**Global updating of supply-use tables with parameter calibration of estimation errors**

Xesús Pereira López1

Manuel Fernandez-Grela2

Melchor Fernández Fernández3

**Abstract**

Supply-Use tables (SUTs) are increasingly used in input-output analysis. Consequently, in recent years there is a growing interest in obtaining convergent updates not only of square matrices, but also of rectangular ones. The RAS technique is the most widely used updating method. The basic RAS procedure presents a number of advantages, but it suffers also from one important limitation: it requires information about the row and column sums of the matrices that will be updated.

The global procedure to distribute estimation differences presents an alternative for matrix updating that avoids this limitation. It consists in a scale algorithm that varies the restrictions of the optimization program one at a time in each iterative stage. This procedure allows using any variant of RAS to update SUTs without the need to know in advance the values of row and column sums.

This paper presents a mechanism that improves the performance of this procedure. The mechanism consists in going back in time through adjustments of survey SUTs to detect influential row parameters for the distribution of the difference in the estimation of the individual elements of the matrices. Once a plausible calibration of these parameters is reached, SUTs can be projected forward in time in a simple and efficient way.

1Universidade de Santiago de Compostela, Instituto Universitario de Estudos e Desenvolvemento de Galicia (IDEGA) & Departamento de Economía Cuantitativa. Email: xesus.pereira@usc.es.

2Universidade de Santiago de Compostela, GAME-IDEGA. Email: mf.grela@usc.es.

3Universidade de Santiago de Compostela, GAME-IDEGA. Email: melchor.fernandez@usc.es.

**1. Introduction**

The importance of supply-use tables (SUTs) for input-output analysis is continuously growing. A parallel interest on convergent updating, both of rectangular and square matrices, is developing in recent years. Well-known techniques, like the Euro (EU) method, as well as more recent ones, like SUT-RAS, bear witness to the methodological advances in this field. One of the main factors limiting the validity of these techniques is availability of information.

Improvements in matrix updating methods are desirable to fill the need for non-survey methods that can provide a solid alternative to the building of input-output frameworks through surveys. Survey methods are too costly to be implemented on a yearly basis. RAS is currently the most popular tool for updating. It is a technique based on successive multiplications of row and column elements of a base matrix and adjustment coefficients. This bi-proportional procedure was first proposed by Stone & Brown (1962), and after that a good number of extensions were developed in subsequent years (see, among others, Bacharach 1970, Allen & Lecomber 1975, or Szyrmer 1989). There are multiple alternatives to RAS, many inspired by it, and they are described in e.g. Lahr & Mesnard (2004), Jackson & Murray (2004), or Pavia *et al.* (2009). However, basic RAS presents a strong limitation, as row and column sums of the matrices to be adjusted need to be known in advance.

Global approaches to updating contribute to the stock of available methods by circumventing some of the difficulties found in other approaches. In particular, these approaches do not need information about row and column sums. The adoption of a wider framework allows global approaches to make adjustments based on any of the magnitudes recorded in the input-output framework.

A global procedure of allocation of estimation errors is a suitable, low-cost, alternative for matrix adjustments. It consists of a scale algorithm that includes the dynamics of RAS in a particular way, modifying at least one of the constraints in the associated optimization program in each iterative stage. Besides, it provides the advantage of being easily adaptable to various contexts, even where information is scarce. This procedure allows the possibility of using RAS or other alternative techniques to update SUTs without knowing in advance the row and/or column sums of the intermediate demand matrix, thanks to a computational development that can deal with the various magnitudes included in SUTs.

This paper aims to improve one variant of the global procedure of allocation of estimation errors. In particular, it articulates a mechanism to detect influential parameters by row to formalize *a posteriori*, in the iterative stage, the allocation of the errors in the estimation of the elements of the intermediate demand matrix. The mechanism consists in going backwards in time through successive adjustments of survey-based SUTs and comparing estimations with actual data. This comparison will provide a plausible calibration of the influential parameters that can then be used to project the SUTs forward, within a short-run horizon.

The paper is structured as follows. Section 2 describes the notation used and the information required to use our suggested procedure. Section 3 discusses briefly the methods available for updating of SUTs with partial information. Section 4 is the central section of the paper, describing the procedure. Section 5 presents a simple example of application of the procedure, making use of the German SUTs for year 2006. Finally, Section 6 provides some brief final remarks.

