Complexity as Interrelatedness: an Input-Output Approach

João F. Amaral, João Dias and João C. Lopes

UECE-ISEG
Technical University of Lisbon
Rua do Quelhas, nº 6, 1200-781 Lisboa - Portugal

e-mail: jcflopes@iseg.utl.pt

Paper prepared for the EcoMod/IIOA Conference:

Input-Output and General Equilibrium: Data, Modelling and Policy Analysis

September 2-4, 2004
Brussels, Belgium

Abstract:

In this paper we make a first attempt to link two subjects with a potentially useful, but as yet not conveniently explored, connection: the study of complexity and the (Leontief) input-output analysis. In this context, we consider economic complexity as interrelatedness between the different parts or sectors of an economy, as represented by an input-output system and one interesting question emerges, namely: should we expect to find a natural move to higher complexity as the economy grows and develops? And a related one: is a larger economy necessarily more complex than a smaller one? In a first attempt to answer these questions we propose a new measure of complexity as interrelatedness that combines a network effect and a dependency effect. In the empirical part of the paper we apply this measure of complexity to the inter-industry tables of several OECD countries, and discuss some interesting findings.

JEL Classification Codes: C67, D57

Keywords: economic complexity, input-output analysis

1. Introduction

The main purpose of this paper is to make an explicit link between two subjects with a potentially useful, and yet in the relevant literature largely implicit, connection: the study of complexity and the (Leontief) input-output analysis.

We don't discuss here, but are aware of identification problems (Durlauf, 2003) and other difficulties related to the economic applications of the complexity notion (Rosser, 1999).

Broadening the scope of analysis, we are also familiar with the difficulty, not to say impossibility, in obtaining a comprehensive and universal definition of this notion, so that a researcher can be confronted with a choice from this large (but not at all exhaustive) menu: computational, statistical, structural, functional, hierarchical, sequence, Kolmogorov, informational, effective, physical complexity [1].

However, it appears to emerge as one of the most prominent characteristics of this concept, common in several systems (physical, biological, political, social or economical), the interaction between different components (or agents) of a whole. As Brian Arthur (1999, p. 107) puts it, "Common to all studies on complexity are systems with multiple elements adapting or reacting to the pattern these elements create".

From this perspective, it appears almost obvious the gain of studying economic complexity within an input-output framework. In fact, not yet sufficiently explored and certainly well before the phase of marginal diminishing returns, this area of research has recently been enriched with interesting contributions. Some examples: Sonis and Hewings (1998) define economic complexity as an emerging property of the process of network complication that can be studied by means of a structural path analysis; Dridi and Hewings(2002) make a decomposition of economic complexity into finite stages, using a data analysis technique known as dual scaling; Aroche-

[1] For a quick survey of some of these definitions, most of them proposed by physicists and biologists, see Adami (2002).

Reyes (2003) equates a growing economic complexity to an increasing number of important connections between industries, capturing these by means of increasingly complicated graphs of the economy, using the so-called Qualitative Input-Output Analysis (Campbell, 1975; Schnable, 1994).

Most of these studies are mainly concerned with an analysis of structural change[2], the quantification of economic complexity being in a certain sense a secondary or by-product result.

The aim of this paper is to present a method for quantifying complexity, considering it explicitly as (the level of) interrelatedness between the parts or sectors of an economy, represented here by an input-output system.

In section 2, we propose an index of complexity that combines two dimensions or effects in a linear system, a network effect and a dependency effect, distinguishing two possible ways of perception of complexity: one the *inside view*, that is the point of view of those that are immersed in the system and have only partial information about it; and the other the *outside view* of those that from the outside have all the relevant information about the system.

This distinction is important to clarify the frequent confusion made in the literature about complexity of a system between these two different points of view: complexity is one thing to someone outside the system having all the relevant information about it; complexity is (or at least should be) another thing to someone immersed in the system and having only a limited information about it.

From the first point of view we can say, for example, that a linear system is less complex than a non linear system with chaotic behaviour. From the second point of view things are not so clear. A system can have low complexity for those that look at it from the outside and great complexity for those inside that deal with problems such as those that living beings (or firms) face in their environment (or markets).

