Twelfth Annual Conference
on Global Economic Analysis
Santiago, Chile
June 10-12, 2009

Comparison of Different Updating Procedures and Identification of Temporal Patterns

Casiano Manrique-de-Lara-Peñate (ULPGC)
Dolores Santos-Peñate (ULPGC)

1 Introduction

Our paper deals with the comparison of different procedures for the updating and adjusting of National Accounting Matrices. We will concentrate our efforts in updating a series of National Accounting Matrices (NAM) for different European countries. We have generated these NAMs for 21 countries using the national accounts data provided by Eurostat. Some of these NAMs are not conveniently balanced what have forced us to reduce the number of countries considered. For the time being we have already balanced the NAMS for 11 countries referred to the period 2000-2006 (66 NAMs).

We will essentially test the appropriateness of different updating procedures, through the comparison between the "real" Nam and the obtained trough the adjustment of previous NAMs available for a certain country.

We believe this effort could help also to identify whether using different reference years really makes any difference in the adjustment results obtained.

The rest of the paper is organized as follows. In the next Section we define the updating problem studied. In Section 3, we present some computational experience and Section 4 includes some concluding remarks. The appendix1 contains the measures used to calculate the dissimilarity between the initial matrix and the final matrix.

2 The updating problem presented

A National Accounting Matrix (NAM) is a square matrix whose rows and columns represent the resources and uses of separate economic accounts which together summarize the basic flows within the different groups of transactions or categories of transactions in an economy. The incomings (rows) and outgoings (columns) are necessarily equilibrated following basic economic accounting standards, therefore the sum of columns and rows for each account must be the same. This matrix presents an accounting system that provides information about aspects of an economy like the structure, composition and level of production, the value added generated by the factors of production and the distribution of income among different groups of households. A more disaggregated NAM can be constructed using an Input-Output Table as starting point and includes figures about consumption, the structure of production, and the flows related to foreign trade, savings and investment.

The updating of a NAM consists of determining the NAM corresponding to period t=T starting from a known NAM, which corresponds to moment t=0, and from certain available information about the NAM in moment t=T. It is a question of finding a NAM X^T whose structure is, in a certain sense, as close as possible to the structure of the NAM in moment t=0, X^0. Normally this structure is defined in terms of the *row (horizontal) coefficients* and the *column (vertical) coefficients*, assuming that if element x^0_{ij} of the initial matrix is null then the corresponding element x^T_{ij} of the final matrix is also null. If $X=(x_{ij})_{1 \le i,j \le n}$ is a NAM and $S^{\neq 0}=\{(i,j) : x_{ij} \neq 0\}$, then the matrix of row coefficients, $A = (a_{ij})_{1 \le i,j \le n}$, and that of column coefficients, $B = (b_{ij})_{1 \le i,j \le n}$, can be defined as

$$a_{ij} = \frac{x_{ij}}{\sum_j x_{ij}} \qquad \text{if (i,j)} \in \boldsymbol{S}\text{, and } a_{ij} = 0 \text{ if (i,j)} \notin \boldsymbol{S}$$

$$b_{ij} = \frac{x_{ij}}{\sum_i x_{ij}} \qquad \text{if (i,j)} \in \boldsymbol{S}\text{, and } b_{ij} = 0 \text{ if (i,j)} \notin \boldsymbol{S}$$

We consider that the adjustment objective consists of determining a NAM X^T such that the ratios $\frac{a_{ij}^T}{a_{ij}^0}$, or $\frac{b_{ij}^T}{b_{ij}^0}$, for (i,j) belonging to S={(i,j) : 1 ≤i,j≤n}, are as similar as possible to unity, distinguishing between two types of adjustment, the horizontal and the vertical one, depending on whether we approximate the ratios of horizontal or vertical coefficients to unity. The definition of the concept of proximity determines the updating criterion. Assuming that all elements in the matrices are non-negative, for horizontal coefficients, the criteria used in this paper are defined as follows (for the vertical updating, the horizontal coefficients are replaced by the vertical ones):

- Criterion L_1

$$F_1(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} \left| \frac{a_{ij}^T}{a_{ij}^0} - 1 \right|$$

- Criterion L_2

$$F_2(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} \left(\frac{a_{ij}^T}{a_{ij}^0} - 1 \right)^2$$

- Criterion (L_1,L_∞) (weighted sum of L_1 and L_∞)

$$F_3(A^T, A^0) = w_1 \sum_{(i,j)\in S^{\neq 0}} \left| \frac{a_{ij}^T}{a_{ij}^0} - 1 \right| + w_2 \max_{(i,j)\in S^{\neq 0}} \left| \frac{a_{ij}^T}{a_{ij}^0} - 1 \right|$$

where w_1 and w_2 are positive weights.