**2. Basic notions and information requirements**

In order to apply the procedure, we need to know the base year SUTs.[[1]](#footnote-1) The vectors and matrices included in them will be denoted as follows:

* *U(0)* – intermediate demand matrix *(m×n).*
* *V(0)* – output matrix *(m×n).*
* *Y(0)–* final demand matrix *(m×f).*
* *u(0) –* industry intermediate demand vector *(n×1).*
* *w(0) –* product intermediate demand vector *(m×1).*
* *v(0) –* industry value added vector *(n×1).*
* *m(0) –* product imports vector *(m×1).*
* *x(0) –* industry output vector *(n×1).*
* *q(0) –* product output vector *(m×1).*

We will denote as *i* a column matrix of unit values, as *n* the number of industries, as *m* the number of products, and as *f* the number of components of the final demand vector.

We need also some magnitudes corresponding to the year to which the SUTs will be updated. We could consider various situations regarding the availability of data, but in this case we opt for an intermediate scenario. In particular, we will make use of the following information relative to the target year:

* *x(1) –* industry output vector*.*
* *v(1) –* industry value added vector*.*

*z(1) =Y(1)´i –* components of final demand total values vector (the information needed is their variation relative to the base year).

* *i´m(1)–* products total imports (again, what we need is variation relative to the base year).

Less information would be required if the objective is not a complete updating of the SUTs. Sometimes, like e.g. for simulation exercises, only the inverse matrix needs to be updated. In this case, just one of the variation magnitudes would be required, because the other could be deducted from the rest of the information.

In order to develop a scale algorithm, this should respect he basic accounting relationships associated with the rectangular format. In any particular year, e.g. the base year, products’ total output (domestic plus imported) must be equal to the sum of intermediate and final demand:

On the supply side, industry output must be equal to the sum of intermediate and primary inputs:

**3. Global updating of SUTs with partial information**

If we know the growth rates of output, value added and the totals of the components of final demand, we can update the SUTs. The adjustment procedure is based on the accounting relationships that are characteristic to the SUTs format, and it can be applied to RAS or any other partial adjustment technique. For the sake of simplification we will not disaggregate flows according to their origin, implicitly assuming that there are no imports.[[2]](#footnote-2)

The adjustment of the intermediate and final matrices takes place in parallel, being the only constraints those imposed by the available information. Simple RAS could not be applied with this information, because the row and column sums of the intermediate demand matrix are unknown.[[3]](#footnote-3) Eurostat (2008) describes an alternative procedure based on the EU method proposed by Beutel (2002). EU is a global updating method, but it can be applied only to square matrices and, as indicated by Temurshoev *et al.* (2011), it does not always converge.

Our proposed global updating procedure estimates two paths based on two hypotheses: the stability of technical coefficients, and the respect of the accounting relationships supporting the SUTs format. The intermediate demand vector is estimated repeatedly following both ways, and their results are compared. The estimation differences obtained in each stage are distributed in function of the relative weights of the rows in the intermediate demand matrix.

We will make systematic use of correcting coefficients matrices. The elements on the main diagonal of these matrices are initially equal to the growth rates of the corresponding flows to be updated. In subsequent iterations, the latest estimations of final and intermediate demand will be used to modify the values in the main diagonal.[[4]](#footnote-4) Row constraints will be variable, while column constraints remain fixed.

Approximation paths will follow opposite senses. We will make the usual assumption that changes in the variables and in the elements of the matrices to be updated are small. Actually, if any of the instrumental rates is substantially different from unity, the results should be interpreted with caution. The process should be terminated when the differences between the intermediate demand estimations will be zero or nearly zero. By increasing the number of iterations, convergence of the various magnitudes included in the SUTs will be simultaneous.

**4. A global procedure with distribution parameters**

In each iterative stage it is necessary to assign a value to the distribution coefficient of the estimation errors. This choice is highly relevant: the updating results will be different according to the value chosen. Our suggestion is to implement a backwards test mechanism that will allow us to find which are the best values and then use them for forward projections.

