42nd Congress of the European Science Association

August 27-31, 2002

Dortmund, Germany

SPATIAL PROJECTION OF INPUT-OUTPUT TABLES FOR SMALL AREAS

Rubén Álvarez Herrero

Ana Salomé García

Carmen Ramos Carvajal*

ABSTRACT

The purpose of this paper is double: firstly, we aim to estimate input-output tables of the Asturian districts for 1995 and, secondly, analyse the industrial structure of the Asturian region from a structural study of their districts.

To achieve the former, an undirected method of estimation of input-output coefficients based on information theory, known on the literature of many fields as "cross entropy", will be released. Whereas to get the latter, tools related to graph theory will be submitted.

* Departament of Applied Economics
University of Oviedo,
E-Mail: rherrero@correo.uniovi.es, asgarcia@correo.uniovi.es, cramos@correo.uniovi.es

1. INTRODUCTION AND MOTIVATION

It is not a task without complexity to elaborate an Input-Output Table (IOT), in that both a great deal of statistic information and a method for relating such information, is required. This entire process, which is generally defined as "direct" method, yields reliable results though, as a counterpart, it implies to make use of a great amount of economic resources and time. There is an important discrepancy between the published data –the year of releasing- and the benchmark table -the year of measuring- due to these issues. In order to overcome these time-lags we believe that the available tables should be updated using any approximation method which enables us to decrease the required information and, at the same time, give consistent results.

Given that IOT is very important for regional analysis, bureau and statistic offices all over the world have devoted many efforts to elaborate tables, not only on a national but also on a regional level. Nonetheless, many smaller geographic areas still do not have a reference table, which allows us to analyse their industrial structure. According to this, we believe it will be necessary overcome this lack or scarcity of tables in an efficient way, that is, to obtain tables for small areas (in our case, Asturian Districts) using an approximation method, which gets to reduce the manufacture cost and, at the same time, preserve, as much as possible, the quality on the estimates with respect to the direct method. The major pioneer works on this subject are the following: Schaffer & Chu 1969; Round 1978; Sasaki & Shibata 1984.

Consequently, the IOT approximation through indirect methods can be planned with a double perspective: either as an update problem of an existent matrix, or as a problem of spatial projection, estimating regional (small areas) tables from other economy table.

In our case, we propose to estimate IOT of the eight Asturian Districts for 1995, taking as starting point the table of the Asturian Region in the same period, in that the last IOT of Asturias is referred to this year.

There are several indirect techniques that allow us to cover the problem of spatial projection, and so the first step that must be given consist on select one of the possible methods to achieve our target. In pursuit of it, we have decided to employ the cross entropy procedure, which is closely related to information theory, on the base of two aspects:

a) From a theoretical perspective, it turns out to be one of the most consistent procedures to estimate tables from incomplete information (Kapur & Kesavan 1992). Nonetheless, we are aware that it has certain drawbacks, which have been taken into account thoroughly in prior works (McDougall 1999), and we consider that more efforts should be still devoted to compare the theoretical properties of cross entropy in relation to other methods like RAS procedure.

b) The obtained results on several empiric analyses tend to verify the cross entropy procedure as the one that gives the closest results to the direct method (Golan, et alia 1994; Robinson, Cattaneo, et alia 2000).

On the other hand, the structural analysis is an essential requisite for prediction tasks and involves an important aid to take political economic decisions. An interesting side of this study is the comparison between different economic structures, which can be done in space or in time. In the first case, we know the differences and/or similitudes on two economies and in the second, we can introduce the temporal evolution. In order to analyse such structure we propose to apply a methodology based on graph theory, which allow us to group Asturian Districts in areas with similar characteristics.

2. THE CROSS ENTROPY PROCEDURE AS INDIRECT METHOD OF PROJECTION FOR SMALL AREAS

In this section we present, on the one hand, the formal treatment of entropy procedures based on information theory, as well as the treatment of statistics information. On the other hand, the results of the applied method: cross entropy.

2.1 THE GENERAL PROBLEM OF ESTIMATION

The general problem, which is widely known in the literature of many fields as "constrained matrix adjustment" (Deming & Stephan 1940; Bacharach 1970; Macgill 1970), is focused on, given prior information y_{ij}, $i = 1,...,n$, $j = 1,...n$, finding a new estimations z_{ij} which must be as close as possible to **Y** and, at the same time, have to be consistent with the given total sum by columns $z_{.j} = \mathbf{v}$ and rows $z_{i.} = \mathbf{u}$. We denote **Y** as the prior matrix of intermediate flows and z_{ij} as any pair (ij) of the unknown matrix **Z**.

Consequently, we seek to recover, from incomplete data, a new matrix **Z** that accomplishes with the prior **Y** and the given **u** and **v** information.

Within input-ouput framework and in our specific case, **Z** is the intermediate consumptions table of any Asturian districts, which we want to estimate, **Y** is the counterpart to Asturian Region, which we consider as a starting point, whereas **u** (**v**) is the given sum vector by rows (columns) of each one of the **Z** Asturian Districts tables, which we impose as problem constraints. In addition, it is important that we are aware that rows (columns) of table are the sell (purchase) sectors.