[2] A pioneering and strong influential article for much of these studies and our own is Yan and Ames (1965).

This case, which is a source of much confusion in the literature, will be exemplified later in this paper.

For those inside the system, complexity arises mainly from the inter-relations between the parts of the system, in the sense that the behaviour of a part can be strongly affected by the behaviour of many others. If this is the case, a rational agent inside the system who has information only about the part where he is located or about a limited neighbourhood of it lacks information about important factors affecting the behaviour of that part of the system. To behave rationally under such conditions becomes a very complex problem for this agent.

So, the interrelatedness between the parts of a system is certainly an essential feature of its complexity, from the *insider* point of view. As we are particularly interested in this point of view, the indicator we propose for measuring complexity is essentially an interrelatedness indicator.

After the construction of our relevant index of 'inter-industry complexity', we apply it in section 3 to a set of countries using the OECD input-output database. We are particularly interested in measuring the evolution of 'quantitative complexity' as the countries develop and grow.

Finally, we provide in section 4 some concluding remarks.

2. A measure of (inter-industry) economic complexity

In order to construct an index of complexity as interrelatedness, we must consider two effects:

a) a "network" effect, that gives us the extent of direct and indirect connections of each part of the system with the other parts; more connections correspond to more complexity;

b) a "dependency" effect, that is, how much of the behaviour of each part of the system is determined by internal connections between the elements of that part – which means more autonomy and less dependency – and how much that behaviour is determined by external relations that is, relations with other parts of the system – which means less autonomy and more dependency.

Both effects will be measured with an index of their own. We will begin constructing an index for the dependency effect, and then an index for the network effect.

2.1 The degree of dependency

Let us first consider a system represented by a square matrix A, of order N and with all values non negative.

A part of the system of order *m* *(m = 1, ..., N-1)*, is a square block A* of order *m* which has its main diagonal formed by *m* elements of the main diagonal of A.

Let A* be a part of the system. For example:

$$A^* = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

We can consider A* as a sub-system of the system A. This sub-system is the more autonomous (or, equivalently the less dependent) the greater the values of its elements (a_{11}, a_{12}, a_{21}, a_{22}) are relative to the elements (a_{1j}, a_{2j}, a_{j1} a_{j2}), for all j>2.

In order to measure the greater or lesser autonomy of the sub-system A*, we define the autonomy degree of A* as:

$$(2.1) \qquad G_a(A^*) = \frac{\|A^*\|}{\|A^*\| + \|A^{**}\| + \|A^{***}\|},$$

where $\|M\|$ means "sum of the elements of matrix M", A** is the block of all the elements of the columns belonging to A* with the exception of the elements of A* and A*** means the same for the rows. For example, if A* is the block defined above, we will have:

$\|A^{**}\| = \sum(a_{j1} + a_{j2})$ and $\|A^{***}\| = \sum(a_{1j} + a_{2j})$ for j = 3, 4, …, N.

Based in the autonomy degree we can define a block dependency degree as:

(2.2) $G_d(A^*) = 1 - G_a(A^*)$.

We don't consider blocks of A for which A*, A** and A*** are null matrices. In such an extreme case we would not even be able to define an autonomy degree and the block of A would be irrelevant.

We call *admissible* a matrix A that has no irrelevant blocks. In all that follows, we consider only admissible matrices. And of course we loose nothing with this restriction, because a block of A that is irrelevant is not really a part of a system.

It's easy to see that in a matrix A of order N there are $2^N - 2$ blocks A* (because there are $\sum \binom{N}{k}$ blocks A* with k = 1, ... , N-1).

So, we can define the (raw) dependency degree of system A as:

(2.3) $G^*(A) = \dfrac{\sum_k G_d(A_k)}{2^N - 2}$.

for which k varies from 1 to $2^N - 2$ and A_k represents a square block that includes the main diagonal.

It is also easy to see (Amaral, 1999) that:

a) $G^*(kA) = G^*(A)$ for k > 0
b) $0 \le G^*(A) \le 1$
c) $G^*(\hat{A}) \ge G^*(A)$, where $\hat{A}$ is obtained from A, making null the main diagonal elements.
d) $G^*(A) = 0$, iff A is diagonal.