In the first place, the intermediate demand matrix is estimated assuming stability of the technical (non-homogeneous) coefficients, i.e., we modify the initial matrix on the basis of the variation rates of intermediate inputs,

Actually, this is an implementation of one of the stages of the RAS method. It allows obtaining an approximation to the intermediate demand vector following the path S:

Alternatively, we can obtain the output matrix using a specialization coefficient stability hypothesis, in order to estimate intermediate demand ** starting from the output vector, obtaining . Adding the output matrix elements by row we get a product output vector:

We can estimate the final demand matrix, through a double adjustment based on the variation rate of the totals of the components of final demand. Then, the approximation of the final demand vector through path R can be expresses as:

.

The first correction by rows, proportional to the rate of change in output, will be highly influential on the final result. Anyway, advancing through the process this influence will be smoothed by other instrumental rates with cyclical incidence. Hereafter, we will note by *g* the rates of change of the various magnitudes considered.

The estimation of intermediate demand through path R is: Usually, and will not coincide. Some elements will be overvalued and others undervalued in both paths. We define the differences vector as:

where because of

Elements of this vector may be positive or negative, but their sum has to be equal to zero.

These differences have to be distributed to correct the intermediate and final demand matrices. We modify in proportion to the weight of such demands on the row totals:

The distribution coefficient (or parameter) may be constant or it may be variable. For the sake of simplification, let it be constant, so, In this case

If there is additional information available regarding the evolution within the time interval of total product outputs, this assumption could be avoided. *C* is a scalar with values between 0 and 1, both included. If the coefficient is variable, the following constraint has to be satisfied:

because the equilibrium of the system must not be altered. The scalar is the influential coefficient on the elements of the *i* row in this first stage.[[5]](#footnote-5) In general, different values for these scalars give different estimations for the intermediate demand vectors.

Actually, this is again an application of RAS, but this time in a particular version conditioned by the available information. In practice, the role of statistical institutes in the determination of the best treatment of these coefficients is paramount, as they usually possess additional unpublished information about the SUTs.

The characteristical element of is as follows:

In order to preserve the economic meaning of the estimations, intermediate demand elements must be non-negative. If the differences turn out to be negative,then

in order to avoid incoherent results. Thus, special attention must be given to values close to zero. Anyway, updating should be attempted in those situations when changes in the magnitudes are low, which in practice eliminates the possibility that these kind of problems will arise.[[6]](#footnote-6)

In order to highlight the similitude with the RAS approach, we note that the characteristic element of can be expressed in simplified form as:

where is the *i*-th element of the intermediate vector between both estimations of product *i*’s intermediate demand. In particular,

The use of an intermediate vector implies to assume that the estimation error is distributed in a particular way. Continuing with the iteration, we can observe that ≈.

Matrix can be rewritten as follows:

.

Note again that the lack of knowledge about the row sums of the intermediate demand matrix causes no problem for the application of RAS. It is not advisable to correct matrix on the basis of output estimations. Output may change in different ways than intermediate demand, and at different rates. Maybe this is the reason why the EU method cannot always warrant a convergent solution.

Focusing on the estimations of the intermediate inputs vector, ,in general they will not coincide, either. Thus, we can estimate an alternative intermediate demand matrix, in the following way:

.

Double adjustment of the final demand matrix gives an approximate vector of final demand, as the difference between output, and the previous intermediate vector, .

Thus, we obtain the matrix

,

where

So we get where .We then get an alternative estimation of intermediate demand through the other path by adding the rows of i.e.,Once again we get another difference vector:

And from this vector, we can arrive to a new estimation of the intermediate demand matrix:

or, in more abbreviated form:

 where

Advancing in the adjustment process, column corrections will be expressed as:

Finally, we can get generic expressions for matrix estimations. In the case of final demand:

,

where

Then with .The last estimation of intermediate demand was We can use the difference *)* in order to get a new intermediate demand matrix:

=.

And then we can again make a backward correction for the columns of the intermediate demand matrix:

=.

This process would be iterated until the differences between estimations of intermediate demands by product were zero or almost zero.