As a general rule this problem is indeterminate, in that generally we have more variables (n^2) than given parameters (2n). We can pose this one as an optimisation problem imposing linear constraints to make this problem determinate. It is only one of the possible options, since it could also be solved like an iterative process.

In this paper, we have decided, following authors like (Golan, et. alia 1994; Sherman, et alia 2000), present our problem as optimisation one, making use of the entropy principles, which are taken from information theory. The major of these principles, which guide all the exposure, will be the following: "we should only use the information given to us and should strictly avoid using any information not given to us". So, our basic philosophy is use only and all the information available to us.

2.2 THE CROSS ENTROPY PROCEDURE

From the entropy principles point of view, it can be considered that an IOT of intermediate flows (consumptions) has certain inherent entropy. Accordingly, we must begin defining what is meant by "entropy" and we have to continue introducing any measures related to it. The entropy is basically a measure of probabilistic uncertainty.

In order to formalize this issue let us consider a discrete variable **X** with probability distribution $\mathbf{p} = [p_1, ..., p_n]$, which satisfies the following:

$$\sum_i p_i = 1$$

$$p_i \geq 0$$

Where i=1,...,n .

Given **X,** its entropy (Shanon´s measure) is defined as:

$$H(\mathbf{X}) = -\sum p_i \log p_i$$

Where i=1,..,n.

Let $\mathbf{p} = [p_1,...,p_n]$ and $\mathbf{q} = [q_1,...,q_n]$ be two probability distributions. Then, the relative entropy (Kullback-Leibler´s distance) between two probability distributions is given by:

$$D(\mathbf{p}|\,|\mathbf{q}) = \sum_i p_i \log \frac{p_i}{q_i}$$

It is a measure of the divergence that assumes that the given distribution is **q**, whereas the unknown distribution is **p**. An important property of relative entropy is $D(\mathbf{p}|\,|\mathbf{q}) \geq 0$, with zero if and only if **p**=**q.** Another interesting property is $D(\mathbf{p}|\,|\mathbf{q})$ is a convex function.

In pursuit of being able to determine "constrainted matrix adjustment" problems within our specific context, according to other authors (Golan, Judge, et alia 1994), we release several procedures based on the entropy principles. From *Laplace's principle of insufficient reason* it is possible to point out that if there is no information of **p** distribution we might assume that each element unknown of **p** is equally likely to be selected, i.e., we should choose the uniform probability distribution as our estimation since we seek to maximize the uncertainty of the information unknown to us. Otherwise, if information related to **p** distribution is known, then the problem can be reformulated in an optimisation context that involves maximizing a non-linear function subject to certain constrains. Thus the Shanon´s uncertainty measure defined out of **Z** and linear constrains related to given information (**u** y **v**), are introduced:

$$H(\mathbf{Z}) = -\sum_{i,j} b_{ij} \log b_{ij}$$

Where b_{ij} coefficients denote the $\frac{z_{ij}}{\sum_i z_{ij}}$ shares, being z_{ij} the intermediate consumptions (flows) of the matrix **Z** unknown, which is any table of Asturian districts. In addition, if it is taken into account the **Z** probabilistic character and the given $z_{.j} = \mathbf{v}$ and $z_{i.} = \mathbf{u}$ of any **Z** table, then we introduce the following constraints:

$$\sum_i b_{ij} = 1$$

$$b_{ij} \geq 0$$

$$\sum_j b_{ij} z_{.j} = z_{i.}$$

Using the maximum entropy principle: "out of all probability distributions satisfying given constraints, choose the distribution that maximizes the entropy of the information unknown", the maximum entropy procedure (MEP) will consist on maximizing H(**Z**) subject to the latter constraints set, which ensures the estimations reliability.

As some authors have demonstrated (Golan, et alia 1994), the MEP has a formal solution. To achieve it they form the Lagrangian function and solve the equations resultant from the first order conditions.

We note that the objective function H(**Z**) is not single entropy but a sum of entropies. Thus, the H function can be translated into H_j components consistent with the b_{ij} coefficients sum by columns, which must be one, and provides an uncertainty measure for each column.

In the MEP only information related to the tables of Asturian districts, was used. In our case, there is an IOT of the Asturian Region, **Y**, which provides potential information that could be used in getting the results of **Z** unknown table, but only if we consider that the additional information **Y** is relevant, that is, the structure of the table resulting from estimation procedure is similar to that of the reference table.

Accordingly, the cross entropy procedure (CEP), which takes into account both regional and districts information, is introduced. There are a few definitions of cross entropy. According to these authors (Kapur & Kesavan 1992), one of them defines the cross entropy as the same as the relative entropy $D(\mathbf{p}||\mathbf{q})$. It is a directed divergence from one distribution **p** to another **q**.

The cross entropy procedure is derived from *the minimun relative entropy principle*, which is started as: given a priori distribution **q**, out of all probability distributions satisfying the given constraints, choose the distribution that minimized the cross entropy of **p** from **q**. The constraints come from the probabilities law and the given information.