To reach our final definition of the degree of dependency we must correct the above one by a scaling factor that is a function of N. To see this we may note that, with N>2, no matrix A has $G^*(A) = 1$. In fact, we can prove the following theorem:

THEOREM 1: For any matrix A of order N, the maximum value of $G^*(A)$ is:

$$\frac{2^N - 2^{N-2} - 1}{2^N - 2}.$$

Proof:

The proof proceeds in two steps. First, we verify that we have:

$$G^*(A) \leq \frac{2^N - 2^{N-2} - 1}{2^N - 2}, \tag{2.4}$$

for any matrix A. Next, we will give an example of a matrix A of order N for which the equality applies.

As we are dealing with the maximum of G*(A), by item c) above we can consider only matrices that have null main diagonal elements.

Let A_k^i be a part (i) of A of order k. Then:

$$G_d(A_k^i) = 1 - \frac{m_{ik}}{m_{ik} + n_{ik}}, \tag{2.5}$$

where m_{ik} is the sum of all the elements of the block and n_{ik} is the sum of all the elements of A that belong to one column or one row of the block but don't belong to it. That is,

$$G_d(A_k^i) = \frac{n_{ik}}{m_{ik} + n_{ik}}. \tag{2.6}$$

If we add to the numerator and to the denominator of this expression the quantity p_{ik}, which is the sum of all the elements of A that are not terms of m_{ik} and n_{ik}, we obtain, given that all the values are non negative,

$$\frac{n_{ik} + p_{ik}}{m_{ik} + n_{ik} + p_{ik}} \geq \frac{n_{ik}}{m_{ik} + n_{ik}}. \tag{2.7}$$

As the left side denominator doesn't change either with i or k, because it is the sum ||A|| of all the elements of A, we can add in i for each k>1 (for k=1 each block of an element, which is null, contributes with one unity for the degree of dependency of A) and then add for all the k > 1 to obtain:

$$(2.8) \quad \frac{\sum_k \sum_i (n_{ik} + p_{ik})}{\|A\|} \geq (2^N - 2)G^*(A) - N,$$

in which N is the contribution to G*(A) of the N blocks composed of one single element only (null) in the main diagonal.

For each k>1, each element of A not belonging to the main diagonal enters $\binom{N}{k} - \binom{N-2}{k-2}$ times in the correspondent term of the summation, because it enters in all the terms of $\sum_i$ except in those corresponding to the blocks it belongs to.

So, as the elements of the main diagonal are null the expression (2.8) can be written, for k from 2 to N-1, as:

$$(2.9) \quad \sum_k \binom{N}{k} - \binom{N-2}{k-2} \geq (2^N - 2)G^*(A) - N, \quad \text{that is}$$

$$(2.10) \quad (2^N - N - 2) - (2^{N-2} - 1) \geq (2^N - 2)G^*(A) - N, \text{ and finally:}$$

$$(2.11) \quad \frac{2^N - 2^{N-2} - 1}{2^N - 2} \geq G^*(A), \text{ as we wanted to prove.}$$

As the second step of the proof it will suffice to give an example of a matrix for which the maximum is attained. Consider the case of a matrix A such as:
$a_{ij} = 0$, for i ≠ j, i > 1, $a_{ii} = 0$ and $a_{1j} \neq 0$ for j > 1.

This is an admissible matrix that attains the maximum value of the dependency degree. And so the theorem is proved. □

With this theorem, we can finally define the dependency degree of an admissible matrix A.

Dependency degree $G(A)$ of A is the number:

$$(2.12)\quad G(A)=\frac{(2^N-2)G^*(A)}{2^N-2^{N-2}-1}.$$

We have $0 \le G(A) \le 1$ and for any order N there are matrices A that have $G(\cdot) = 0$ and $G(\cdot) = 1$. Besides, it's obvious that all the properties a) to d) above defined for G^* also apply to G.