If we want to keep the differentiation between domestic and imported flows, we have to introduce some modifications in the procedure. Adjustments of the domestic and imported intermediate demands, and respectively, can be made jointly. We consider the intermediate demand matrix, as composed of another two matrices: y . The same is true of the final demand matrix, which includes and . In general, and

If our aim is to update the domestic intermediate demand matrix, it is advisable to aggregate imported flows into one single vector. In this case, the above matrices will contain *m+1* rows, with the last row corresponding to the column sums of imported intermediate and final demands. Thus, we have:

 and

Also, the outputs vector, is composed of and And intermediate demand, is composed of and (*m+1* vectors), while by industries

**5. An example**

In order to illustrate the workings of our suggested procedure, we will update the intermediate demand matrix of the German supply table, taking 2006 as the base year. The backward calibration will be made for the year 2005, and the forward projection for the year 2007. The table used is square, but the procedure would work in a similar way for a rectangular table.

The ideal tables for this procedure would be survey-based SUTs. But these are built with a low (usually at least five-year) periodicity. Tables corresponding to consecutive years usually include both survey and non-survey ones. Calibration procedures, then, implicitly discover the adjustment tool used by the statistical institute that built the tables (usually RAS or any other bi-proportional technique).

Once the matrices are updated and tested, we will detect the right coefficient value by computing the Weighted Absolute Percentage Error (WAPE)[[7]](#footnote-7) for the intermediate demand matrix depending on the values assigned to the characteristic parameters of the proposed algorithm, as in Figure 1.

Source: Own elaboration from EUROSTAT

**Figure 1.** WAPEs associated with different distribution coefficients (2006→2005)

We can see that when the value of the coefficient diminishes, the error increases. Thus, in the hypothetical situation where the coefficient has a value of 0, the corresponding value of WAPE is 6.0.[[8]](#footnote-8)

Widening the scale on the vertex of the curve shown in Figure 1 in order to see with more precision which is the coefficient that minimize errors, we get Figure 2:

Source: Own elaboration from EUROSTAT

**Figure 2.** WAPEs in the surroundings of the coefficient value 0.95 (2006→2005)

It is easy to confirm visually that the coefficient value giving the best results is 0.950. Thus, for the *p-*th iteration we get that the intermediate vector considered in the adjustment by rows:



is closer to the estimation of intermediate demand obtained through path R, *wR(p)*. In other words, intermediate demand presents higher instability than final demand.

In order to project forward the intermediate matrix from 2006 to 2007 we now use only coefficient values close to 0.950. The corresponding WAPEs are shown in Figure 3:

Source: Own elaboration from EUROSTAT

**Figure 3.** WAPEs associated with different distribution coefficients (2006→2007)

In this projection WAPEs are slightly lower than in the backward projection. The optimum coefficient value is very close to the value found in the backward projection, i.e. 0.959. This will be the value that would provide the best updating within the limits imposed by the available information.

Finally, there is an almost absolute coincidence in the distribution coefficient value that minimizes estimation errors both backwards and forward in time. When dealing with total flows, it seems this value will be close to 1, indicating that the technical coefficient matrix shows less stability than the final demand matrix. When dealing with domestic flows, instead, the minimizing value could be in a different range, due to the possibility of substituting imports for domestic products and vice versa.

An interesting feature of these estimations is convergence. Even if 12 iterations were needed to get a total adjustment using the coefficient value 0.950, approximation is easily visualized already in the first iterations. Distances between paths R and S in the estimations for year 2005 diminish very quickly, as shown in Figure 4.[[9]](#footnote-9)

Source: Own elaboration

**Figure 4**. Convergence in intermediate demand estimations

The number of iterations needed is sensible to the value of the distribution coefficient. In general, if the coefficient diminishes its value, more iterations are needed.

In SUTs updating, it is crucial to warrant convergent results. Our suggested procedure allows simultaneously an optimal exploitation of available information and a convergent adjustment process, both highly useful properties.[[10]](#footnote-10)

**6. Final Remarks**

The value assigned to the distribution coefficient of estimation errors has a strong influence on the results of SUTs-updating procedures. This makes the choice of the right value highly relevant. We proposed in this paper the implementation of a mechanism of backward testing to detect the right values that then can be used for short-run forward projections. The mechanism is simple and frugal in information requirements (but note that the existence of survey-based SUTs is necessary).

It is easy that some incongruence may arise in the row corrections, especially when intermediate demand of a product is zero. Some typical examples are products like Public Administration, Defence, Social Security or Non-Market Activities. These exceptions require a specific procedure. The best one is to exclude known data about intermediate or final demand matrices from the iterative process.[[11]](#footnote-11) This procedure, similar to the one used in modified RAS, makes computational programming of the algorithm easier and avoids *ad hoc* manipulations. Implementation is also easy: SUTs have to be decomposed into a sum of two matrices, one including known data about intermediate and final demands, the other including unknown data, and updating has to proceed with just the second one, adding known data at the end of the process.

