

TEMPORAL PROJECTION OF AN INPUT-OUTPUT TABLES SERIES FOR THE REGION OF ASTURIAS

Ramos Carvajal, Carmen⁽¹⁾

Fernández Vázquez, Esteban⁽²⁾

Departamento de Economía Aplicada. Universidad de Oviedo.

(1) cramos@correo.uniovi.es

(2) evazquez@correo.uniovi.es

ABSTRACT

The interest on regional economics has strongly progressed in the latest decades; however, and in spite of the great extension of the information that statistical sources of data offer to economic researchers, one of the major problems arising with these kind of studies still keep being the lack of data.

Regional studies can be focused from several points of view, one of these is the input-output framework. This method allows the economic researcher to analyse the intersectoral relationship underlying on a economy in a joint way, as well as the aggregate demand, so it makes feasible to get a integrated knowledge of economic activity. This technique is a key part on the knowledge of a region, because supplies the necessary information to study the economic situation of the mentioned region. However, its use now at days is quite restricted because the discontinuity in the publications of the input-output (IO) tables by statistical agencies, as well as the long time is necessary to wait between each published table. Specifically, focusing the problem on the region of Asturias (north of Spain), the last available input-output table was made for 1995, having been published in 1999. Having account this situation, we believe that for an effective and rigorous application of this kind of analysis a previous step we must take is the input-output tables series estimation.

The input-output table elaboration is a work that implies a great effort to get statistical information, as well as a very high cost on every kind of resources. So, the IO tables obtaining by the use of indirect methods of estimation would reduce the needs for information and material and human resources.

The information theory is being applied in the latest decades as a very flexible tool that allows to estimate the IO tables coefficients when the available data are not perfectly consistent.

Our aim in this work is to make a comparison between the cross entropy method, an information theory derived technique, and the biproportional RAS method, with a long tradition on applied works (Robinson et al. 1998, Mc Dougall, 1999). On a second stage, and basing on the conclusions for the previous comparison and our information availabilities, we will estimate the input-output tables series corresponding to the region of Asturias for the years 1995-2000. For this estimation, we will apply a cross-time analysis taking as starting point the IOT for Asturias in 1995.

1. INTRODUCTION AND MOTIVATION

In spite of the progress of regional economic studies in the last decades, in Spain the statistical information they are based on still presents important deficiencies.

A very useful tool in regional economic analysis is the input-output approach, which allows researching the relationship between economic sectors as well as their final demand. However, this approach has a very restrained use because of the discontinuity in the publication of input-output tables (IO tables), in addition to the lengthy time lag between each table published¹. For these reasons, in search of an efficiently and rigorously applied use of this kind of analysis, a previous step must be the estimation of a sufficiently large series of IO tables.

The estimation process of these tables may be based on a spatial or temporal approach. The former lies in the assumption of equality between the national and regional economic structures, while the latter takes the regional economic structure as a starting point but in a moment prior to the estimation.

In preceding works², it has been empirically proven that the use of temporal information, if we wish to make a short-run prediction, generally provides more accurate estimates than the spatial approach. This result appears because the economic structures suffer very little changes in short intervals of time; for this reason, a better fit can be achieved than translating the national economic pattern for a region.

Another important fact, when the objective is the estimation of IO tables, is the choice of a good estimation method that allows us to reach that goal³. A classical and very often used method for this kind of problem is the RAS technique, which presents

good behavior from both the theoretic and applied viewpoint. On the other hand, the mathematical information theory has been being applied in the last decades as a very flexible instrument for reaching coefficient estimations when the available data are not perfectly consistent.

In this work, we pursue a double objective: first, we will make a comparison between the cross entropy method (CE), an information theory derived technique, and the biproportional RAS method, with a long tradition of applied works (Robinson et al. 1998, McDougall, 1999). In a second stage, and based on the conclusions of the previous comparison and the information available, we will estimate the input-output table series corresponding to Asturias for the years 1995-2000, following a temporal approach.

2. COMMENTS ABOUT THE ESTIMATION METHODS OF COEFFICIENTS

In this section, comments about the fit techniques pointed out above, RAS and CE, will be formulated. In both cases, we focus our attention on the temporal approach.

2.1 THE RAS METHOD

This technique, owing to Richard Anthony Stone (1962), needs the knowledge of an initial coefficients matrix as the starting point to estimate a new matrix referred to at a later moment, for which the sum of its rows (columns) is known. In general, this method lies in changing the starting matrix, multiplying it by adjustment coefficients over rows as well as columns. So, the sums (horizontal and vertically) of the estimated matrix elements are the “closest” ones to the real values. This technique will be explained later on in detail⁴.