Another interesting result, which we will use later on, arises from the following concept. Let A be a matrix of non negative elements such as $a_{ij}+a_{ji}=k>0$ for some pairs (i,j) and $a_{ij}+a_{ji}=0$ for the remaining pairs and some of the a_{ii} equal to k and others equal to 0. A matrix C is called **congruent (m,k)** with A if:
$c_{ij}+c_{ji}=m>0$ for all the pairs (i,j) whose sum $a_{ij}+a_{ji}$ is equal to k in matrix A, and $c_{ij}+c_{ji}=0$ for the remaining pairs, $c_{ii}=m$ if $a_{ii}=k$, and $c_{ii}=0$ if $a_{ii}=0$.

THEOREM 2: Let A be a matrix satisfying the above conditions. Then, if C is congruent with A, $G(C) = G(A)$.

Proof:
Let C be congruent (m,k) with A. Let us first multiply A by the value m/k. Then, we obtain a matrix A^* such as $G^*(A^*) = G^*(A)$ and where $a^*_{ij}+a^*_{ji}=m$ or 0 and the same for the a_{ii}, A* being congruent (m,k) with A. On the other hand, since in the calculation of all the terms of $G^*(C)$, c_{ij} enters in the same terms as c_{ji}, if we sum the values of the two we don't change the value of $G^*(C)$ relatively to $G^*(A^*)$ even for very different values of c_{ij} or c_{ji}, since the sum of the two has the same value (m) as the corresponding sum in A*. Then, G*(C) = G*(A*) and as G*(A*) = G*(A) we have G*(C) = G*(A) and the same for G. □

We can now find an upper limit for the degree of dependency of a matrix A of order N.

The contribution for G(A) of each block A^i_k of A containing the corresponding elements of the main diagonal is:

(2.13) $$G_{ik} = 1 - \frac{\|A^i_k\|}{\|A^i_k\| + c_{ik}},$$

where c_{ik} is the sum of the elements not belonging to A^i_k but belonging to a row or column of A^i_k.

We have:

(2.14) $$G_{ik} = \frac{c_{ik}}{\|A^i_k\| + c_{ik}}.$$

Adding to the numerator and to the denominator all the other elements of A we have

(2.15) $$G_{ik} \le \frac{\|A\| - trace(A^i_k) - d_{ik}}{\|A\|},$$

where d_{ik} is the sum of the elements of the block A^i_k that are not elements of the main diagonal.

Summing for each k and each i (that is, for $2^N - 2$ terms) the denominator doesn't change either with k or i, so that we get:

(2.16) $$G^*(A) \le \frac{(2^N - 2)\|A\| - (2^{N-1} - 1)trace(A) - (2^{N-2} - 1)(\|A\| - trace(A))}{\|A\|}$$

Explanation for the term $(2^{N-1} - 1)$ trace (A): each element a_{ii} of the main diagonal enters only once in a (1 x 1) block and $\binom{N-1}{m-1}$ times in an (m x m) block for m = N-1.

The other term that needs explanation is $(2^{N-2} - 1)$ ($\|A\|$ - trace(A)).

Each term of d_{ik}, for instance a_{jl} ($j \neq l$) enters once in the block of order 2 formed with lines or columns j and l; then enters $\binom{N-2}{m-2}$ times in blocks of order m (m = 3, …, N-1).

Therefore, for each term a_{jl} of d_{ik} the sum gives for m = 2, ..., N-1:

$\sum \binom{N-2}{m-2}$, that is $(2^{N-2} - 1)a_{jl}$.

From (2.16) we get:

$$(2.17)\quad G^*(A) \le \frac{(2^N - 2^{N-2} - 1)\|A\| + (2^{N-2} - 2^{N-1})trace(A)}{\|A\|}, \text{ so that:}$$

$$(2.18)\quad G(A) \le 1 - \frac{2^{N-1} - 2^{N-2}}{2^N - 2^{N-2} - 1}\left(\frac{trace(A)}{\|A\|}\right).$$

For high values of N, this upper limit is approximately:

$$\approx 1 - \frac{1}{3}\frac{trace(A)}{\|A\|}$$

After considering the dependency effect we introduce next an index for the *network* effect.

