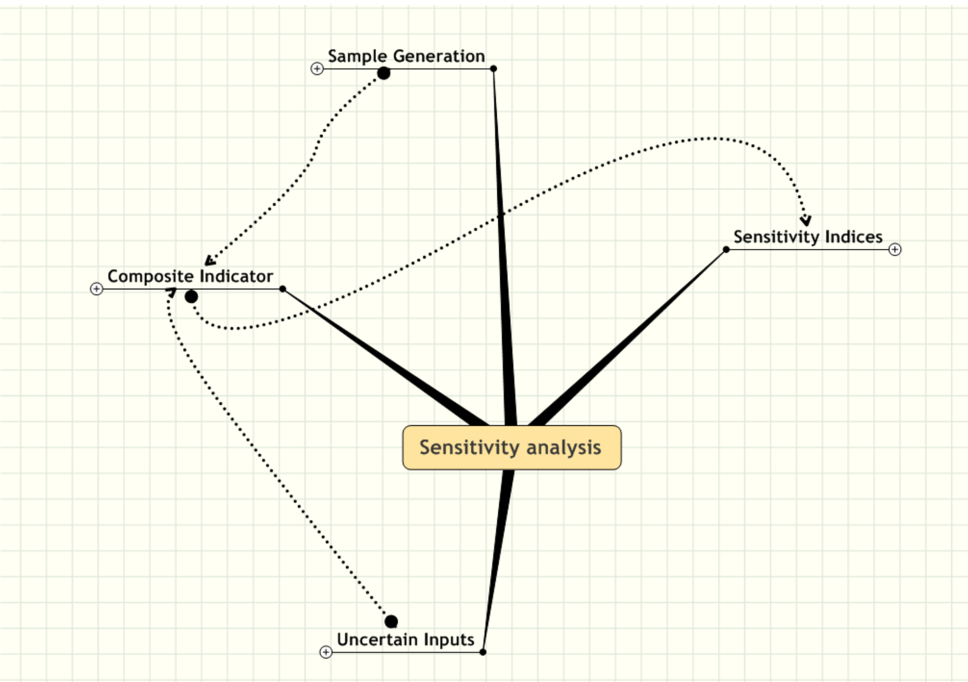

Figure 1.1: General structure of the sensitivity analysis

1.1 Sample Generation

The initial step for running the programme is to construct a decision matrix which contains random numbers. Hence, the decision matrix has a dimension of 10240 x 4. The dimension of the matrix will be explained in the next Section. This decision matrix is the key tool in creating uncertain inputs.