The set of three elements $(\mathbf{A}, \mathbf{x}, \mathbf{y})$ is called *problem*, where \mathbf{A} is a $n \times n$ matrix and $\mathbf{x}(X_1, X_2, \dots, X_n)$ and $\mathbf{y}(Y_1, Y_2, \dots, Y_n)$ are positive element vectors⁵ with $n \times 1$ and $1 \times n$ as respective dimensions. The first stage of the RAS method lies in fitting the starting matrix by changes on the rows. Let us consider at this stage $t=0$, so \mathbf{A}^0 is the starting matrix of the estimation process. Now, the first adjustment will be $\mathbf{A}^1 = \hat{\mathbf{R}}^1 \mathbf{A}^0$, where the $\hat{\mathbf{R}}^1$ elements are:

$$r_i^1 = \frac{X_i}{\sum_j a_{ij}^0}$$

and $\mathbf{A}^1 \mathbf{i} = \mathbf{x}$, where \mathbf{i} is a row vector of ones and $\hat{\mathbf{R}}^1$ is a diagonal matrix with r_i^1 elements in the main diagonal and all other entries equal zero. The next step will be to change the resulting matrix from the previous step by columns: $\mathbf{A}^2 = \mathbf{A}^1 \hat{\mathbf{S}}^1$, where the main diagonal elements of the (diagonal) matrix $\hat{\mathbf{S}}^1$ are now:

$$s_j^1 = \frac{Y_j}{\sum_i a_{ij}^1}$$

From these expressions, we can write $\mathbf{A}^2 = \hat{\mathbf{R}}^1 \mathbf{A}^0 \hat{\mathbf{S}}^1$, with $\mathbf{i}' \mathbf{A}^2 = \mathbf{y}$. This process will be successively repeated. In general, the changes by rows will be obtained from the expression $\mathbf{A}^{2t+1} = \hat{\mathbf{R}}^{t+1} \mathbf{A}^{2t}$, where the elements of matrix $\hat{\mathbf{R}}^{t+1}$ will be:

$$r_i^{t+1} = \frac{X_i}{\sum_j a_{ij}^{2t}}$$

with $\mathbf{A}^{2t+1} \mathbf{i} = \mathbf{x}$.

Regarding column adjustments, we will have $\mathbf{A}^{2t+2} = \mathbf{A}^{2t+1} \hat{\mathbf{S}}^{t+1} = \hat{\mathbf{R}}^{t+1} \mathbf{A}^{2t} \hat{\mathbf{S}}^{t+1}$, where the elements of matrix $\hat{\mathbf{S}}^{t+1}$ will be:

$$s_j^{t+1} = \frac{Y_j}{\sum_i a_{ij}^{2t+1}}$$

and where $\mathbf{i}' \mathbf{A}^{2t+2} = \mathbf{y}$, with $a_{ij}^{2t}, a_{ij}^{2t+1} \geq 0, \forall t=1,2,\dots$

$\hat{\mathbf{R}}^{t+1}$ and $\hat{\mathbf{S}}^{t+1}$ are diagonal matrices with r_i^{t+1} and s_j^{t+1} , respective coefficients in the main diagonal. Given a problem $(\mathbf{A}, \mathbf{x}, \mathbf{y})$, we can say that \mathbf{A}^* is the problem's solution if and only if $\mathbf{A}^* = \lim_{h \rightarrow \infty} \hat{\mathbf{R}}^{h+1} \mathbf{A}^{2h} \hat{\mathbf{S}}^{h+1}$, with $\mathbf{A}^* \mathbf{i} = \mathbf{x}$ y $\mathbf{i}' \mathbf{A}^* = \mathbf{y}$. It can be proved that "if $(\mathbf{A}, \mathbf{x}, \mathbf{y})$ is a feasible⁶ problem, then it has a solution and this is the only one".

2.2 THE CROSS ENTROPY METHOD

In the last decades, the mathematical information theory has been being applied as a useful tool in the input-output coefficient estimation. One of the measures from this approach is the cross entropy that is based on Kullback-Leibler's divergence (1951). In developing this approach, let us start from an experiment that has n possible outcomes (states) E_1, E_2, \dots, E_n with respective probabilities q_1, q_2, \dots, q_n , which will be considered as prior probabilities. Let us suppose that new information embodied in a message is

received, which implies that prior probabilities are transformed into posterior probabilities p_1, p_2, \dots, p_n . If the message is just an outcome E_i , the amount of received information will be $-\ln(p_i)$. Therefore, because of this new information for each outcome E_i , an “information increase” will be observed and it will be measured as:

$$-\ln \frac{p_i}{q_i} = -[\ln p_i - \ln q_i]$$

So, the expected information can be defined as:

$$-I(p,q) = -\sum_{i=1}^n p_i \ln \left(\frac{p_i}{q_i} \right)$$

where $I(p,q)$ is the Kullback-Leibler’s divergence measure. This measure allows us to quantify the distance between two probability distributions: before and after the new message (information) is received.