Within our estimation context, the CEP is posed as:

minimise

$$\sum_i \sum_j b_{ij} \log \frac{b_{ij}}{a_{ij}}$$

subject to the same MEP constraints.

Where the b_{ij} coefficients are the same as in MEP and the a_{ij} coefficients represent the $\frac{y_{ij}}{\sum_i y_{ij}}$ shares to **Y** matrix, which is the given table of Asturian Region.

Formally, this problem can be solved using Lagrange formulation. According to these authors (Golan, et alia 1994), the CEP solution is:

$$b_{ij} = \frac{a_{ij}}{\Omega_j(\lambda_i)} \exp[\lambda_i x_j]$$

where

$$\Omega_j(\lambda_i) = \sum_i a_{ij} \exp[-\lambda_i z_{.j}]$$

and λ_i parameters represent the Langrange multipliers.

The CEP has some interesting properties that can be overviewed in (McDougall 2000):

a) *Solution unity*: given the data no too scarce, it can provide a solution, and the solution if exists is unique.

b) *Conservation of zeros and sign*: the coefficients that are zero and positive in the **Y** prior table are also zero and positive in the **Z** estimated table.

c) *Similarity of structure*: small (large) coefficients a_{ij} are represented by small (large) b_{ij}, which correspond to the same fix (i,j).

d) *Relationship between CEP and MEP*: when **Y** is no available to us it can be seen that the minimization of Kullback-Leibler´s distance, subject to some constraints, is equivalent to maximization Shanon´s uncertainty or entropy measure subject to the same restrictions. Thus, the latter is a special case of the CEP (Kapur & Kesavan 1992).

2.3 ESTIMATION OF INPUT-OUTPUT TABLES FROM CROSS ENTROPY

In order to estimate the districts matrices have been necessary to “cross” the information coming from Asturias IOT with the obtained from The Income of the Asturian Boroughs (TIAB). The same institution, SADEI (Asturian Society of Regional and Economic Studies) makes both statistics, what ensures the homogeneity of data.

2.3.1 TREATMENT OF THE STATISTICAL INFORMATION

Nonetheless, we have found that the benchmark periods of both sources are different, in that the TIAB is properly bi-annual and is concerned to pair years, with which we would have information for 1994 and 1996, but not for 1995. To aim to give a solution to this problem we have estimated the necessary data of The Income of the Asturian Boroughs for 1995, assuming as such value the intermediate point of the last two years.

In order to be able to use the information derived from the TIAB of different years, we have proceeded to homogenize it at 9 sectors using the Hermes ranking. As we have already pointed out such information has had to be "crossed" with the information obtained from the IOT, and so we have also had to aggregate at 9 sectors the AIOT-95. Next, we release the aggregation done:

Table N°1 Homogenisation of TIAB & AIOT-95 at 9 sectors accordingly to Hermes ranking

HERMES SECTORS	**TIAB SECTORS 1994**	**TIAB SECTORS 1996**	**AIOT-95 SECTORS**
(A)	1	1	1-2-3
(E)	2-11	2-9	4-5-17-32-33
(Q)	7-8-9	4-5-6	6-7-8-18-20-21
(K)	10	7	22-23-24-25-26-27-28-29-30 (part)
(C)	3-4-5-6	3-8	9-10-11-12-13-14-15-16-19-30 (part)
(B)	12	10	34
(Z)	14	13	39-40-41-42-43
(L)	13	11-12-14	31-35-36-37-38-44-45-46-47-48-49-50 (part)-51-53 (part)-54 (part)
(G)	15-16	15-16	50 (part)-52-53(part)-54(part)-56-57-58-59-60

Where (A) is Agriculture, (B) Energy, (Q) Manufacture of intermediate products, (K) manufacture of equipment products, (C) Manufacture of consumptions products, (B) Construction, (Z) Transport and communications, (L) Commercial services, (G) Non commercial services.

2.3.2 DISTRIBUTION BY DISTRICS OF THE REGIONAL AGGRETATES

Having homogenised the information of employed statistic sources, the following stage in our work will be the achievement of the district aggregates to proceed to the estimation. The Asturian Region is divided in 8 districts: Eo-Navia, Narcea, Aviles, Oviedo, Gijón, Caudal, Nalón and Oriente.

In order to realize the estimation it is necessary to have both regional and districts information, in particular, and given that our aim will be the construction of tables by districts for 1995, we will have to know the region and districts values, what is possible since they are published, and we will have to access to the "real" values of the aggregates of intermediate consumptions by rows and columns. The following step will be to proceed to the achievement of such vector for each district.

a) Intermediate consumptions by columns can be obtained from the difference between Value produced at market prices (VPmp) and Gross value added at producers cost (GVApc). To realize the distribution by districts, we have adopted the GVApc as *proxy* variable that is released by sectors in TIAB. The proportion of this variable for each district in 1994 and 1996 has been calculated. These values have been used to carry out the delivery by districts of the intermediate consumptions aggregates.

b) Estimation of intermediate consumptions by rows. These values can be determined as difference between the total outputs and final demand.

$$\mathbf{u}=\mathbf{q}\text{-}\mathbf{d}$$

Where **u, q, d** represent the intermediate consumptions sum by rows, the total output and the final demand, respectively.