1.2 Generation of the random numbers

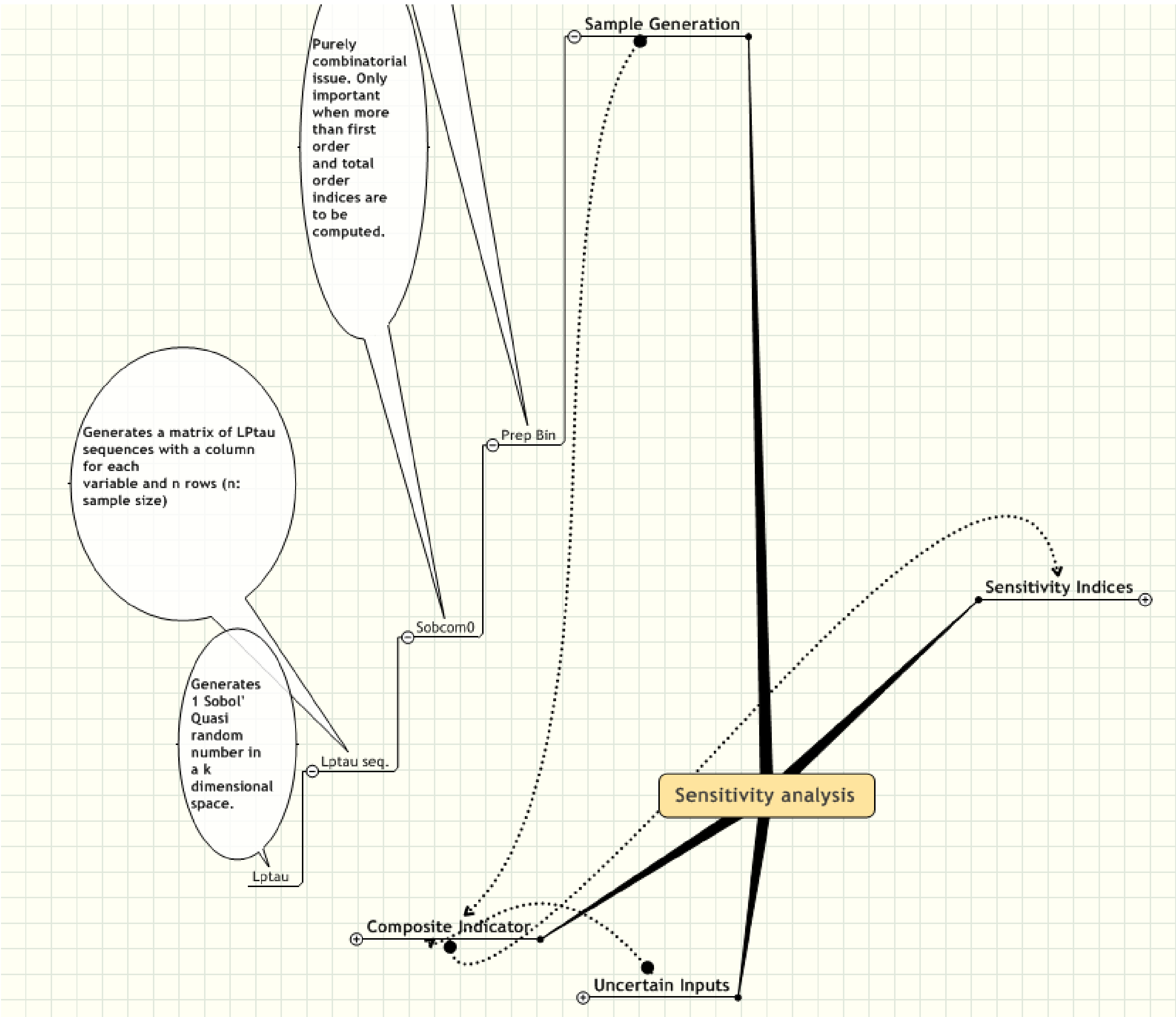

Figure 1.2: generation of the random numbers

First, equally distributed random numbers between 0 and 1 are created by the LP_τ programme. These random numbers are produced with the help of a sampling scheme named Sobol' LP_τ sampling. The resulting sequences of LP_τ vectors are quasi-random sequences which are defined as sequences of points without inherent random quality. In general, sequences of quasi-random vectors $V_1, ..., V_n$ should satisfy the following requirements:

- The uniformity of the distribution is to be optimal when the length of the sequence tends to infinity.
- Uniformity of vectors $V_1, ..., V_n$ should be observed for fairly small n.
- The algorithm used for the computation of the vectors should be simple.

Indeed LP_τ sequences do fulfill these constraints. In many cases they result in better convergence than random points in the MC algorithm with finite constructive dimension. There are widely available programmes, written in C and FORTRAN77, containing the algorithm to generate the sequences. The theoretical development of the LP_τ sequences is described in Bratley and Fox (1988). The sample points are uniformly distributed in a hypercube. A two-dimensional pattern of a LP_τ sequence is shown in the following plot.

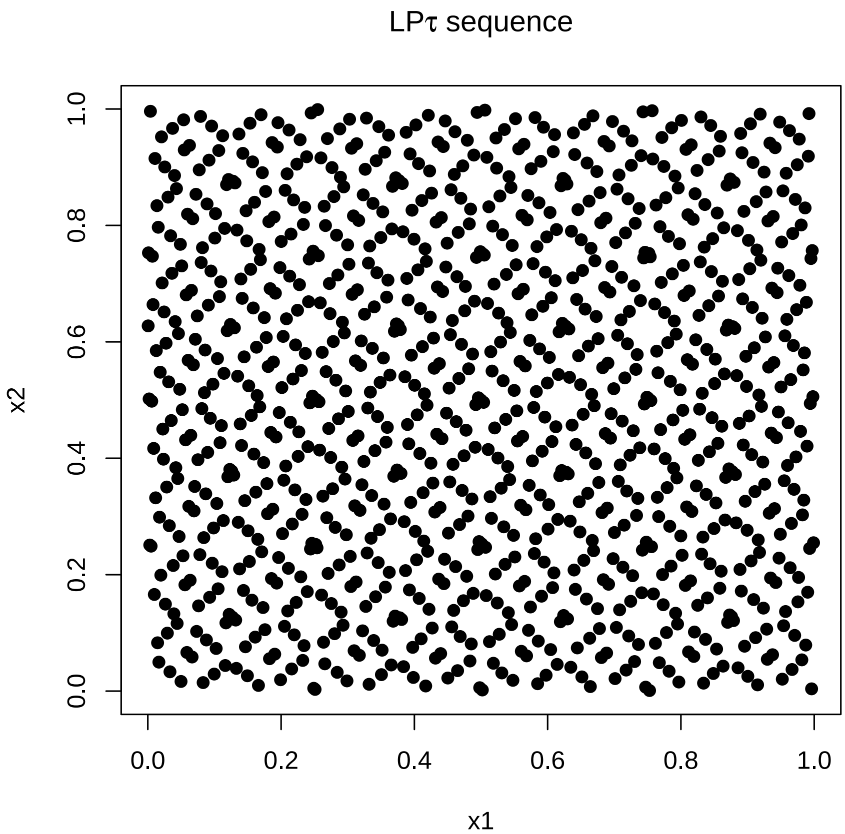

Figure 1.3: Sequence with 1000 points

The main avantage of using quasi-random points is a fast rate of convergence. For large n the approximate error can approach $1/n$, compared with the error of standard MC methods, which is of the order $1/\sqrt{n}$. There are several applications for the use of LP_τ points. For example multidimensional integration or trial points in multicriteria decision making. Other possible methods are for instance simple random sampling and latin hypercube sampling. The programme named LP_τseq creates a sequence of LP_τ points which are used in the PrepBin programme to create the decision matrix. The sample size is always 1024, because that is the base sample size required for the computation of an integral by the Sobol method (see Saltelli et al., 2000 p. 190).

1.3 Execution of the programme PrepBin

In case only the first order and total variance of the output is calculated, PrepBin generates a LP_τ matrix with 8 columns ($number\ of\ input\ factors \cdot 2$) and $n = 1024$ rows. Afterwards, PrepBin divides this matrix in two equal submatrices, say SubM1 and SubM2. It also generates an empty matrix of dimensions 10240×4 ([$number\ of\ input\ factors \cdot 2 + 2] \times$ *sample size x number of input factors*). This matrix is systematically filled out with: SubM1, SubM2, SubM2 with the first column replaced by the first of SubM1 and so on. This program implements a procedure, described in detail in Saltelli et al. (2008) pp. 164-167, that provides all pairs of first-order and total effect at a cost of ($number\ of\ input\ factors + 2$) $\times$ *sample size* model runs. Any interaction term between two input factors is computed at the additional cost of model evaluations per sensitivity measure.