Starting from the previous expression, Golan, Judge & Robinson (1994) suggest the following measure to estimate the IO table coefficients:

$$\sum_i \sum_j a_{ij} \ln \left[\frac{a_{ij}}{a_{ij}^0} \right]$$

where a_{ij} represents a (normalized by columns) coefficient from the matrix that is going to be estimated and a_{ij}^0 is a generic coefficient from the matrix taken as a starting point.

Now, the problem is minimizing the “distance” between the initial matrix and the one we wish to estimate, subject to a restriction set. So, the non-linear program will be:

$$\begin{aligned} \text{Min} \quad & \sum_i \sum_j a_{ij} \ln \left[\frac{a_{ij}}{a_{ij}^0} \right] \\ \text{s.t.} \quad & \sum_j a_{ij} Y_j = X_i \\ & \sum_i a_{ij} = 1 \\ & 0 \leq a_{ij} \leq 1 \end{aligned}$$

where X_i is the by row and Y_j is the by column respective “real” sums. That is, the first restriction group guarantees that the estimated coefficients are consistent with observed by row and by column sums, and the remaining restrictions refer to the average expense properties (taking values between zero and one). The program's result is the following⁷:

$$a_{ij} = \frac{a_{ij}^0 \exp(\lambda_i Y_j)}{\sum_i \sum_j a_{ij}^0 \exp(\lambda_i Y_j)}$$

where λ are the Lagrange restriction multipliers and the quotient's denominator is a normalization factor.

2.3 RAS AND CROSS ENTROPY: A COMPARISON

In this section, a comparison between both methods of estimation will be made. The comparison will be developed for three points or approaches: the incorporated information approach comparison, the analytic approach and the applied approach.

2.3.1 A COMPARISON FROM THE INCORPORATED INFORMATION APPROACH

In applying any of the previously explained estimation methods (RAS and CE), two kinds of information are needed: on the one hand, prior information that may be national (spatial approach) as well as regional (temporal approach), and on the other, specific regional data relating to the moment when one wishes to estimate the IO tables. In applying a spatial approach and considering prior information, it will be embodied in the IO tables holding the flows and relationship at the national level in the same t period where one wants to make the estimation, that is, NIO TABLES _{t} . Under the temporal approach, the starting point matrix would be regional IO tables (RIO tables), relating to a previous moment $t-k$, that is to say, RIO TABLES _{$t-k$} . In this work, based on the reasons explained before, this last approach will be used.

Briefly, both methods need quite similar information: the starting matrix that can be spatial as well as temporal, and the knowledge of by row and by column sum vectors (\mathbf{x} and \mathbf{y}).

Robinson et al. (2000) point out the major flexibility of the CE method as opposed to the RAS method in the following way: with the CE method one is allowed to include additional information or to make estimations, even with a lack of data (information).

Let us suppose that a few intersectoral flows are known in the estimating period t , and that these are some of the IO table entries. This information can be included using additional equality restrictions. Assuming k restrictions, we will obtain:

$$\sum_i \sum_j g_{ij}^k z_{ij} = \gamma^k$$

where g_{ij} are the elements of matrix G , which will be one or zero. They will be one in the respective entries of known values and zero in other cases. The elements z_{ij} are flow values between the i and j sector, while γ^k is the known value of the aggregate.

Similarly, perhaps the by row or column sum values are not precisely known. In this case, inequality restrictions will be added to the problem considering margins bounding the values, that is:

$$X_i^1 \leq \sum_j z_{ij} \leq X_i^2$$

where X_i^1 and X_i^2 are the margins for the X_i value.

We consider that, even though the “basic” RAS method does not allow for the inclusion of additional information to realize estimations with incomplete information, these difficulties can be overcome by applying some of the extensions of the RAS method. So, the extension proposed by Paelinck & Waelbroeck can be pointed out because this method implies the incorporation of known matrix coefficients, specifically, most problematic coefficients, that is to say, the ones with a major “divergence” with regard to the real ones.

Being \mathbf{A} a matrix that has been partitioned into submatrices \mathbf{A}_{11} , \mathbf{A}_{12} , \mathbf{A}_{21} and \mathbf{A}_{22} ; that is:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

considering that, for the sake of simplicity, the elements of submatrix \mathbf{A}_{22} are known. A residual matrix \mathbf{E} is built as the difference between \mathbf{A} and \mathbf{C} :

$$\mathbf{A} - \mathbf{C} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}$$

where in matrix \mathbf{C} all entries are zero except for \mathbf{A}_{22} composed of known coefficients. So, matrix \mathbf{E} is formed by initial coefficients excluding the problematic ones, which will have been replaced by zeros⁸. The next step is to apply RAS in matrix \mathbf{E} and, finally, the matrix is estimated as $\mathbf{A}^* = \mathbf{E}(\text{RAS}) + \mathbf{C}$. Therefore, an estimated element matrix will be available, which is formed by submatrices \mathbf{A}_{11} , \mathbf{A}_{12} and \mathbf{A}_{21} .