From the model proposed by Tilanus (1966) we can derivate the following equality:

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

Where (I-A)-1 is the Leontief inverse.

Replacing in the first equality and operating suitably is obtained:

$$\mathbf{u}=[(\mathbf{I}\text{-}\mathbf{A})^{-1}\mathbf{d}\text{-}\mathbf{d}]=[(\mathbf{I}\text{-}\mathbf{A})^{-1}\text{-}\mathbf{I}]\mathbf{d}$$

That is, we have expressed the intermediate consumptions by rows depending on the final demand, nonetheless, we do not dispose information by districts of this variable. Thus, we have had to use a *proxy* variable of its behaviour and we have decided for compensation of employees, in that we consider it relevant in the demand analysis and are available, both by districts and by activity branches. Proceeding in the same way as in the case of intermediate consumptions by columns, the aggregates for districts by rows has been estimated.

2.3.3 RESULTS OF ESTIMATION

We have applied the cross entropy procedure to the Asturian IOT of 1995; the sum vectors of intermediate consumptions by rows (**u**) and columns (**v**) are the one referred to as Asturian districts in 1995. From the theory of non-linear optimisation, it is relatively straightforward to estimate a matrix subject to restrictions using the cross entropy procedure, in that the measure of relative entropy implied in the procedure ensures a unique global solution. So, we have made use of a non-linear computational program such as GAMS to solve this problem. The estimated tables are gathered on the annex I.

3. STRUCTURAL ANALYSIS OF THE ASTURIAN DISTRICTS FROM GRAPH THEORY

Once we have estimated the matrixes, we then go on to study the productive structure of the Asturian region from them. In order to analyse the structures we propose appling a methodology based on graph theory.

The application of graph theory on structural analysis has a lot of potential since it focuses on issues like the relative positions of sectors, their orientation or the paths through which they revolve round the economic influence in the considered structure.

3.1 THE STUDY OF DIRECT RELATIONS BETWEEN SECTORS

Intuitively, a graph can be defined as a representation of the actual relations between the elements of an objects set (Pelegrin et al. 1992). More formally, a directed graph is defined by the pair (V, A) where V is a set of elements known as nodes or vertices and A is a set made up of actual arranged arcs between analysed elements.

The study of direct relations between different productive branches makes for a first approximation in the analysis of the sectorial interchange structure; with this aim, given the matrix of input-output coefficients we raise the absolute graph of influence.

Let us define an associated matrix, such as the transposed boolean matrix resulting from input-output coefficients (b_{ij}), then their elements will be:

$$z_{ij} = \begin{cases} 1 & \text{si } b_{ji} \neq 0 \\ \\ 0 & \text{si } b_{ji} = 0 \end{cases}$$

From this graph, we can analyse several aspects taking into account the degrees that have been reached by every branch. Thus, the output degree is defined by the number or lines going out of the vertex:

$$d^{+}(z_i) = \sum_j z_{ij}$$

They are the relations of direct causality produced by the sector *i*, or number of branches to which it makes purchases. So, it is about a measure of integrations by purchases.

It is understood by input degree the number of lines going out of the vertex:

$$d^{-}(z_i) = \sum_j z_{ji} = \sum_i z_{ij}$$

This can be interpreted as number of sectors that buy to i-esimo branch, becoming a measure of the sells integration of the referred to sector.

From the aggregation of these two expressions, it is derivated the all degree, the total number of lines, coming in or going out of the vertex:

$$d(z_i)= d^{+}(z_i)+d^{-}(z_i)$$

That is the total number of relations which *i-esimo* sector has.

On the basis of the previous presented measures, it is possible to establish a classification of the different branches with regard to his degree of integration, such as it is showed in the next table:

Table Nº2. Classification of sectors according to its degree of integration

Classification	**Conditions**
Little integrated	If $d^{+}(z_i)<\mu$ y $d^{-}(z_i)<\mu$
Integrated mainly by buys	If $d^{+}(z_i)>\mu$ y $d^{-}(z_i)<\mu$
Integrated mainly by sells	If $d^{+}(z_i)<\mu$ y $d^{-}(z_i)>\mu$
Much integrated	If $d^{+}(z_i)>\mu$ y $d^{-}(z_i)>\mu$

Where the average is represented by μ.

The output degrees can be interpreted as an index of direct influence and the input degrees as indicators of direct dependence, and so it is possible introduce an index of net direct influence whose expression will be:

$$\sigma_i = \frac{d^+(z_i)}{d^-(z_i)} \qquad \forall\ d^-(z_i) \neq 0$$

This expression is delimited lower by cero and upper by the number of sectors, that is, $0 \leq \sigma_i \leq n$. The bigger the direct net influence the higher the index values and vice versa.

One inconvenient of this kind of studies is the cut-off point which it be considered for making the graph, given that we can distort the analysis taking into account as equal all the actual relations independently of its numeric value. The applied criteria in the coefficients elimination are frequently subjective. One possible solution to this problem is change the cut-off point systematically in order to look into the structural evolution.