2.2 The network effect

First of all, it is convenient to note that we limit the definition of this indicator to non negative, admissible matrices that satisfy the following condition:

$\sum_i a_{ij} \le 1$, with the inequality strictly verified, for at least one j. Matrices under these conditions are called *productive matrices.*

In order to define the network indicator for a productive matrix we need to recall the concept of a decomposable matrix. A is a *decomposable matrix* if and only if by permutations of rows and columns it can be put as:

$A = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}$, in which A_1 and A_3 are square blocks (not necessarily of the same order) and *0* is a block of zeros.

In matrix theory it can be proved that for a productive, indecomposable matrix A, the inverse $(I-A)^{-1}$ exists and all its elements are positive. A decomposable matrix also have the corresponding non-negative $(I-A)^{-1}$ matrix, although not all of its elements are positive.

When the system is represented by an A matrix such as X = AX + Y, as the Leontief model is, then the fact that the matrix A is indecomposable means that all the sectors are connected directly or indirectly. The sectors may be not directly connected, that is, we can have some elements a_{ij} null, despite the fact that A is indecomposable. However, we know for sure that if A is indecomposable, a variation in any component of vector Y causes a variation in all the components of vector X.

This brings us to the definition of **decomposability degree** of the productive matrix A, h(A):

(2.20) $h(A) = \frac{Z(A)}{N^2 - N}$, in which Z(A) is the number of zeros of matrix $(I-A)^{-1}$.

It is easy to prove that h(A) is equal to zero if and only if A is indecomposable and it is equal to one if and only if A is a diagonal matrix with the elements of the main diagonal satisfying the condition of a productive matrix. In the other cases h(A) is between zero and one.

The **network effect indicator**, H(A) will be:

(2.21) H(A) = 1 – h(A).

When all the sectors are directly or indirectly connected (matrix A indecomposable) the network effect will be maximum (H(A) =1). When they are not connected, either directly or indirectly (matrix A diagonal) the network effect will be minimum (H(A) = 0).

However, to make this indicator useful, we need to verify if for every positive k that keeps the condition of productive matrix, that is, for every positive $k < \frac{1}{\max_j \sum_i a_{ij}}$,

we have H(kA) = H(A).

This is obviously true for an indecomposable matrix, because kA is also an indecomposable matrix. If A is decomposable it can be verified in the following way.

Let A be a decomposable matrix such that by an appropriated permutation of rows and columns the blocks A_1 and A_2 are indecomposable. Then, using block inversion we have:

$$(I-A)^{-1} = \begin{bmatrix} (I-A)^{-1} & (I-A_1)^{-1}A_2(I-A_3)^{-1} \\ 0 & (I-A_3)^{-1} \end{bmatrix}, \text{ and}$$

$$(I-kA)^{-1} = \begin{bmatrix} (I-kA)^{-1} & (I-kA_1)^{-1}kA_2(I-kA_3)^{-1} \\ 0 & (I-kA_3)^{-1} \end{bmatrix}$$

Since $(I-kA_1)^{-1}$ and $(I-kA_3)^{-1}$ have only positive values, the only null elements if any present in both inverses, $(I-A)^{-1}$ and $(I-kA)^{-1}$, besides the null blocks, are those related to the case of A_2 being null in both cases. Therefore, the multiplication by k doesn't change the decomposability degree.

In case A_1 or A_3 are decomposable we can proceed the same way, successively until needed, that is, until we find two sub-blocks indecomposable within a decomposable sub-block. This means that this decomposable sub-block preserves the zeros of the inverses from the case A to the case kA, so we can conclude that the decomposable sub-block containing this one also preserves it and so on until we reach the blocks of the matrices A and kA. So, H(kA) = H(A), which is of course a desirable property for the index.

We have at last all we need in order to define the complexity index that includes both the dependency and the network effects.

2.3. The complexity index

The complexity indicator, in the sense of the intensity of interrelatedness between the parts of productive matrix A is:

(2.22) I(A) = G(A) x H(A)

As can be easily seen by what was said before, 0 ≤ I(A) ≤ 1; I(A) = 0 if and only if A is diagonal; I(kA) = I(A) for all k positive that keep the condition of productive matrix.