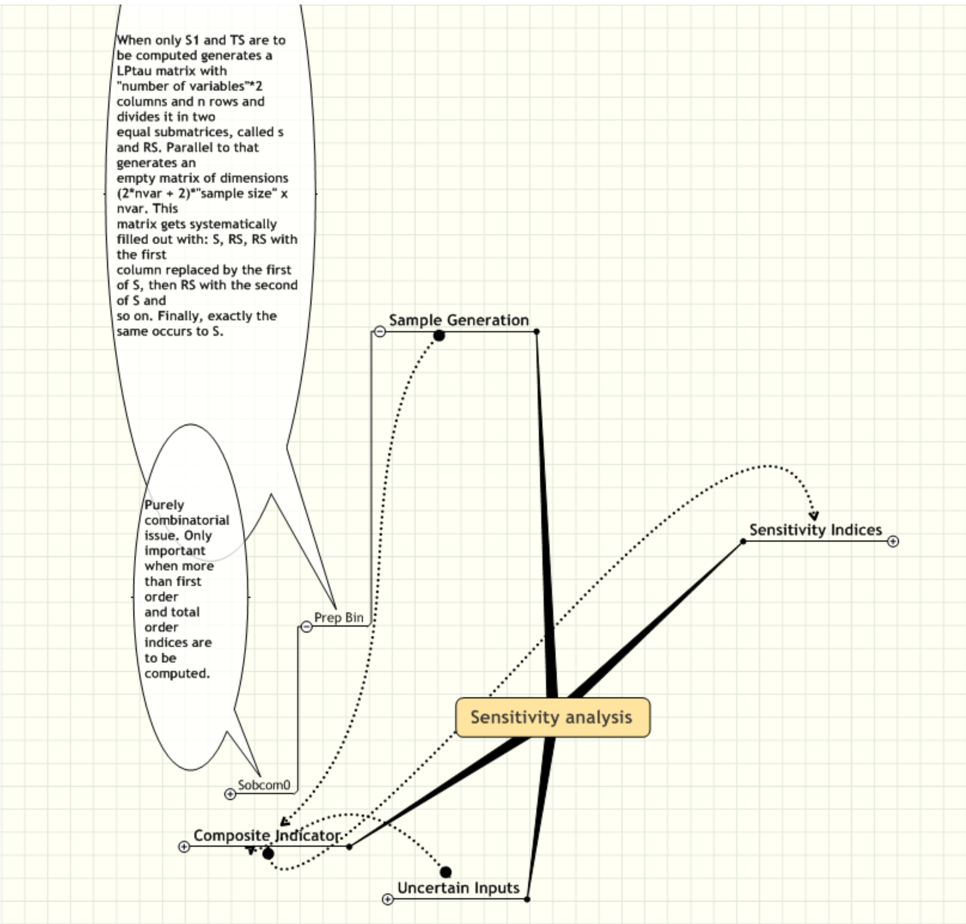

Figure 1.4: Generation of the Decision Matrix

Finally exactly the same is done to SubM1. The resulting matrix is the decision matrix (called www) mentioned above. This is one of the inputs of the composite indicator programme; the others are uncertain inputs which are explained in the following chapter.

Chapter 2

Uncertain Inputs

The goal of the `Uncertain Inputs` programme is to calculate the relevance of the variance caused by different uncertain inputs. Indeed uncertain inputs are action alternatives to do a determined task. The set of action alternative of each method is called input factor or input trigger.

2.1 Aggregation

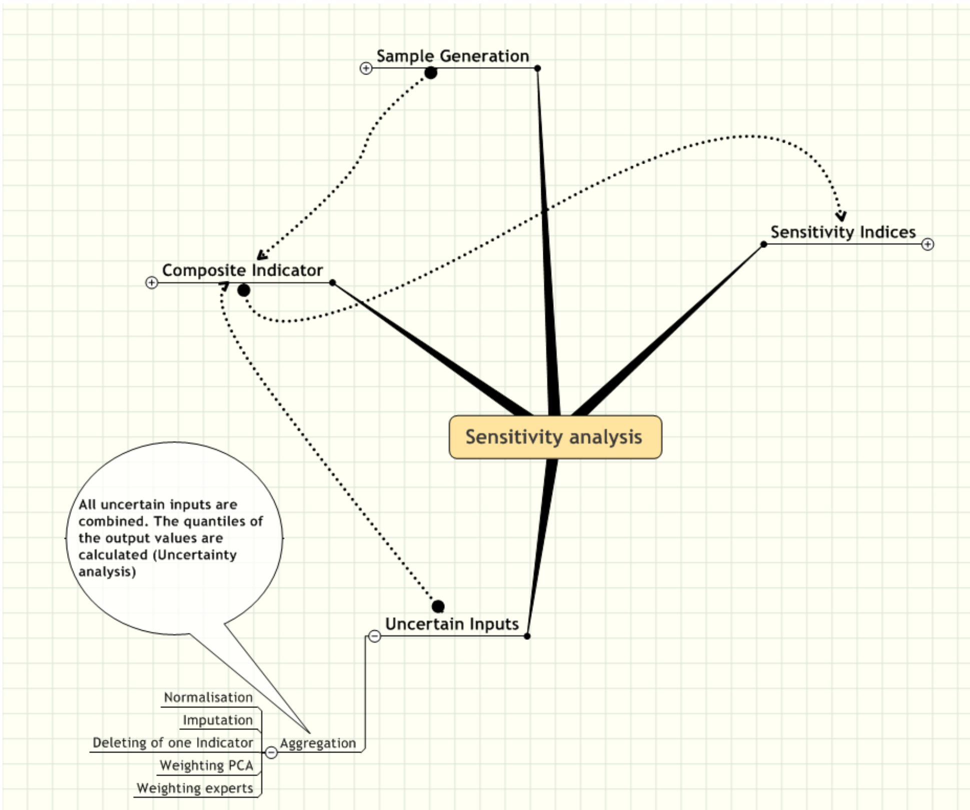

Figure 2.1: Uncertain Inputs

Aggregation is conducted by the composite indicator programme. It includes four input triggers:

- Imputation,
- normalisation,
- delete one indicator, and
- weighting.

The information provided by the four input triggers determines the way the indicators included in the KEI dataset is combined into a single number per country and year. In any case, the aggregation rule that is used is the weighted arithmetic average of the normalized indicators. We will explain this in more detail in the coming sections.