The developments, which we will be next are based on a partial graph, given that the average over averages of every graph as applied cut- point. Thus, one represents those coefficients with greater values, whereas the rest will be identified with zeros.

From the I-O Tables we can observe that the density of graphs is between 0.1975, in the case of the minor value corresponding to Oviedo, and the upper number 0.308 for Avilés. It means the rate with one sector has relations with the others in its district through transactions are larger than the mentionated cut-point. In this sense, there is not a wide range of density degree between these small areas. The level of economic relations carried out in their economies is very similar.

An analysis in detail of their direct relational measures lets us assemble several districts round four areas.

The districts known as Narcea, Caudal and Nalón form one group, which has been called Area I. It is a zone which shows us a big weight of little integrated branches: Non-commercial services (G), Construction (B) and Transport and communications (Z). It exhibits high indexes of direct net influence corresponding to the branches Manufacture of equipment products (K), Manufacture of consumption products (C) and Agriculture (A). Thus, they are sectors with a clear orientation to their purchaser side.

Table Nº3 Area I

Sector classification	Narcea	Caudal	Nalón
Little integrated	A, C, B, Z, G	B, Z, G	G
Integrated mainly by buys	K	A, K, C	A, K, C, B
Integrated mainly by sells	E, Q, L	E, Q, L	E,Q
Much integrated	---	---	Z, L

Source: Own elaboration from IOAT-95.

The Area II, constituted by Eo-Navia and Oriente, has Transport and Communications (Z) as a much integrated branch in the economy of theses Asturian places. Its purchase axis is based mainly on Energy (E), Manufacture of equipment products (K) and Construction (B).

Table Nº4 Area II

Sector classification	Eo- Navia	Oriente
Little integrated	A, G	G
Integrated mainly by buys	E, K, B, Z	E, K, B
Integrated mainly by sells	Q, C	A, Q, C, Z
Much integrated	L	L

Source: Own elaboration from IOAT-95.

The rest of the districts, Oviedo, Gijón and Avilés; are the main towns in the Asturian region. Oviedo and Gijón shape the Area III in which the branches Manufacture of consumption products (C) and Manufacture of equipment products (K) are very integrated and offer high indexes of direct net influence as well. Whereas Construction (B) and Transport and Communications (Z) are oriented chiefly to purchases.

Table Nº5 Area III

Sector classification	Oviedo	Gijón
Little integrated	G	E, G
Integrated mainly by buys	A, K, B, Z	C, B, Z
Integrated mainly by sells	E, Q, L	A, Q, L
Much integrated	C	K

Source: Own elaboration from IOAT-95.

The Area IV is formed by Avilés. It is a different district from the other two main cities of Asturias. It has hardly any characteristics in common with the others. Thus, inside group constituted by the much integrated branches are the Transport and communications (Z) and Commercial services (L). Energy (E) and Manufactures of intermediate products (Q) are sectors, which are orientating to their purchaser side.

Table Nº6 Area IV

Sector classification	Avilés
Little integrated	C, G
Integrated mainly by buys	E, Q, B
Integrated mainly by sells	A, K
Much integrated	Z, L

Source: Own elaboration from IOAT-95.

3.2 THE STUDY OF INDIRECT RELATIONS

The study of direct relations can also be enriched by considering the indirect transactions. In order to analyse these total relations it is usual to introduce the concept of path matrix and distance matrix.

The matrix of paths is defined as a matrix of zeros and ones, in such way that if an element takes the unit value, this supposes the existence of a route between sector *i* and *j*, so the branch *i* influences either directly or indirectly through others branches in sector *j*.

The minimum distance there is between two connected sectors *i* and *j* fixes the elements of the distance matrix. It informs us about the number of transformation processes that one good or service of sector *j* has to suffer until it is used through the shortest path by the sector *i* giving information about the intermediate branches.

The analysis of these itineraries of the disturbance propagations in the economy can be carried out by raising complementary measures to which we have done as the betweenes.

The degree of mediation or betweenes indicates the number of times that it is necessary to pass for every node with the purpose of connecting another two through the geodesic paths or the shortest routes.

Given g_{ik} be the number of geodesics between actor *i* and *k*, we can assume, that all geodesics are likely to be chosen equally by *i* and *k*, furthermore the probability that *i* and *k* choose a specific geodesic is $\frac{1}{g_{ik}}$. Moreover, g_{ik} (n_j) is the number of geodesics between actor *i* and *k* which contain actor *j*. The probability b_{ik} (n_j) that *i* and *k* choose a geodesic, which contain actor *j* is:

$$\frac{g_{ik}(n_j)}{g_{ik}}$$

The betweeness of a sector j will be then, the sum of the probabilities for all dyads in the network:

$$C(n_j) = \sum_i \sum_k c_{ik}(n_j) \quad j \neq i \neq k$$

and standardized betweeness:

$$C(n_j) = \frac{2C(n_j)}{n^2 - 3n + 2}$$

So the index of betweeness will measure to what extent a sector is located on the geodesics. It means the power of sectors for connect resources and in this sense, they are relevant sectors so they allow the connexion of groups that in any other way they would be kept isolated.