On the other side, I(A) can reach the value one, but only when the matrix A is indecomposable (although the inverse is not true). An example is the matrix A defined by: $a_{11} = 0$, a_{ji} = 0, (j≠1 and i ≠1), $a_{1j} = b$, $a_{j1} = b$, with $b < \frac{1}{N-1}$, which is indecomposable (so that H(A) = 1) and has G(A) = 1, since a matrix A^* such that $a^*_{1j} = 1$, $a^*_{11} = 0$ and $a^*_{ij} = 0$ for i > 1 is admissible and congruent (1, 2b) with A (see theorem 2), and has G(A^*) = 1 (see page 8).

This calls also our attention to an important issue when we deal with complexity. A system that has a structural matrix such that matrix A above has the maximum of complexity: each part suffers the influence of the others (network effect) and this influence is a relatively high one (dependency effect). The complexity of the system, viewed as interrelatedness between its parts is, therefore, very high. And, notwithstanding, the matrix describing the system is very simple.

We have here an example of what we said in the introduction about the difference between complexity for someone outside the system and having all the relevant information about it and complexity for those inside the system. This is, we think, an important point: not always the simplicity of the functional describing a system for someone outside it is a good indicator of its complexity for those immersed in it.

3. An application to the OECD countries

In this section we apply our measure of complexity as interrelatedness (I) to a number of OECD countries for which data on input-output matrices were available on a comparable basis. The original industries are listed in Appendix 1. In order to increase the comparability of data and to avoid a prohibitive number of computations, we aggregated the original data to a smaller number of industries. The list of seventeen industries used in most of this section is presented in Appendix 2. All the computations are made using domestic input-output matrices in current prices.

Let us first consider the case of the United States. The results of our index of complexity for this country are presented in table 1, for different level of aggregation and from 1972 to 1990. Given the high level of aggregation, for both the USA and the other countries considered in this paper, the network effect is equal to unity, and so doesn't influence, in this case, the complexity index.

Table 1. Index of Complexity for the United States

	Number of industries			
	3	10	15	20
1972	0.588	0.675	0.712	0.723
1977	0.618	0.681	0.718	0.731
1982	0.639	0.704	0.732	0.739
1985	0.637	0.709	0.742	0.751
1990	0.614	0.699	0.731	0.740

The results in table 1 suggest two broad comments. First, complexity increases, in general, with the level of disaggregation, that is, the number of industries considered in the input-output matrices. For example, if we compare the results for three with those for twenty industries, in the last case the index augments by around twenty per cent. Second, the evolution during the last twenty years is not linear: while there is an increase in complexity until mid-eighties or so, the index for 1990 is always below the level in 1985.

The results for the ten OECD countries, for seventeen industries, are presented in table 2. The years available are not exactly the same for all the countries and the appendix 3 shows the available year in each case.

In what concerns evolution the results are mixed. While there are cases, like the United States, where the index increased from early seventies to 1990, there are also cases (Japan and Canada) with a reduction in the measured complexity during these two dates.

Table 2. Complexity index, 17 industries

	Early-70's	Mid-70's	Early-80's	Mid-80's	Early-90's
Australia	0.747	0.731		0.756	0.767
Canada	0.779	0.778	0.765	0.766	0.765
Denmark	0.786	0.792	0.787	0.795	0.791
France	0.737	0.741	0.758	0.757	0.748
Germany		0.732		0.742	0.753
Italy				0.764	
Japan	0.740	0.740	0.742	0.723	0.719
Netherlands	0.754	0.769	0.777	0.782	
United Kingdom	0.729	0.770		0.751	0.737
United States	0.712	0.720	0.731	0.741	0.732

The index obtained for the ten countries is not dramatically different from one to another. The lower value is obtained for the United States in 1972 and the higher value is for Denmark in 1985. That is, the largest country in our database tends to have the lowest index of complexity. And the smallest country tends to have the highest value. Is there a negative correlation between dimension and complexity of an economy?