2.1.1 Imputation

Within the trigger *imputation* no real imputation is conducted itself. Within the given context one imputed table of data is selected. Within this study one of five different datasets can be chosen which is due to the application of a multiple imputation framework based on Rubin's work (see RUBIN, 1978). The implementation of the multiple imputation on knowledge economy indicators can be drawn from HUERGO et al. (2008).

The associated indicator function for the five imputed datasets can be described as:

$$D_{imp} = \begin{cases} \text{dataset}_1 & \text{for } x \in [0, 0.2] \\ \text{dataset}_2 & \text{for } x \in (0.2, 0.4] \\ \text{dataset}_3 & \text{for } x \in (0.4, 0.6] \\ \text{dataset}_4 & \text{for } x \in (0.6, 0.8] \\ \text{dataset}_5 & \text{for } x \in (0.8, 1] \end{cases}$$

where x is a random number of the decision matrix. For m datasets, the ith table is chosen with $i = \lfloor x \cdot m + 1 \rfloor$.

2.1.2 Normalisation

The second input trigger is normalisation. Here, five methods are used. Each normalization method is then applied to each of the five imputed datasets

The five methods are as follows:

1. Standardisation with subtraction of the mean:

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

   Where x_{qc}^{t} is the mean value of all countries.

2. Standardisation without subtraction of the mean:

$$I^t_{qc} = \frac{x^t_{qc}}{\sigma^t_{qc=\overline{c}}}$$

3. Distance to a reference country with subtraction of the mean:

$$I^t_{qc} = \frac{x^t_{qc} - x^t_{qc=\overline{c}}}{x^t_{qc=\overline{c}}}$$

4. Distance to a reference coutry without subtraction of the mean:

$$I^t_{qc} = \frac{x^t_{qc}}{x^t_{qc=\overline{c}}}$$

5. Cyclical indicators:

$$I^t_{qc} = \frac{x^t_{qc} - E(x^t_{qc})}{E_t(\left|x^t_{qc} - E(x^t_{qc})\right|)}$$

EU-25 was chosen to be *reference country*.

2.1.3 Deleting one Indicator

The third input trigger is deleting one indicator, similarly to the delete-one-jackknife method. According to the related random number in the decision matrix it is decided whether one indicator gets eliminated or not and in case of elimination which one it is. This input factor is very helpful for data sets in which one single indicator explains a large part out the output variance. Its indicator function can be written as:

$$D_{di} = \begin{cases} \text{all indicators in} & \text{for } x \in [0, 1/(n+1)] \\ \text{first indicator excluded} & \text{for } x \in (1/(n+1), 2/(n+1)] \\ \ldots & \\ \text{last indicator excluded} & \text{for } x \in (n/(n+1), 1] \end{cases}$$

where n is the number of indicators and x a random number of the decision matrix.

As one further option, no indicators are deleted which aims being the standard case of analysis.

2.1.4 Weighting

The fourth input trigger is weighting. In this deliverable, only two weighting methods are used, equal weighting or principle component analysis (pca) weighting (cf. JRC/OECD, 2008, pp. 89). The selection of the weighting is conducted by a random number being either below or not below 0.5.

The weighting, however, is dependant on the last input trigger. In case of a deletion of one indicator, a reweighting has to be applied. Let n be the number of indicators of interest. If no indicator is excluded by the third trigger, equal weighting simply creates a vector of weights w_i with $w_i = \frac{1}{n}\ i = 1, ...n$. In case one indicator is excluded, the equal weights have to be

$$w_i = \frac{1}{n-1} \quad i = 1, ..n\ .$$

pca weighting can only be done for matrices with more rows than columns. In our example the data matrix has 29 rows (representing the countries) and 116 columns (representing the indicators). However, a segmentation into the indicator groups with recursive pca weighting can be applied.

To solve this problems the following road is chosen: First the seven groups of indicators are pca weighted to obtain weights w_i for each single indicator. This step is possible because none of the groups contains more than 29 indicators. Afterwards composite indicators I_g(one for each group) are created:

$$I_g = \sum_{i \in g} I_i \cdot w_i \quad g = 1..7 \;\; i = 1..116$$

Then these group indicators are re-pca-weighted in order to receive group weights w_g. Finally each single indicator weight was multiplied by the weight of the group it belonged to.

Then each single indicator weight is multplied by the indicator group weight of the group it belongs to. Finally each component of the resulting vector of weights was divided through the vector's sum to guarantee that the weights sum up to 1. The indicator function of the weighting trigger can be written as:

$$D_{wei} = \begin{cases} \text{equal weights} & \text{for } x \in [0, 0.5] \\ \text{pca weights} & \text{for } x \in (0.5, 1] \end{cases}$$

where x is a random number of the decision matrix.

Independent of the selected weighting method the data matrix is finally multiplied by the weighting vector (linear aggregation). The results are saved as a matrix of composite indicator scores.