- Criterion RAS:

$$F_4(A^T, A^0) = RAS(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} \frac{\gamma_i}{\gamma} a_{ij}^T \ln \frac{a_{ij}^T}{a_{ij}^0}$$

- Criterion MSCE: Minimum sum of cross entropies

$$F_5(A^T, A^0) = MSCE(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} a_{ij}^T \ln \frac{a_{ij}^T}{a_{ij}^0}$$

- Criterion Kullback:

$$F_6(A^T, A^0) = K(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} a_{ij}^T \ln \frac{a_{ij}^T}{a_{ij}^0} + \sum_{(i,j)\in S^{\neq 0}} a_{ij}^0 \ln \frac{a_{ij}^0}{a_{ij}^T}.$$

Where $\gamma_i = \sum_j x_{ij}$ and $\gamma = \sum_i \gamma_i$. Therefore γ_i is the sum of the elements of row i, which is equal to the sum of the elements of column i, and γ is the sum of all the elements of matrix X^0.

The updating problem is formulated as an optimization problem where the objective function is one of the previously mentioned criterion and the restrictions are determined by the column and row totals, which are supposed to be known, in addition to other given values.

The vertical criterium F_1 has been used by Matuszewski, Pitts and Sawyer (1964). Criterium F_4 corresponds to the well known RAS algorithm and F_5 is based on an entropy measure which has been used by Golan, Judge and Robinson (1994), and Robinson, Cattaneo and El-Said(2000), among others. Criteria F_4 and F_5 are particular cases of the updating criterion of the minimization of the sum of weighted crossed entropies (McDougall, 1999). In contrast to Criterion F_4 and F_5, F_6 is a distance in the mathematical sense. Measures F_1 to F_5 have been used by Manrique de Lara and Santos-Peñate (2004).

The basic updating problem is formulated as follows:

$$\min F(A)$$

subject to:

$$\sum_j x_{ij} = \gamma_i,\ i = 1,...,n \qquad (1)$$

$$\sum_i x_{ij} = \gamma_j,\ j = 1,...,n \qquad (2)$$

$$x_{ij} = 0,\ \forall (i,j) \notin S^{\neq 0}$$

$$x_{ij} \geq 0,\ \forall (i,j)$$

where γ_i are given, F is one of the criterion defined previously and $a_{ij} = \frac{x_{ij}}{\gamma_i}, \ \forall (i,j) \in S$. But we can also consider the vertical coefficients or both simultaneously, what generates three objective functions in total.

More general formulations include other constraints in addition to the balance restrictions. On the other hand, certain situations require some particular treatment; for example, if the structure of the accounting matrices varies in time, which means that the index set with positive values is not constant.

In this paper, we apply the previous criteria to update a series of NAM for 11 European countries. In order to evaluate the results of the updating procedure and compare the criteria used, we calculated the dissimilarities between the initial and final matrices using the measures defined in the appendix 1. The results provided by the updating procedure can be used to classify the results obtained using the different NAMs available for different periods.

2. Computational experience

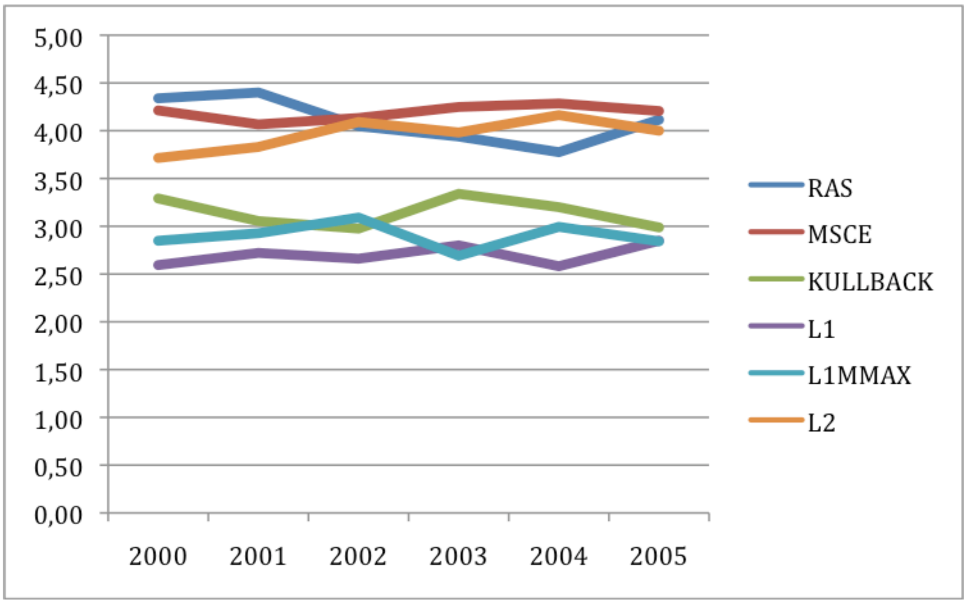
In this research we have prepared the National Accounting Matrices (NAM) of 11 European economies for the years 2000 to 2006 using the data published by Eurostat (2009). The experiment consists in updating the 2006 NAMs for the different countries using the six (2000-2005) previously existing matrices for each country as benchmark reference.