Another extension of the RAS method is proposed by Hitz and Schmid (1978). In this method, limits are set up for the margins without perfectly known values. A set of restrictions is included as follows:

$$\begin{aligned}
 Y_j^1 &\leq \sum_i z_{ij} \leq Y_j^2 \\
 X_i^1 &\leq \sum_j z_{ij} \leq X_i^2 \\
 X^1 &\leq \sum_i \sum_j z_{ij} \leq X^2
 \end{aligned}$$

where $Y_j^1(Y_j^2)$ represent column margins, $X_i^1(X_i^2)$ row margins and $X^1(X^2)$ represent global limits. So, if some of the extensions of the RAS method are taken into account, properties similar to the CE method can be noticed.

2.3.2 ANALYTIC COMPARISON

Considering the comparison from an analytic approach, McDougall (1999) proves that the RAS estimation method is equivalent to a minimization model of cross entropy. From the following expression:

$$\sum_j \frac{Y_j}{X} EC_j = \sum_i \sum_j \frac{Y_j}{X} \frac{z_{ij}}{Y_j} \ln \left(\frac{z_{ij}/Y_j}{z_{ij}^0/Y_j^0} \right) = \sum_i \sum_j \frac{z_{ij}}{X} \ln \left(\frac{z_{ij}/Y_j}{z_{ij}^0/Y_j^0} \right)$$

it is possible to prove that minimizing this objective function is equal to minimizing:

$$\sum_i \sum_j d_{ij} \ln \left(\frac{d_{ij}}{c_{ij}} \right)$$

where $d_{ij} = \frac{z_{ij}}{X}$ and $c_{ij} = \frac{z_{ij}^0}{X^0}$.

With the first order conditions, the following expression is obtained:

$$\ln \frac{d_{ij}}{c_{ij}} = -1 - \lambda_i - \mu_j$$

and with the right mathematical transformations, $d_{ij}=r_i c_{ij} s_j$ is yielded, the RAS method solution, where $r_i = e^{-1-\lambda_i}$ and $s_j = e^{-\mu_j}$.

So, we conclude that the RAS method can be seen not as a technique different from CE, but as a particular case of this method using normalized matrix coefficients.

2.3.3 APPLIED COMPARISON

We thought it would be interesting to conclude this analysis by making an applied comparison between the RAS and CE method results in IOT estimation. Specifically, the 1995 IOT for Asturias has been estimated because it is the last one published for this region. We have taken a temporal approach, which is why the 1990 IOT for Asturias is the starting point of the process.

The first step is to realize that this comparison lies in previous information processing in search of homogenizing data. The 1995 IOT for Asturias has suffered a significant methodological change, because it was elaborated under the 1995 European System of Accounts (1995 ESA). As a consequence of this change, the sector number and composition have been modified. So, in the 1990 table (AIOT-90), one can distinguish 50 sectors while the 1995 table (AIOT-95) is classified into 60, 31, 16 and 4 sectors. For simplicity's sake, we have chosen to use the Hermes classification and to work with nine aggregated sectors. However, two problems arise: the first one refers to the manufacturers of furniture and other products sector, which includes three subsectors (manufacturers of metal products, manufacturers of wood products and manufacturers of other products). The second one, the research and development sector and the health activities sector, is not included in commercial services and non-commercial services, as occurs in the 1985 IOT. In both cases, a similar information treatment has been applied, quantifying the average weight of manufacturers of metal products subsector with respect to the other three in the 1985 and 1990 tables, keeping this weight for 1995. Following this approach, the weights for the education and (commercial and non-commercial) research activities sectors and the health activities sectors (commercial and non-commercial) have been unchanged in relation to their values in past tables. The following chart shows this aggregation:

Chart 1. Classification Correspondences between AIOT-90 and AIOT-95 under the Hermes Classification

HERMES SECTORS	AIOT-90 Sectors	AIOT-95 Sectors
Agriculture (A)	1-2	1-2-3
Energy (E)	3-4-5-6-7	4-5-17-32-33
Manufacturers of Intermediate Products (Q)	8-9-10-11-12-13-14-15	6-7-8-18-20-21
Manufacturers of Equipment Products (K)	16-17-18-19-20-21-22	22-23-24-25-26-27-28-29-30 (partially)
Manufacturers of Consumption Products (C)	23-24-25-26-27-28-29-30-31-32-33-34	9-10-11-12-13-14-15-16-19-30 (partially)
Construction (B)	35	34
Transport & Communications (Z)	39-40	39-40-41-42-43
Commercial Services (L)	36-37-38-41-42-43-44-45-46	31-35-36-37-38-44-45-46-47-48-49-50 (partially)-51-53 (partially)-54 (partially)
Non Commercial Services (G)	47-48-49-50	50 (partially)-52-53 (partially)-54(partially)-56-57-58-59-60