Chapter 3

Composite Indicators

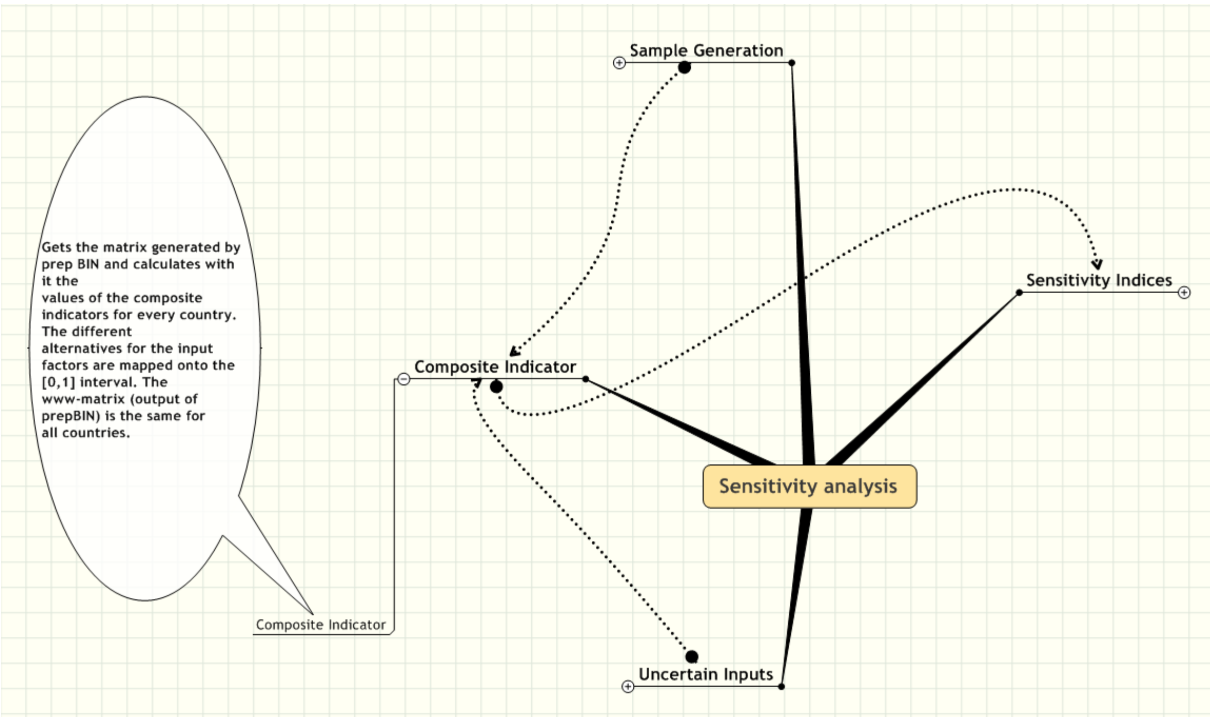

Figure 3.1: Composite indicators

The www-matrix if passed to `Composite Indicators` which runs the aggregation programme 10240 times (number of rows of the decision matrix). So it is closely related to sample generation and uncertain inputs, as shown in Figure 3. The decision matrix is the same for all countries. The different action alternatives are mixed randomly and the resulting composite indicators feed the next core programme.

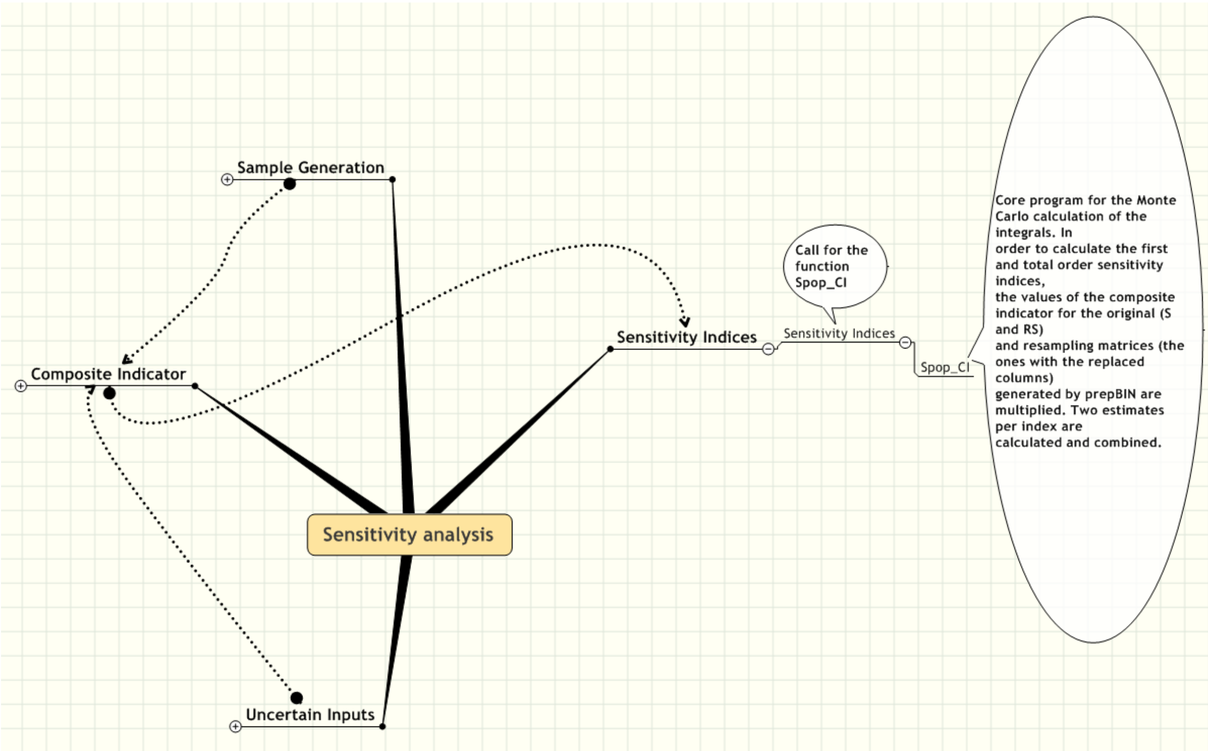

Figure 3.2: Calculation of the Output

`Sensitivity Indices` is the core programme of the sensitivity analysis. According to the results (scores) of the `Composite Indicator` programme SpopCI calculates the first [S] order and total effect [ST] indices for a total of k input factors. Through the use of Monte-Carlo integration and resampling techniques the variance of the simulated composite indicators is divided into each input factor's part of the total variance. The corresponding interactions are also quantified. The first order sensitivity index (S_i) of a factor x_i measures the main effect of x_i on the output (the fractional contribution of $x_i$to the total variance).