While the countries in our sample are not dramatically different in terms of level of development, we made a very simple exercise, regressing the index of complexity on the dimension and the level of development (as measured by *per capita* income). The results are presented in table 3 and they do suggest a negative correlation between complexity and dimension of an economy. Our results are robust to alternative specifications for the dimension variable: population, current GNP or even (not shown in the table) GNP in terms of purchasing power parity (constant prices).

Table 3. Regression of I on dimension and development
(log-log model)

log(POP)	-0.021 *	-0.021 *	
log(GNP)			-0.013 *
log(GNP/POP)	0.00782		
N	42	42	42
R2	0.695	0.664	0.394

* significant at the 5% level

In section 2 we presented un upper limit for G(A), in terms of the trace of matrix A. When, as it is the case here, the network effect is equal to one, this is also a limit for our complexity index. However, given real data, we can, in principle, give a more accurate relation between the two concepts. We further investigate this issue here, for our sample of OECD countries.

So, lets define T* as:

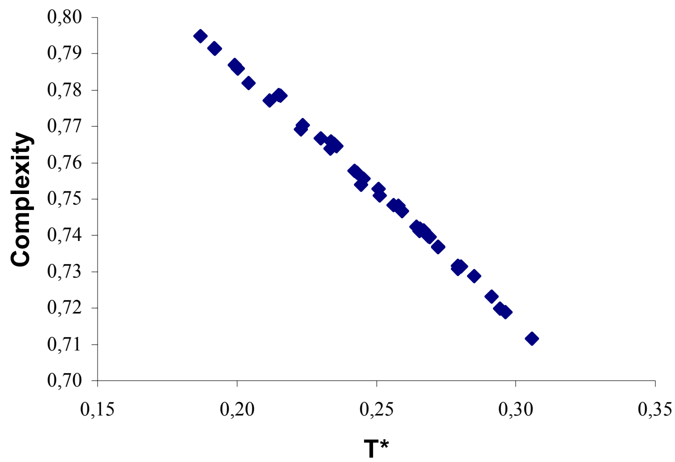
$$T^* = \frac{trace(A)}{\|A\|}.$$

The approximation found for G(A) (that is, our index of complexity (I) with unitary network effect), when the number of industries is very high, is given by:

$$I \approx 1 - \frac{1}{3} T^*.$$

This may be a relatively rude approximation in certain cases. For example, when we have a diagonal matrix, the approximation gives $\frac{2}{3}$ since T*=1. However, we know that, for a diagonal matrix, the correct value is zero.

In our case, the following figure plots our index of complexity against T*, with 42 observations (Germany has an additional year 1988) and for the case of seventeen industries.



Complexity (I) and normalized trace (T*)

We get an almost perfect negative relationship between the two variables, with a correlation coefficient of –0.998.

A simple regression of I on T* gave the following results (standard errors in parenthesis),

$$\boxed{\begin{array}{l} \hat{I} = \;\; 0.93 \;\; - \;\; 0.694T^* \\ \quad\;\; (0.0016) \;\; (0.0062) \\ \quad\;\; N = 42; \; R^2 = 0.997 \end{array}}$$

Imposing the restriction that T*=1 implies I=0, we obtain,

$$\boxed{\begin{array}{l} \hat{I} = 1.00162 \; - \; 1.00162\mathrm{T}^* \\ \qquad\qquad\qquad\quad (0.0019) \\ \quad N = 42; \; R^2 = 0.8 \end{array}}$$

and, with this condition, but allowing for a quadratic term we have,

$$\hat{I} = 0.902 \; + \; 0.407T^* - \; 1.309T^{*2}$$
$$(0.0018)\ (0.0071)$$
$$R^2 = 0.998; \; \hat{\sigma} = 0.001$$

In any case, the normalized trace gives a good predictor for our index, principally for high number of industries where its calculation becomes computationally prohibitive but is very ease to obtain using T*. Table 4 shows how the correlation coefficient changes with the number of industries in our sample.

Table 4. Correlation between I and T* for different number of industries

Number of industries	corr(T*,I)
2	-0.947
3	-0.988
5	-0.989
10	-0.997
17	-0.998

This link between complexity and the trace of an input-output matrix is in agreement with our definition, since increasing dependency augments complexity and increasing "autarky" reduces it. Apparently, there is a tendency for larger countries to have less specialized sectors than smaller ones for countries with similar level of development.