**REFERENCES**

Allen, R.; Lecomber, J. (1975) Some Test on a Generalized Version of RAS, in Allen, R.; Gossling, W. [eds.]: Estimating and Projecting Input-Output Coefficients. London, Publishing Company.

Bacharach, M. (1970) *Biproportional Matrices and Input-Output Change.* Cambridge, Cambridge University Press.

Bachem, A.; Korte, B. (1979) *On the RAS-algorithm*.Computing, Vol. 23, pp. 189-198.

Beutel, J. (2002) *The Economic Impact of Objective 1 Interventions for the Period 2000-2006.* Informe para la Dirección General de Política Regional, Konstanz.

Callealta, F.; López, A. (2005) “Predicciones Armonizadas del Crecimiento Regional: Diseño de un Modelo de Congruencia”*.* *Estadística Española*, Vol. 47, nº 159, pp. 219-251.

Eurostat (2008) *Updating and Projection Input-Output Tables*. Luxemburg, Office for Official Publications of the European Communities.

Jackson, R.; Murray, A. (2004) “Alternative Input-Output Matrix Updating Formulations”. *Economic System Research*, Vol. 16, nº 2, pp. 135-148.

Lahr, M.L.; Mesnard, L. de (2004) “Biproportional Techniques Input-Output Analysis: Table Updating and Structural Analysis”. *Economic Systems Research,* Vol. 16, nº 2, pp. 115-134.

Mínguez, R.; Oosterhaven, J.; Escobedo, F. (2009) “Cell-corrected RAS Method (CRAS) for Updating or Regionalizing an Input-Output Matrix”. *Journal of Regional Science*. Vol. 49, pp. 329-348.

Pavia, J.M.; Cabrer, B.; Sala, R. (2009) “Updating Input-Output Matrices: Assessing Alternatives through Simulation”. *Journal of Statistical Computation and Simulation*. Vol. 79, nº 2, pp. 1467-82.

Stone, R.; Brown, A. (1962) *A Computable Model of Economic Growth*. London, Chapman and Hall.

Szyrmer, J. (1989) Trade-Off Between Error and Information in the RAS Procedure, in Miller, R.; Polenske, K.; Rose, A. [eds.]: Frontiers of Input-Output Analysis. New York, Oxford University Press, pp. 258-278.

Temurshoev, U.; Webb, C.; Yamano, N. (2011) “Projection of Supply and Use Tables: Methods and their Empirical Assessment”. *Economic Systems Research*. Vol. 23, nº 1, pp. 91-123.

1. We do not make distinctions between domestic and imported flows, but they can be made without making substantial changes to the procedure. [↑](#footnote-ref-1)
2. But note that the procedure can deal easily with the distinction between domestic and imported flows. [↑](#footnote-ref-2)
3. See Temurshoev *et al.* (2011). [↑](#footnote-ref-3)
4. We will call these estimations *instrumental rates*. [↑](#footnote-ref-4)
5. See the exposition by Callealta & Lopez (2005) about modification of the RAS technique proposed by Bachem & Korte (1979). [↑](#footnote-ref-5)
6. A different but related question may arise due to the monetary units in which SUTs are expressed. If numbers represent, e.g., thousands of units, rounding to zero may introduce significant alterations in the estimations. [↑](#footnote-ref-6)
7. WAPE formulation and its interpretation can be found in Minguez et al. (2009). [↑](#footnote-ref-7)
8. Note that coefficients with the extreme values of 0 and 1 correspond to the use of simple RAS taking one or the other of the intermediate demand vectors obtained in the first iteration. [↑](#footnote-ref-8)
9. Note also that over- and under-estimation appear to be fairly balanced. [↑](#footnote-ref-9)
10. Estimations of the intermediate demand matrix can be expressed alternatively as infinite matrix products, allowing for an easier analysis of convergence. [↑](#footnote-ref-10)
11. If what is intended is to update only the intermediate demand matrix it is advisable to work with a single final demand vector. [↑](#footnote-ref-11)