The total effect sensitivity index is defined as the sum of all the sensitivity indices involving the factor in question. For example, suppose that we have three factors in our model. The total effect of factor 1 on the output variance, denoted by TS(1), is given by TS(1) $= S_1 + S_{1,2} + S_{1,3} + S_{1,2,3}$, where $S_1$is the first order sensitivity index of factor 1. $S_{1,J}$ is the second-order sensitivity index for the two of factors 1 and j ($j \neq 1$), i.e. the interaction between factors 1 and j ($j \neq 1$) and so on (see SALTELLI et al., 2000 p. 177).

Chapter 4

Output interpretation

4.1 Sensitivity Analysis Results: Variance decomposition

As mentioned above one of the main goals of the programme is to conduct sensitivity analysis in order to apportion the variance of the output (in our case the variance of the countries' composite indicator scores) to the uncertain input factors.

The middle part of Figure 4.1 shows which part of the output variance is caused by the four triggers taken singularly. A general result is that normalisation and/or imputation are responsible for the variance of the countries' scores. On the other hand, the exclusion of an indicator from the dataset or the selection of the weighting method (i.e. equal weighting or PCA-based weighting) does not affect the output variance significantly. All four input factors, taken singularly, explain about half of the output variance in the majority of the countries. For few countries - Greece, Ireland, Luxembourg, Italy, EU25, and Japan - the normalisation method is almost entirely responsible for the output variance and the impact (first order sensitivity measure) is near 1.0. This implies that no interaction effects have an impact on the variance of those country scores and hence the differences of the Total Effect minus First Order sensitivity measures are practically zero (lower part of Figure 4.1).

This is not, however, the case for the remaining countries, for which about half of the variance of the countries scores is due to interactions among the factors themselves. The low part of Figure 4.1 shows the difference (Total Effect - First Order): the greater the difference, the more that factor is involved in interactions with the other factors. We can see that imputation and normalisation method are also dominant here. Another observation is that the chosen weighting method has more influence to the output variance due to the interactions with the other factors than does the exclusion of an indicator from the dataset.

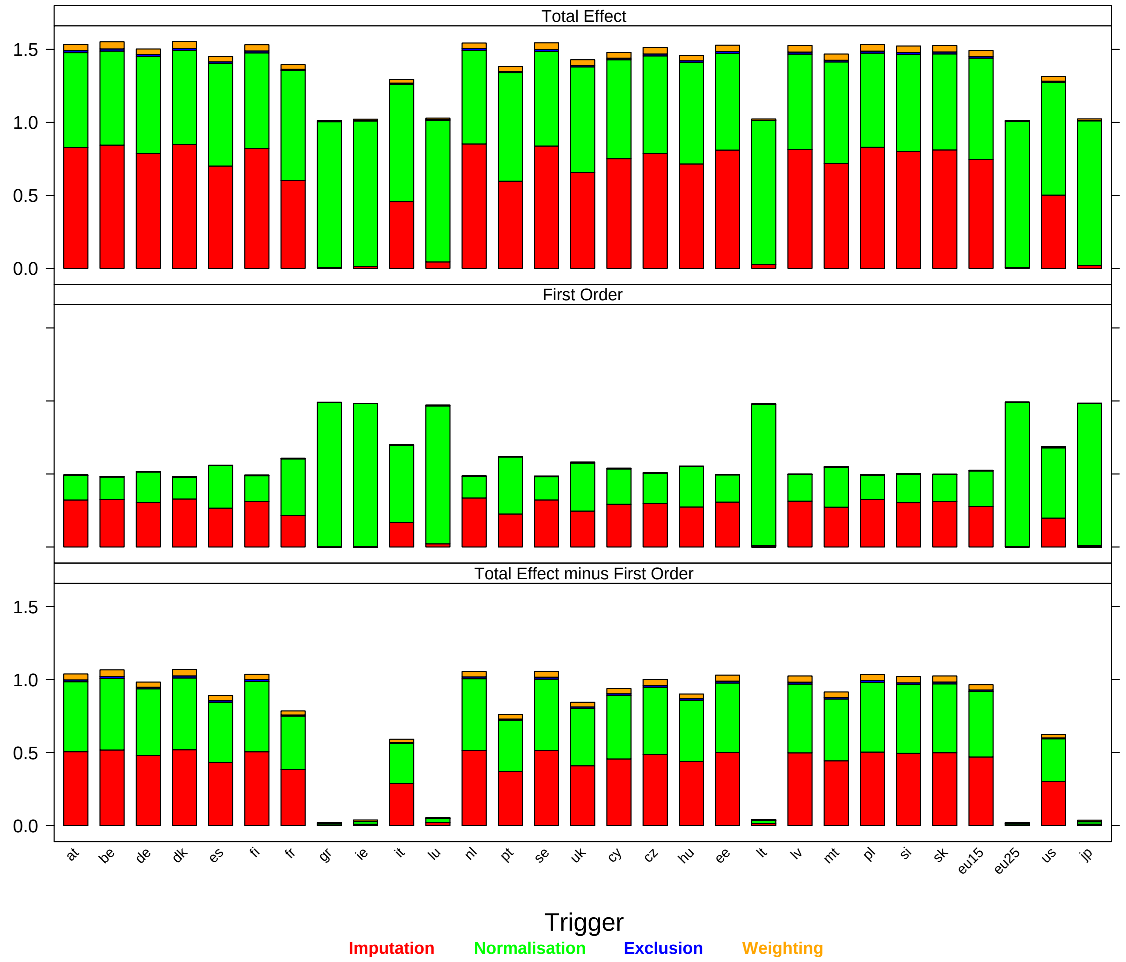

Figure 4.1: First Order and Total Effect Sensitivity Indices for the country scores in the composite indicator due to uncertainties in the Imputation, Normalisation, Exclusion of an indicator and Weighting

4.2 Uncertainty Analysis Results: Frequencies of Countries Ranks

The programme also allows for the estimation of the frequencies of the country ranks. The country scores in the Knowledge Economy Index estimated using all 10240 values resulting from the Monte Carlo execution of the decision matrix is converted to a vector containing the values of 1 to 29, representing the rank of each country according to that composite indicator.