Once the information is homogenized as explained, the AIOT-95 is estimated using the RAS as well as the CE method. To analyze the proximity (similarities) of our estimations, we have applied the following measure proposed by Le Masne:

$$S^{R-C} = 100 \left(1 - 0.5 \sum_i |a_{ij}^R - a_{ij}^C| \right)$$

where (a_{ij}^R) are the estimated coefficients by the RAS method and (a_{ij}^C) are the ones estimated by the CE method. This measure is bounded between 0 and 100, and the more similar the coefficients are, the closer the measure is to 100. The results obtained are shown in Chart 2:

Chart 2. Analysis of the Similarity in the Estimations

Sector	A	E	Q	K	C	B	Z	L	G
S ^{R-C}	95.98	99.32	98.50	97.57	98.80	97.23	94.21	96.63	94.76

Thus, the high similarity between the RAS and the CE methods can be noted. The use of numerical methods in the estimation process is the reason for the observed differences.

3. ESTIMATION OF IO TABLE SERIES FOR ASTURIAS

Taking into account the previous information processing, an IO table series can be estimated in this section. Specifically, we have obtained tables for the 1995-2000 period. Before making this estimation, let us consider a few notes on the statistical information processing applied.

3.1 ESTIMATION OF ROW AND COLUMN MARGINS

To estimate the IO table series, one needs to know the row and column sums of intermediate consumptions and the technical coefficient matrix that will be used as a starting point in the estimation process.

The Gosh model (1958) can be approached from the following equation:

$$\mathbf{w}=(\mathbf{I}-\mathbf{D}')^{-1}\mathbf{g}$$

where \mathbf{w} is the effective production vector, \mathbf{D} is the market coefficient matrix and \mathbf{g} is the gross value added to the purchase prices. Once the effective production is calculated, it is feasible to obtain the column sums of intermediate consumptions (\mathbf{v}), just considering the following equation:

$$\mathbf{w}=\mathbf{v}+\mathbf{g}$$

from which we can obviously state that $\mathbf{v}=\mathbf{w}-\mathbf{g}$. On the other hand, and taking into account the Tilanus (1966) model, it is possible to derive the equation that follows:

$$\mathbf{q}=(\mathbf{I}-\mathbf{A})^{-1}\mathbf{d}$$

where $(\mathbf{I}-\mathbf{A})^{-1}$ is the Leontief inverse matrix. Reordering the terms in the previous expression we obtain:

$$\mathbf{d}=(\mathbf{I}-\mathbf{A})\mathbf{q}$$

and so, the row sums of intermediate consumptions (\mathbf{u}) are obtained from the equation $\mathbf{u} = \mathbf{q} - (\mathbf{I} - \mathbf{A})\mathbf{q} = \mathbf{A}\mathbf{q}$.

That is to say that $\mathbf{u}^{t+h} = \mathbf{A}^{t+h}\mathbf{q}^{t+h}$, and from this, following Tilanus, we can state that $\mathbf{u}^{t+h} = \mathbf{A}^t\mathbf{q}^{t+h}$. In other words, the row sums of intermediate consumptions will be obtained using the expression $\mathbf{u}^{1995+h} = \mathbf{A}^{1995}\mathbf{q}^{1995+h}$, $\forall h=1,2,\dots,5$. However, and because the \mathbf{q} values for time periods after 1995 are unknown, they must be previously estimated. For this, we use the following assumption: the total output / gross value added ratio stands unchanged over the mentioned period. That is,

$$\frac{q_i^{1995}}{g_i^{1995}} = K_i \quad \forall i = 1, 2, \dots, 9$$

and so:

$$\frac{q_i^{1995}}{g_i^{1995}} = K_i \quad \forall i = 1, 2, \dots, 9$$

Following this method, all the information necessary for the estimation process has been obtained. However, a last step in the statistical information process must be undertaken. Since we are going to estimate the IO tables from 1995 to 2000, it is essential to deflate these matrix entries beforehand since we are working with only a few years and the transactions among sectors are expressed in current monetary units (Barriga, 1992; Pulido & Fontela, 1993).

3.2 OBTAINING A DEFLATOR SERIES

The technical coefficients are quotients expressed in monetary values, and the numerator and denominator values come in different prices. Let us consider the production structure for the initial year in the following terms:

$$a_{ij}^{*0} = \frac{z_{ij}^0}{z_j^0} = \frac{x_{ij}^0 p_i^0}{X_j^0 p_j^0} = \frac{p_i^0}{p_j^0} a_{ij}^0$$

where a_{ij}^{*0} is the technical coefficient of the initial year (in current monetary values), p_i^0 and p_j^0 are the goods and service prices for the same year, x_{ij}^0 is the transaction values between sectors i and j , and finally, X_j is the total input for sector j .