In total we generated 1188 updated matrices (11 countries, 6 criterion, 3 objective functions and 6 reference years). In order to maintain a coherent and simple updating structure, we introduced in the optimization program only those values that where positive in both versions simultaneously. The rest of the values where fixed to the observed values beforehand. We implemented the different updating approaches in GAMS and used Matlab to calculate the dissimilarity measures between pairs of matrices.

3. Concluding remarks

There can be obtained two main conclusions from this experience so far. The first one is that we can obtain a certain hierarchy of the different updating methods. In figure 1, we show the position of each of the criterion for each of the reference years for all the countries considered. The nearer to the value one the better are the results obtained. There are two clearly distinguishable groups. The worse of both groups is formed by the RAS, the MSCE and the L2 criteria. The best performing group is the one composed by the Kullback, the L1MMAX and the L1 criteria. The use of different reference years does not affect the contents of each of the groups and slightly alters the individual position within the group only for the RAS and the L1MMAX criteria.

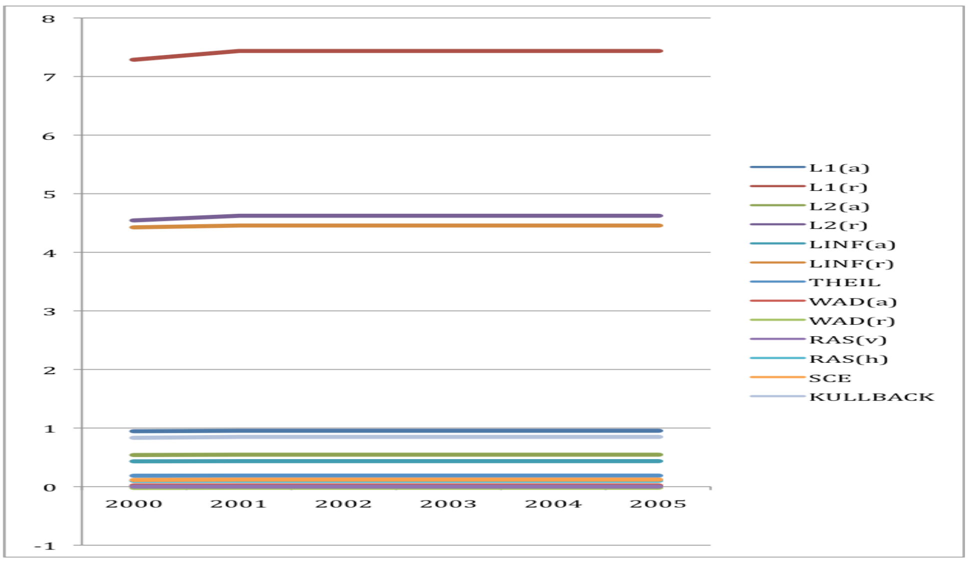
Figure 1: Hierarchy of the different criteria



Once we have obtained the estimated final matrices for any year, a comparison between any pair of matrices, initial, final and final estimated, is possible. We can compare the updating criteria and the importance of the number of years elapsed between the estimated matrix and the one used as initial reference of the updating procedure.

The other main conclusion can be stated as the rather insignificant role in using different reference years. Figure 2 can complement the information given by Figure 1 in this sense. In this figure we show the different values of the dissimilarity measures calculated for each of the years for the case of Germany. It is pretty obvious that there are no crucial gains in the measures by using different years.

Figure 2: Dissimilarity measures for Germany



The first step in the future research will be focused to a deeper analysis of the data, the updating criteria and results. Other subjects are the definition of macroeconomic behaviours by means of cluster analysis, the treatment of different scenarios defined by the set of data we know, and the structural analysis.

5. References

Bacharach, M. (1965) Estimating nonnegative matrices form marginal data. *International Economic Review*, 6, 3, 294-310.

Golan, Judge and Miller(1996). Maximum entrophy econometrics, robust estimation with limited data. John Wiley & Sons.

Golan, Judge and Robinson (1994). Recovering information from incomplete or partial multisectorial economic data. *The Review of Economics and Statistics*, 76, 541-549.