Figure 4.2 shows how often each country achieves which rank. The darker blue a square is the more often a country takes this place. White squares mean the dedicated country never takes the associated rank. For example in 47,8 per cent of the cases Finland is ranked first. According to the ranks which are taken most often, the top countries are Finland, Sweden and the United Kingdom. In general, interpretation of ranks in such a way is complicated, because some countries exhibit a wide range of scores, as Malta does.

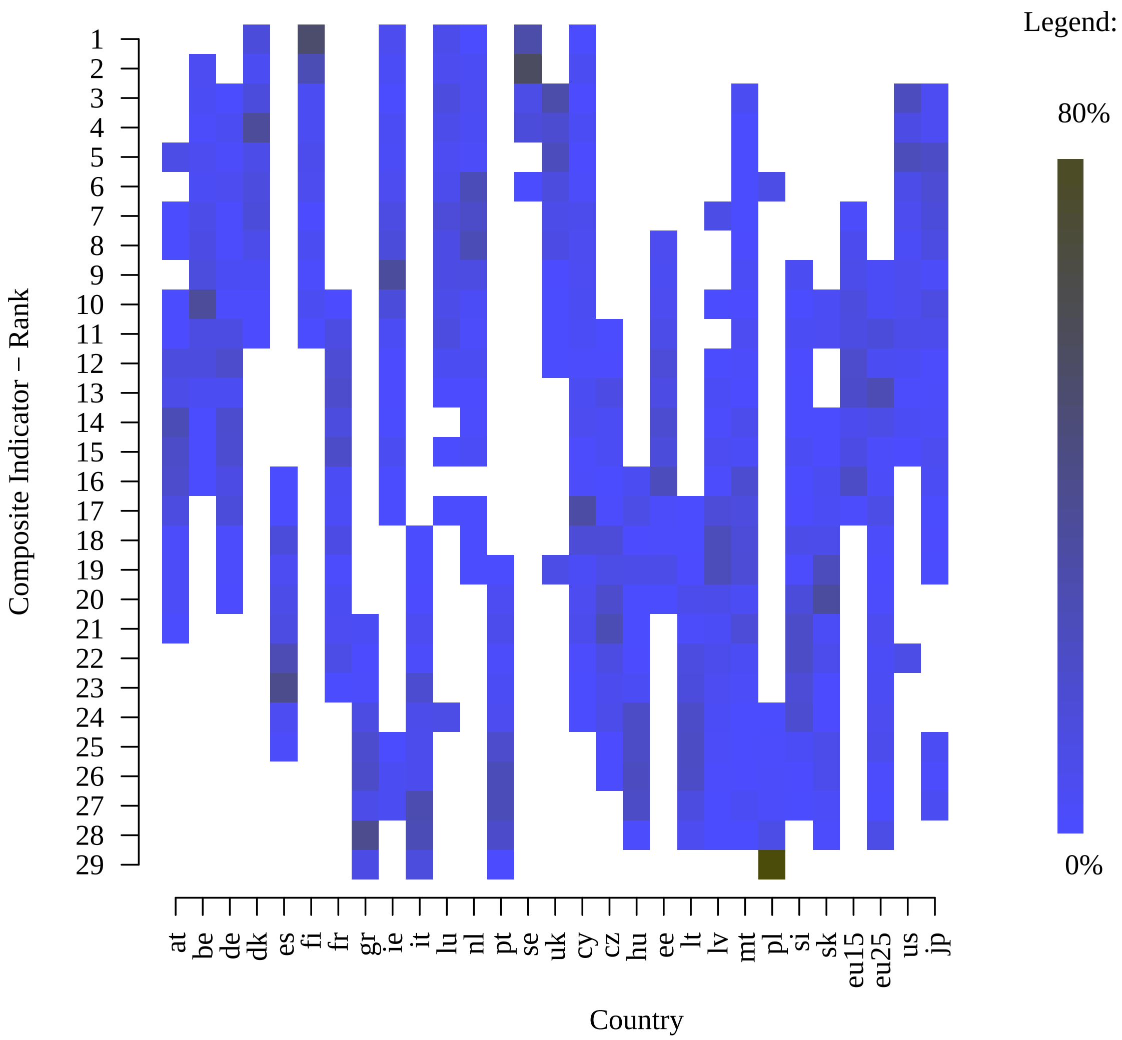

Figure 4.2: Country Ranking

Also outliers can appear: though the United Kingdom is ranked third most often, it takes only the 19th place at some scores.

An alternative way to create a rankings can be achieved by considering all combinations of triggers. Indeed, in this case a decrease of computation burden is achieved. Instead of using a decision matrix, each possible trigger combination is selected exactly once which yields 5850 combinations in this example. This follows from

$$5850 = (5\,imputed\ datasets \times 5\,normalisation\ methods \times (116+1)\,sets \times 2\,weighting\ methods)$$

The ranking of these 5850 composite indicators is shown in Figure 4.2.