The technical coefficients for the following year would be:

$$a_{ij}^{*1} = \frac{z_{ij}^1}{z_j^1} = \frac{x_{ij}^1 p_i^1}{X_j^1 p_j^1} = \frac{p_i^1}{p_j^1} a_{ij}^1$$

where all the terms refer to this year.

It seems reasonable to compare the coefficients in constant prices, because the observed variation among years may be caused by price changes only, without an economic structure change. That is,

$$\Delta a_{ij}^* = \Delta \left(\frac{p_i}{p_j} \right) \Delta a_{ij}$$

The deflators used⁹ refer to intermediate consumption, the *Paid to Farmers Prices Index* for Agriculture; the *Residential Construction Investment Deflator* for Construction; the corresponding *Gross Value Added (GVA) at Purchase Price Deflators* for Transportation & Communications, Commercial Services and Non Commercial Services; the corresponding *Industrial Price Index* for Energy, the Manufacture of Intermediate, Consumption and Equipment Products. The GVA's have been deflated using the implicit deflators.

3.3 ESTIMATION OF THE IO TABLE SERIES

To perform the estimation of the IO table series, it was essential to use information from the Regional Accounts of Asturias (elaborated by SADEI, the regional statistical office) and from the Spanish Regional Accounts (SRA, elaborated by the Spanish Statistics Institute, INE). For this reason, it was necessary to aggregate the sectors included in the SRA, following the Hermes classification. Chart 3 shows this aggregation:

Chart 3: SRA Sectors according to the Hermes Classification

HERMES SECTORS	SRA SECTORS
Agriculture (A)	01
Energy (E)	06
Manufacturers of Intermediate Products (Q)	13-15-17
Manufacturers of Equipment Products (K)	24-28
Manufacturers of Consumption Products (C)	36-42-47-50
Construction (B)	53
Transport & Communications (Z)	60
Commercial Services (L)	58-69-74
Non Commercial Services (G)	86

The necessary gross values added to purchase prices to estimate the series have been obtained starting from the 1995 value multiplied by the growth rate of this value every year according to SRA data. From this quantity and using the model proposed in the previous section, row and column margins have been calculated. Once these data are available, the IO table series can be estimated.

For that goal, we have used the EViews software. This software allows the creation of specific program to estimate the IO tables. Taking the AIOT-95 as a starting point, the row (\mathbf{u}) and column (\mathbf{v}) intermediate consumption margins and production (\mathbf{w}) are the ones corresponding to Asturias for this year. The convergence criterion used is the one proposed by Pedreño & Muñoz (1986):

$$\left| \frac{u_i(1) - u_i^k}{u_i(1)} \right| < \varepsilon, \forall i \quad \text{and} \quad \left| \frac{v_j(1) - v_j^k}{v_j(1)} \right| < \varepsilon, \forall j$$

where \mathbf{u}^k and \mathbf{v}^k are the estimated vectors in the last iteration. The ε value has been set at 10^{-11} ; that is, the relative differences between real and estimated vectors, measured in absolute values, must be smaller than this arbitrary and rather small number. The estimation results are included in Appendix 1.

4. CONCLUSIONS

This work compares the two most frequently used methods in IO analysis estimation: the RAS and the CE method. The CE method is a more flexible technique than the “basic” RAS method, because it allows the inclusion of information related to some of the matrix entries, such as the use of row or column sums (even when these are not exact). However, this potential advantage may be reduced if some of the extensions of the RAS method are considered, such as the ones proposed by Paelinck & Waelbroeck or by Hitz & Schmid. When this occurs, both methods behave similarly.

On the other hand, McDougall proves that the RAS and the CE techniques are strongly connected. Even more so, the RAS technique may be considered a particular CE case when starting from normalized matrix coefficients. For the sake of comparing the results from both methods, the IO table of Asturias for 1995 has been estimated. The temporal approach has been chosen and, so, the IO table for 1990 is our starting point. The comparison has been undertaken using the Le Masne measure, which shows the results are quite similar, as we can expect.

Finally, we have estimated an IO table series for the period 1995-2000. Therefore, a sufficiently wide set of information is available to be used in future works on our regional economy.

5. REFERENCES

- ALVAREZ, R (2001): *Métodos de estimación indirecta de coeficientes input-ouput: una aplicación a la comarcalización de tablas*. Trabajo de Investigación, Universidad de Oviedo.
- BACHARACH, M. (1970): *Biproportional matrices and input-output change*. Cambridge University Press.
- FERNÁNDEZ, E. y C. RAMOS (2001): Técnicas de estimación de matrices de contabilidad social: una comparación. *XV Reunión ASEPELT España*, La Coruña.
- GOLAN, A., G. JUDGE y D. MILLER (1996): *Maximum Entropy Econometrics, Robust Estimation with Limited Data* (John Wiley & Sons).
- GOLAN, A., G. JUDGE y S. ROBINSON (1994): Recovering information from incomplete or partial multisectoral economic data. *The Review of Economics and Statistics*. N° 76, pp. 541-549.