Thus, throughout the several districts there are some common traits in their degree of mediation. From the Table Nº8, it follows that Non-commercial services (G) is the key branch to access to other branches, contrary to Commercial services (L). On the other hand, there are distinct regional features too, in the following central branches.

Table Nº7 Betweeness Centrality

	Area I			**Area II**		**Area III**		**Area IV**
Branches	**Narcea**	**Caudal**	**Nalón**	**Eo-Navia**	**Oriente**	**Gijón**	**Oviedo**	**Avilés**
A	0.027	0	0	0	0	0.036	0	0.25
E	0.292	0.19	0.173	0.009	0	0	0	0
Q	0.077	0.054	0.036	0.125	0.089	0	0.018	0
K	0.021	0.021	0.021	0.009	0.036	0.208	0.107	0.232
C	0.042	0.027	0.045	0.384	0	0.048	0	0
B	0.021	0.021	0.021	0.009	0.036	0.054	0.107	0.036
Z	0	0	0	0	0.25	0.381	0.25	0.125
L	0	0	0	0	0	0	0	0
G	0.449	0.509	0.455	0.321	0.518	0.452	0.839	0.464

Source: Own elaboration from IOAT-95.

In the mining Area I, Energy (e) attaches great importance. The Manufacture of intermediate products (Q) is an important mediator branch in Area II. Whereas Area III puts its emphasis on Transport and communications (Z). In contrast to the other main Asturian cities, Avilés has Agriculture (A) and Manufacture of equipment products (K) as great mediators.

A representation of these betweenness measures for each district can help to identify the central nodes. Take the following example for Oviedo produced by software Netminer:

Figure Nº1

Source: Own elaboration.

It shows us clearly the degree of mediation of the different sectors in this district.

Thus, and to sum up, the applied concepts based on graph theory have allowed us to determine four areas or groups with similar commercial behaviour and shape their features regarding orientation and centrality.

4. CONCLUSIONS

The major conclusions that are drawn from the paper are the following:

- Using regional and districts information the cross entropy procedure has enabled us to approximate tables for the districts of the Asturian Region in 1995.
- Given that we do not have all the information required for this technique, a slice of it have had to be estimated exogenously by a procedure of proportional delivery between districts of the regional aggregates, which are available to us from official statistics.
- Furthermore, from estimated tables we have been able to realise a structural analysis of these Asturian Districts. For this purpose, we introduce the graph theory as an appropriate research method for the study of intersectorial relations in each district.
- The general result of this structural analysis has been that we have assembled the districts round four well-known areas with similar characteristics.
- Area I is formed by the districts known as Narcea, Caudal and Nalón. They are all mining districts whose most important sectors are Non commercial services and Energy.
- Area II, constituted by Eo-Navia and Oriente are periphery zones in the North of Asturias its Non-commercial services and Manufactures of intermediate products are important mediator sectors in this area.
- The remaining districts, Oviedo, Gijón and Avilés are the main districts of this Region. Oviedo and Gijón form the Area III which is clearly different to Area IV formed by Avilés, a district with a strong industrial past. Furthermore, the two first cities focus on Transport and communications, while Avilés focuses on Agriculture.

5. BIBLIOGRAPHIC REFERENCES

ANTILLE, G., FONTELA, E. Y GUILLET, S. (2000): Changes in technical coefficients: "The experience with Swiss I/O Tables", Thirteen International Conference on Input-Output Techniques, Macerata, Italia.

BACHARACH, M. (1970): *Biproportional Matrices and Input-Output Change*. Cambridge.

DEMING, W.E. & STEPHAN F.F. (1940): On a least squares adjustment of a sampled frequency table when the expected marginal totals are known, *Ann. Math. Statist.* 11(1), 427-444.

GARCÍA, A.S., RAMOS, C. (2001): " Análisis de la realidad económica asturiana: un enfoque desde la teoría de grafos", XXVII Reunión de Estudios Regionales, Madrid.

GOLAN, A., G. JUDGE & S. ROBINSON (1994): Recovering information from incomplete or partial multisectoral economic data. *The Review of Economics and Statistics*. Nº 76, 541-549.

GUERRERO CASAS, F.Mª , ORDAZ S., J.A. (2000): "Aplicación de la teoría de grafos al análisis input-output: Andalucía 1995", *VIII Jornadas ASEPUMA*, Sevilla.

KAPUR, J.N & KESAVAN, H.K. (1992): *Entropy Optimization Principles with Applications* (Academic Press).

MACGILL, S.M. (1977): Theoretical Properties of Biproportional Matrix Adjustements. *Environment and Planning*, A 9, Nº 6, 687-701.

MCDOUGALL, R. A. (1999): *Entropy Theory and RAS are friends.* Extracted from the following address: http://www.sjfi/publ/workshops/papers/McDougall.pdf .

MOLINA, J.L. (2000): *El análisis de redes sociales. Una introducción*, Edicions Bellaterra, Barcelona.

MOLINA, J.L., MUÑOZ, J., LOSEGO, P. (2000): "Red y realidad: aproximación análisis de Redes científicas", *VII Congreso Nacional de Psicología* Social, Oviedo.