Harrigan, F.J. and I. Buchanan (1984). A quadratic programming approach to IO estimation and simulation. *Journal of Regional Science*, 24, 3, 339-358.

Macgill, S.M. (1978). Convergence and related properties for a modified biptoportional matrix problem. *Environment and Planning A*, 11, 499-506.

Manrique de Lara, C. and D.R. Santos Peñate (2004). New nonlinear approaches for the adjustment and updating of a SAM. *Economis of Planning*, 36, 259-272.

Macgill, S.M. (1977). Theoretical properties of biproportional matrix adjusment. *Environment and Planning A*, 9, 687-701.

Matuszewski, T.I., P.R. Pitts and J.A. Sawyer (1964). Linear programming estimates of changes in input coefficients. *Canadian Journal of Economics and Political Science* 2, 203-210.

McDougall, R.A. (1999). Entropy theory and RAS are friends. *http: //www.sjfi.dk /gtap /papers /McDougall.pdf*}.

Robinson, S., A. Cataneo and M. El-Said (2001). Updating and estimating a social accounting matrix using cross entrophy methods. *Economic Systems Research*, 13,1, 47-64.

Acknowledgements

This work has been partially supported by the *Ministerio de Educación y Ciencia* (grant ECO2008-05589).

The assistantship of Héctor Rodríguez and Ivonne Pérez is greatly acknowledged.

Appendix1

The following measures are defined for horizontal coefficients. The definitions for vertical coefficients are similar. Additionally, we calculated the correlation coefficient. As in the previous sections, the notation $S^{\neq 0}$ represents the set of indices (i, j) such that $a_{ij} \neq 0$.

- Metrics L_1, absolute and relative:

$$L_1(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} \left| a_{ij}^T - a_{ij}^0 \right|$$

$$L_{1R}(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} \left| \frac{a_{ij}^T}{a_{ij}^0} - 1 \right|$$

- Metrics L_2, absolute and relative:

$$L_2(A^T, A^0) = \sum_{(i,j)\in S} \left(a_{ij}^T - a_{ij}^0 \right)^2$$

$$L_{2R}(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} \left(\frac{a_{ij}^T}{a_{ij}^0} - 1 \right)^2$$

- Metrics L_∞, absolute and relative:

$$L_\infty(A^T, A^0) = \max_{(i,j)\in S} \left| a_{ij}^T - a_{ij}^0 \right|$$

$$L_{\infty R}(A^T, A^0) = \max_{(i,j)\in S^{\neq 0}} \left| \frac{a_{ij}^T}{a_{ij}^0} - 1 \right|$$

- Theils measure:

$$T(A^T, A^0) = \sqrt{\frac{\sum_{(i,j)\in S} \left(a_{ij}^T - a_{ij}^0 \right)^2}{\sum_{(i,j)\in S} \left(a_{ij}^0 \right)^2}}$$

- Weighted absolute difference (WAD), absolute and relative:

$$WAD(A^T, A^0) = \frac{\sum_{(i,j)\in S} \left|a_{ij}^T + a_{ij}^0\right| \left|a_{ij}^T - a_{ij}^0\right|}{\sum_{(i,j)\in S} \left|a_{ij}^T + a_{ij}^0\right|}$$

$$WAD_R(A^T, A^0) = \frac{\sum_{(i,j)\in S^{\neq 0}} \left|\frac{a_{ij}^T}{a_{ij}^0} + 1\right| \left|\frac{a_{ij}^T}{a_{ij}^0} - 1\right|}{\sum_{(i,j)\in S^{\neq 0}} \left|\frac{a_{ij}^T}{a_{ij}^0} + 1\right|}$$

For the following definitions, we will assume that all matrix elements are equal or greater than cero.

- RAS:

$$RAS(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} \frac{\gamma_i}{\gamma}\, a_{ij}^T \ln \frac{a_{ij}^T}{a_{ij}^0}$$

- Sum of cross entropies (SCE):

$$SCE(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} a_{ij}^T \ln \frac{a_{ij}^T}{a_{ij}^0}$$

- Kullback measure:

$$K(A^T, A^0) = \sum_{(i,j)\in S^{\neq 0}} a_{ij}^T \ln \frac{a_{ij}^T}{a_{ij}^0} + \sum_{(i,j)\in S^{\neq 0}} a_{ij}^0 \ln \frac{a_{ij}^0}{a_{ij}^T}$$

Where $\gamma_i = \sum_j x_{ij}$ and $\gamma = \sum_i \gamma_i$. Therefore γ_i is the sum of the elements of row i, which is equal to the sum of the elements of column i, and γ is the sum of all the elements of matrix X^0.