- HITZ, P. y B. SCHMID (1978): *Computerprogram Entrop M Studienunterlagen zur Orts-Regional und Landesplanung OR-L-Institut, ETM, Zurich.*
- INE (Several years): *Contabilidad Regional de España.*
- MCDOUGALL, R. A. (1999): *Entropy Theory and RAS are friends.* Discussion Paper N° 6, Global Trade Analysis Project.
- PAELINCK, J. y J. WAELBROECK (1963): *Etude Empirique sur l'Evolution de Coefficients Input-Output. Essai d'application de la Procédure RAS de Cambridge au Table Interindustrial Belgue. Économie Appliquée.* Vol. XVI, N° 1, January-March.
- PARDO, L. (1997): *Teoría de la información estadística.* Hespérides.
- PEDREÑO, A. (1986): *Deducción de las tablas input-output: consideraciones críticas a través de la contrastación "survey-nonsurvey". Investigaciones Económicas,* Vol. X, N° 3, pp. 579-99.
- PULIDO, A. y A. FONTELA (1993): *Análisis input-output. Modelos, datos y aplicaciones.* Pirámide.
- RAMOS, C., M.J. PRESNO y R. PÉREZ, (1999): *Estimación de tablas input-output: un enfoque espacial-temporal. XIII Reunión ASEPELT España,* Burgos.
- RAMOS, C. y E. FERNÁNDEZ (2001): *La teoría de la información un nuevo enfoque en la estimación de matrices de contabilidad social, XXVII Reunión de Estudios Regionales,* Madrid, 2001.
- ROBINSON, S., A. CATTANEO y M. EL-SAID, (1998): *Estimating a Social Accounting Matrix Using Cross Entropy Methods.* Discussion Paper N° 33, Trade and Macroeconomics Division, International Food Policy Research Institute, USA.
- ROBINSON, S., A. CATTANEO y M. EL-SAID, (2000): *Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods.* Discussion Paper N° 58, Trade and Macroeconomics Division, International Food Policy Research Institute, USA.
- RUIZ, F (1996): *Análisis comparativo de las distintas comunidades autónomas en base a las tablas input-output: un enfoque multivariante. Ph. D. Thesis,* Universidad de Valencia.
- SADEI (Several years): *Cuentas regionales de Asturias.*
- TILANUS, C. B. (1966): *Input-output Experiments.* Rotterdam University Press

6. APPENDIX 1

Chart A.1: AIOT 1996

	A	E	C	Q	K	B	L	G	Z
A	0,215686	0,00796	0,235167	0,001187	0	0,000337	0,006099	0,00088	3,25E-06
E	0,029371	0,488482	0,019152	0,261067	0,038009	0,046114	0,035815	0,034289	0,092135
C	0,197599	0,004217	0,313735	0,028329	0,065579	0,12442	0,090174	0,032444	0,005279
Q	0,0055	0,010376	0,008777	0,131171	0,161114	0,144824	0,002902	0,00505	0,004613
K	0,014414	0,019114	0,007793	0,038537	0,314767	0,075668	0,018281	0,019758	0,025882
B	0,003566	0,026145	0,001208	0,013786	0,004187	0,001026	0,026503	0,012663	0,012741
L	0,009927	0,006232	0,012189	0,013554	0,037887	0,047505	0,051529	0,020282	0,020095
G	0,056219	0,012469	0,014298	0,02278	0,022934	0,031461	0,109849	0,262869	0,016138
Z	0,001204	0,003878	0,005354	0,009198	0,006765	0,01371	0,003548	0,002755	0,006767

Chart A.2: AIOT 1997

	A	E	C	Q	K	B	L	G	Z
A	0,21365	0,00829	0,22625	0,00119	0	0,00032	0,00569	0,00079	3,29E-06
E	0,02715	0,47492	0,01719	0,24341	0,0328	0,04121	0,0312	0,02868	0,0871
C	0,20468	0,0046	0,31564	0,0296	0,06341	0,12461	0,08804	0,03041	0,00559
Q	0,00576	0,01143	0,00892	0,13851	0,15744	0,14657	0,00286	0,00478	0,00494
K	0,01569	0,02188	0,00824	0,04231	0,31979	0,07962	0,01875	0,01946	0,02881
B	0,00376	0,02898	0,00124	0,01465	0,00412	0,00105	0,02632	0,01208	0,01373
L	0,01068	0,00705	0,01274	0,01471	0,03805	0,04942	0,05225	0,01975	0,02211
G	0,06421	0,01498	0,01586	0,02624	0,02445	0,03474	0,11824	0,27168	0,01885
Z	0,00131	0,00444	0,00566	0,0101	0,00688	0,01444	0,00364	0,00272	0,00754