MORILLAS R., A. (1983): *La teoría de grafos en el análisis Input-Output. La estructura productiva andaluza*, Editorial Universidad de Málaga.

ROBINSON, S., A. CATTANEO & EL-SAID M. (2000): *Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods*. Discussion paper Nº 58, Trade and Macroeconomics Division, International Food Policy Research Institute, USA.

SADEI (1998): *Cuentas regionales de Asturias, 1995*. Servicio de Publicaciones del Principado de Asturias.

SADEI: *La Renta de los Municipios Asturianos* Servicio de Publicaciones; diversos años Caja de Asturias.

TILANUS, C. B. (1966): *Input-output Experiments.* Rotterdam University Press.

WASSERMAN, S., Y FAUST, K. (1994): *Social Network Analysis. Methods and Applications*, Structural Analysis in the Social Sciences, Cambridge University Press, United States of America.

ANNEX I : INPUT-OUTPUT TABLES FOR ASTURIAN DISTRICTS (1995)

DISTRICT 1. EO-NAVIA

	A	E	Q	K	C	B	Z	L	G
A	0.628	0.071	0.04	0	0.509	6.24E-04	2.22E-05	0.069	0.006
E	5.13E-05	0.325	0.00E+00	0.025	3.31E-04	0.003	0.124	1.50E-04	0.045
Q	0.017	0.24	0.571	0.524	0.027	0.56	0.005	0.029	0.118
K	5.00E-05	3.23E-04	0.00E+00	0.185	0.00E+00	6.68E-05	0.004	0.00E+00	0.003
C	0.337	0.027	0.385	0.071	0.41	0.193	0.142	0.628	0.127
B	3.61E-04	0.075	6.51E-06	0.004	2.85E-04	4.37E-04	0.044	0.008	0.043
Z	0.004	0.124	0.003	0.038	0.022	0.075	0.214	0.036	0.086
L	0.013	0.13	4.09E-04	0.151	0.029	0.166	0.46	0.22	0.456
G	0.001	0.008	1.94E-05	0.004	7.27E-04	0.002	0.007	0.009	0.115

DISTRICT 2. NARCEA

	A	E	Q	K	C	B	Z	L	G
A	0.48	1.37E-05	8.51E-04	0	0.598	4.04E-04	1.19E-05	0.013	0.004
E	0.389	0.999	0.62	0.028	0.064	0.095	0.471	0.531	0.24
Q	0.034	6.00E-06	0.337	0.505	0.061	0.563	0.003	0.016	0.095
K	1.61E-05	0.00E+00	1.55E-04	0.225	0.012	0.022	0.025	8.83E-06	0.013
C	1.00E-05	0.00E+00	1.68E-06	0.052	0.054	0.002	0.006	0.00E+00	0.003
B	0.003	1.30E-06	0.005	0.004	0.003	9.64E-04	0.041	0.022	0.046
Z	7.41E-05	0.00E+00	9.40E-04	0.033	0.035	0.02	0.056	1.39E-04	0.019
L	0.083	0.00E+00	0.033	0.149	0.168	0.293	0.392	0.399	0.459
G	0.01	0.00E+00	0.002	0.004	0.004	0.004	0.006	0.019	0.121

DISTRICT 3. AVILÉS

	A	E	Q	K	C	B	Z	L	G
A	0.568	0.107	0.013	0	0.755	0.001	1.94E-05	0.206	0.006
E	0.037	0.339	0.089	0.002	0.005	0.011	0.175	0.005	0.071
Q	0.023	0.046	0.313	0.094	0.009	0.186	0.002	0.003	0.058
K	0.05	0.164	0.198	0.657	0.026	0.366	0.149	0.13	0.131
C	0.185	0.005	0.011	0.021	0.096	0.054	0.055	0.033	0.057
B	0.007	0.092	0.031	0.005	0.001	0.001	0.047	0.068	0.055
Z	0.018	0.16	0.251	0.127	0.057	0.173	0.208	0.177	0.1
L	0.1	0.081	0.088	0.092	0.049	0.205	0.358	0.358	0.41
G	0.011	0.006	0.006	0.003	0.001	0.003	0.006	0.02	0.112

DISTRICT 4. OVIEDO

	A	E	Q	K	C	B	Z	L	G
A	0.385	0.004	1.68E-05	0	0.12	6.27E-05	4.35E-06	3.66E-05	9.53E-04
E	0.066	0.692	0.101	0.024	0.042	0.028	0.253	0.014	0.092

Q	0.041	0.09	0.736	0.516	0.1	0.48	0.004	0.016	0.088
K	0.052	0.033	0.007	0.212	0.026	0.099	0.08	0.002	0.055
C	0.261	0.004	0.003	0.057	0.342	0.054	0.053	0.011	0.045
B	0.01	0.066	0.023	0.004	0.005	0.001	0.053	0.063	0.056
Z	0.016	0.015	0.001	0.026	0.024	0.022	0.073	3.16E-04	0.025
L	0.154	0.091	0.123	0.157	0.333	0.312	0.478	0.866	0.514
G	0.016	0.005	0.005	0.004	0.008	0.004	0.007	0.026	0.123