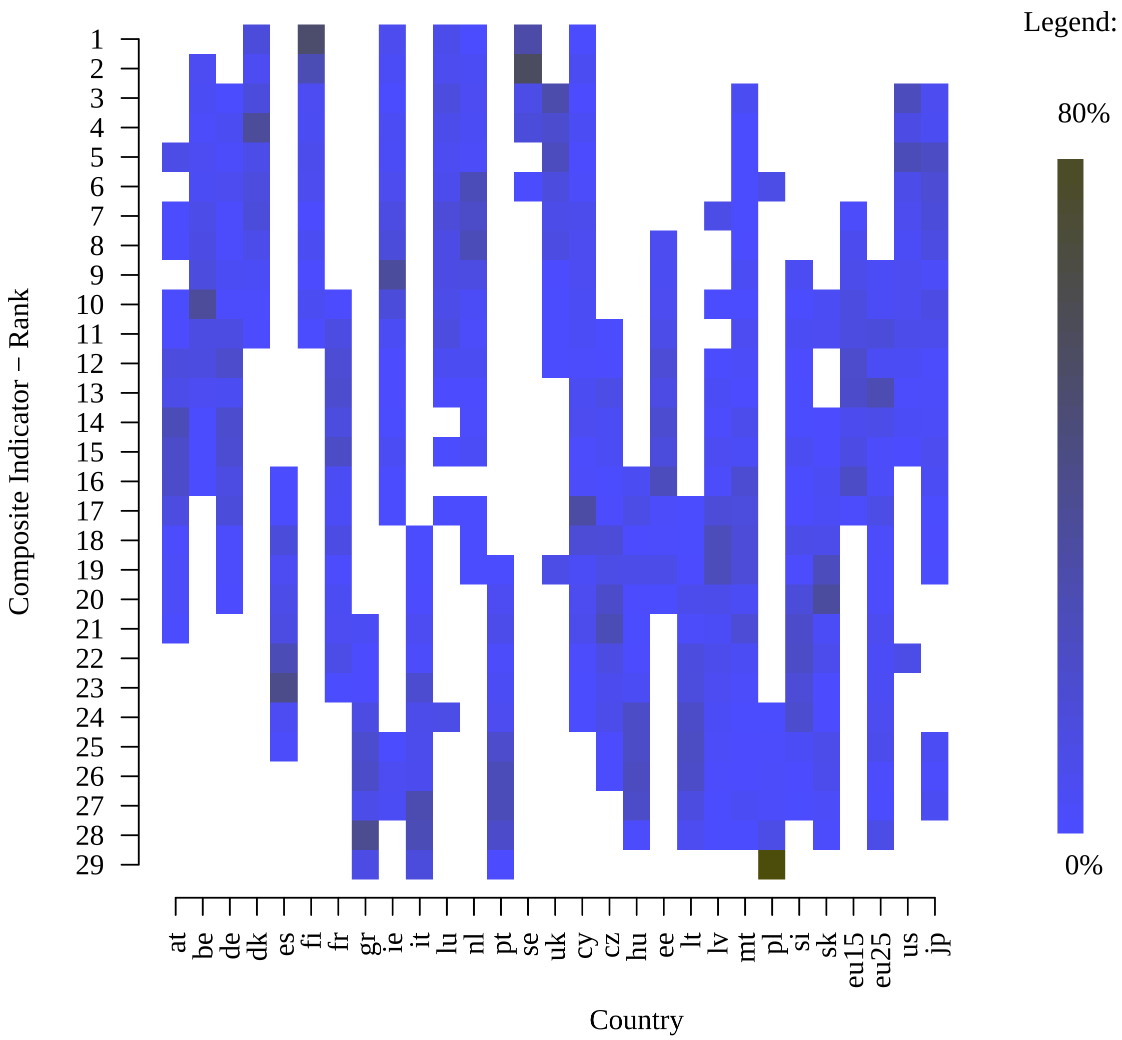

Figure 4.3: Country Ranking

It is obvious that the two rankings almost do not differ from each other. This implies that the ranking schemes are independent from the different ranking computations. However, the advantage of the Sobol scores is gained from large scale applications where the computation effort of all combinations is highly non-linear.

Bibliography

Arundel, A. and **Hansen, W.** (**2008**): *Indicators for the Knowledge-Based Economy: Summary Report*. KEI deliverable D2.5, `http://kei.publicstatistics.net`.

Huergo, L., **Enderle, T.** and **Münnich, R.** (**2008**): *Imputation of Knowledge Economy Indicators*. KEI deliverable D3.2, `http://kei.publicstatistics.net`.

JRC/OECD (**2008**): *Handbook on Constructing Composite Indicators. Methodology and user Guide*. Oecd statistics working paper, `http://browse.oecdbookshop.org/oecd/pdfs/browseit/3008251E.PDF`.

Nardo, M., **Saisana, M.**, **Saltelli, A.**, and **Tarantola, S.** (**2005**): *Input to Handbook of Good Practices for Composite Indicators' Development*. KEI deliverable D5.2, `http://kei.publicstatistics.net`.

Rubin, D. B. (**1978**): *Multiple Imputation in Sample Surveys - A Phenomenological Bayesian Approach to Nonresponse*. Proceedings of the Survey Research Methods Sections of the American Statistical Association, pp. 20–40.

Saisana, M., **Saltelli, A.** and **Tarantola, S.** (**2005**): *Uncertainty and Sensitivity Analysis as tools for the quality assessment of composite indicators*. Journal of the Royal Statistical Society Series A, 168(2), pp. 307–323.

Saisana, M., **Tarantola, S.**, **Schulze, N.**, **Cherchye, L.**, **Moesen, W.** and **Puyenbroeck, T. V.** (**2005**): *State-of-the-Art Report on Composite Indicators for the Knowledge-based Economy*. KEI deliverable D5.1, `http://kei.publicstatistics.net`.

Saltelli, A., **K.Chan** and **E.M.Scott** (**2000**): Sensitivity Analysis: Gauging the Worth of Scientific Models. Wiley Publisher.

Saltelli, A., **Ratto, M.**, **Andres, T.**, **Campolongo, F.**, **Cariboni, J.**, **Gatelli, D.**, **Saisana, M.** and **Tarantola, S.** (**2008**): Global Sensitivity Analysis: The primer. Wiley Publisher.

Thees, N., **Åkerblom, M.**, **Arundel, A.**, **Hansen, W.**, **Magg, K.**, **Ohly, D.** and **Münnich, R.** (**2008**): *Summary of Selected KBE Indicators and Categories: Data Annex*. KEI deliverable D2.3, `http://kei.publicstatistics.net`.

Vose, D. (**2006**): Risk analysis: a quantitative guide. Wiley Publisher.