Chart A.3: AIOT 1998

	A	E	C	Q	K	B	L	G	Z
A	0,21449	0,00863	0,22533	0,00119	0	0,00032	0,00573	0,0008	3,43E-06
E	0,02632	0,4769	0,01654	0,23478	0,03189	0,03966	0,03031	0,02802	0,08772
C	0,2068	0,00481	0,31636	0,02976	0,06427	0,12497	0,08914	0,03096	0,00587
Q	0,00578	0,01189	0,00889	0,13841	0,15863	0,14613	0,00288	0,00484	0,00515
K	0,01559	0,02253	0,00812	0,04184	0,31885	0,07855	0,01868	0,01949	0,02975
B	0,00369	0,02948	0,00121	0,01432	0,00406	0,00102	0,02591	0,01195	0,01401
L	0,0106	0,00725	0,01255	0,01453	0,0379	0,0487	0,05199	0,01976	0,02281
G	0,0636	0,01537	0,01559	0,02587	0,0243	0,03416	0,11738	0,27117	0,0194
Z	0,00129	0,00451	0,00551	0,00985	0,00676	0,01405	0,00358	0,00268	0,00768

Chart A.4: AIOT 1999

	A	E	C	Q	K	B	L	G	Z
A	0,21459	0,00872	0,22311	0,00117	0	0,00032	0,00567	0,0008	3,51E-06
E	0,02611	0,47806	0,01624	0,22928	0,03132	0,03862	0,02973	0,02773	0,08882
C	0,21007	0,00494	0,31806	0,02975	0,06462	0,1246	0,0895	0,03137	0,00609
Q	0,00588	0,01222	0,00895	0,13859	0,15974	0,14593	0,0029	0,00491	0,00535
K	0,01569	0,0229	0,00809	0,04144	0,31759	0,07759	0,01858	0,01956	0,03055
B	0,00368	0,02967	0,00119	0,01404	0,004	0,001	0,02551	0,01187	0,01424
L	0,01068	0,00738	0,0125	0,0144	0,03778	0,04813	0,05175	0,01984	0,02343
G	0,06372	0,01556	0,01545	0,02551	0,0241	0,03359	0,11625	0,27097	0,01983
Z	0,00128	0,00453	0,00541	0,00963	0,00665	0,01369	0,00351	0,00266	0,00778

Chart A.5: AIOT 2000

	A	E	C	Q	K	B	L	G	Z
A	0,21454	0,00867	0,22422	0,00118	0	0,00032	0,0057	0,0008	3,5E-06
E	0,02622	0,47748	0,01639	0,23203	0,03161	0,03914	0,03002	0,02787	0,08827
C	0,20844	0,00487	0,31721	0,02975	0,06445	0,12479	0,08932	0,03117	0,00598
Q	0,00583	0,01205	0,00892	0,1385	0,15918	0,14603	0,00289	0,00488	0,00525
K	0,01564	0,02271	0,00811	0,04164	0,31822	0,07807	0,01863	0,01952	0,03015
B	0,00368	0,02957	0,0012	0,01418	0,00403	0,00101	0,02571	0,01191	0,01413
L	0,01064	0,00732	0,01252	0,01447	0,03784	0,04842	0,05187	0,0198	0,02312
G	0,06366	0,01546	0,01552	0,02569	0,0242	0,03388	0,11681	0,27107	0,01962
Z	0,00128	0,00452	0,00546	0,00974	0,00671	0,01387	0,00354	0,00267	0,00773

NOTES

¹ Specifically for Asturias, there are input-output tables available every five years and a lag in their publication of approximately three years.

² Ramos, C., M.J. Presno & R. Pérez (1999): Estimación de tablas input-output: un enfoque espacial-temporal. *XIII Reunión ASEPELT España*, Burgos.

³ An exhaustive analysis of different estimation methods can be seen in R. Álvarez (2001).

⁴ The exposition line appearing in F. Ruiz (1996) will be followed, but slightly modified to be adapted to the context of social accounting matrices.

⁵ As sums of rows and columns are the same, \mathbf{x} and \mathbf{y} are equal.

⁶ Proofs of different properties regarding this method are made from the concept of a connected matrix. A is a connected matrix if there is no row and column partition in I, I', J y J' , that verifies $a_{ij}=0 \forall i \in I$ y $j \in J'$ y $\forall i \in I' \forall j \in J$.

⁷ This problem does not have a closed-form solution, which is why one must use numerical optimization.

⁸ The RAS method keeps zeros in the estimated entries.

⁹ The indices employed refer to the national economy because the appropriate regional indices are not available.