DISTRICT 5. GIJÓN

	A	E	Q	K	C	B	Z	L	G
A	0.446	0.012	4.95E-04	0	0.212	1.88E-04	9.72E+00	0.004	0.003
E	0.058	0.685	0.097	0.017	0.016	0.024	0.261	0.023	0.101
Q	0.034	0.079	0.653	0.482	0.069	0.407	0.004	0.017	0.084
K	0.054	0.078	0.125	0.295	0.086	0.246	0.153	0.093	0.115
C	0.252	0.006	0.012	0.058	0.414	0.081	0.075	0.096	0.069
B	0.008	0.063	0.027	0.004	0.005	0.001	0.055	0.09	0.056
Z	0.008	3.33E-04	0.00E+00	5.43E-05	0.00E+00	9.47E-05	0.004	0.00E+00	0.002
L	0.126	0.072	0.081	0.14	0.193	0.237	0.441	0.644	0.45
G	0.014	0.005	0.005	0.004	0.006	0.004	0.007	0.033	0.12

DISTRICT 6. CAUDAL

	A	E	Q	K	C	B	Z	L	G
A	0.392	0.00E+00	1.60E-04	0	0.072	1.43E-04	4.58E-06	1.40E-04	0.002
E	0.092	0.999	0.713	0.036	0.391	0.108	0.457	0.652	0.218
Q	0.036	1.50E-05	0.208	0.508	0.037	0.393	0.002	0.002	0.062
K	0.035	0.00E+00	0.002	0.187	0.001	0.05	0.03	3.72E-05	0.027
C	0.258	0.00E+00	0.004	0.066	0.243	0.08	0.046	0.012	0.051
B	0.01	2.27E-04	0.012	0.004	0.005	0.001	0.042	0.036	0.052
Z	0.017	0.00E+00	0.012	0.033	0.019	0.051	0.079	0.001	0.039
L	0.145	1.38E-04	0.045	0.162	0.223	0.311	0.338	0.271	0.423
G	0.017	4.02E-05	0.004	0.005	0.009	0.005	0.006	0.024	0.126

DISTRICT 7. NALÓN

	A	E	Q	K	C	B	Z	L	G
A	0.387	1.00E-05	6.16E-05	0	0.068	8.52E-05	1.00E-05	1.04E-04	0.001
E	0.077	0.987	0.601	0.043	0.143	0.069	0.362	0.398	0.162
Q	0.036	2.01E-04	0.225	0.46	0.046	0.339	0.002	0.005	0.067
K	0.055	1.09E-04	0.024	0.229	0.033	0.163	0.081	0.015	0.082
C	0.263	1.00E-05	0.005	0.058	0.329	0.071	0.046	0.033	0.057
B	0.009	7.35E-04	0.01	0.004	0.003	9.56E-04	0.036	0.042	0.047
Z	0.022	0.012	0.094	0.051	0.198	0.122	0.168	0.14	0.085
L	0.136	5.42E-04	0.039	0.151	0.175	0.231	0.3	0.35	0.395
G	0.015	4.53E-05	0.002	0.004	0.005	0.003	0.005	0.017	0.104

DISTRICT 8. ORIENTE

	A	E	Q	K	C	B	Z	L	G
A	0.832	0.083	0.059	0	0.476	0.001	2.74E-05	0.27	0.008
E	7.45E-04	0.323	1.41E-05	0.021	0.014	0.003	0.099	8.68E-05	0.04

Q	0.012	0.168	0.305	0.525	0.036	0.472	0.004	0.015	0.103
K	1.60E-04	0.011	3.01E-06	0.185	0.012	0.004	0.023	6.19E-06	0.019
C	0.107	0.015	0.011	0.07	0.294	0.119	0.098	0.147	0.096
B	0.001	0.071	6.08E-04	0.004	0.002	4.76E-04	0.04	0.008	0.043
Z	0.018	0.209	0.615	0.042	0.067	0.226	0.319	0.367	0.141
L	0.025	0.114	0.008	0.149	0.096	0.172	0.409	0.187	0.442
G	0.003	0.006	3.42E-04	0.004	0.002	0.002	0.006	0.006	0.109

ANNEX II: GRAPHS

Table NºI Index of net direct influence σ_i

Branches	**Eo-Navia**	**Narcea**	**Avilés**	**Oviedo**	**Gijón**	**Caudal**	**Nalón**	**Oriente**
A	1	1	0.667	3	0.667	3	3	0.333
E	1.5	0.167	1.5	0.5	0.5	0.167	0.167	3
Q	0.5	0.667	1.5	0.333	0.333	0.667	0.667	0.5
K	3	3	0.286	3	1	3	1.5	3
C	0.2857	---	1	1	1.5	1.5	2	1
B	---	---	---	---	---	---	---	---
Z	3	---	0.667	---	---	---	1	0.333
L	0.5	0.333	1	0.143	0.167	0.286	0.428	0.8
G	---	---	---	---	---	---	---	---

Source: Own elaboration from IOAT-95.