4. Concluding Remarks

In recent years there has been in many fields of research a growing interest in studying complexity.

Complexity is an important feature of most of the dynamic systems, physical, biological and social.

Although a universally acceptable definition of complexity is still lacking it emerges as one of the most prominent characteristics of this concept the mutual dependency and interaction between different agents or elements of a whole. That is why by its very nature, the Leontief input-output analysis is a convenient framework for the study of complexity of economic systems. Most of the contributions in this tradition however are mainly concerned with quantifying structural change at the sectoral level and deal with complexity only in an implicit way.

This paper treats economic complexity explicitly, discussing an important issue largely ignored in the relevant literature: the distinction between complexity to someone outside the system having all the relevant information about it (*outside* perspective), and complexity to someone immersed in the system and having only limited information about it (*inside* perspective). The main contribution of the paper is to propose a measure of (inter-industry) complexity as interrelatedness, particularly suited to quantifications related to the inside perspective.

We present in the paper an empirical application of this measure of complexity to several countries based on the OECD input-output database. Apparently, there is a tendency for diminishing complexity as the economies grow, and small countries tend to show a grater complexity than large ones. This surprising result may be explained considering differences in the pattern of sectoral specialization and the evolution of intra-sectoral trade, but this conjecture needs further research.

References

Adami, C. (2002), What is complexity?, *BioEssays*, 24: 1085-1094.

Amaral, J. (1999), "Complexity and Information in Economic Systems", in Louçã, F. (ed.), *Perspectives on Complexity in Economics,* ISEG-UTL, Lisbon Technical University.

Aroche-Reyes, F. (2003), A qualitative input-output method to find basic economic structures, *Papers in Regional Science,* 82, 581-590.

Arthur, B. (1999), Complexity and the Economy, *Science*, 284, 107-109.

Campbell, J. (1975), Application of graph theoretic analysis to inter-industry relationships: The example of Washington State, *Regional Science and Urban Economics,* 5: 91-106.

Dridi, C. and G. Hewings (2002), "Industry Associations, Association Loops and Economic Complexity: Application to Canada and the United States", *Economic Systems Research,* 14(3), 275-96.

Durlauf, S. (2003), "Complexity and Empirical Economics", Santa Fe Institute Working Paper.

Rosser, J. (1999), On the Complexities of Complex Economic Dynamics, *Journal of Economic Perspectives,* 13(4): 169-192

Schnabl, H. (1994), The evolution of production structures analysed by a multi-layer procedure, *Economic Systems Research*, 6: 51-68.

Sonnis, M. and G. Hewings (1998), Economic complexity as network complication: Multiregional input-output structural path analysis, *The Annals of Regional Science,* 32(3): 407-436.

Yan, C. and E. Ames (1965), Economic Interrelatedness, *The Review of Economic Studies*, 32(4): 290-310.

Appendix 1

Industries in the OECD input-output matrices
1 Agriculture, forestry & fishing
2 Mining & quarrying
3 Food, beverages & tobacco
4 Textiles, apparel & leather
5 Wood products & furniture
6 Paper, paper products & printing
7 Industrial chemicals
8 Drugs & medicines
9 Petroleum & coal products
10 Rubber & plastic products
11 Non-metallic mineral products
12 Iron & steel
13 Non-ferrous metals
14 Metal products
15 Non-electrical machinery
16 Office & computing machinery
17 Electrical apparatus, nec
18 Radio, TV & communication equipment
19 Shipbuilding & repairing
20 Other transport
21 Motor vehicles
22 Aircraft
23 Professional goods
24 Other manufacturing
25 Electricity, gas & water
26 Construction
27 Wholesale & retail trade
28 Restaurants & hotels
29 Transport & storage
30 Communication
31 Finance & insurance
32 Real estate & business services
33 Community, social & personal services
34 Producers of government services
35 Other producers
36 Statistical discrepancy

Appendix 2

Aggregation of OECD input-output matrices for 17 industries	
1	Agriculture, mining & quarrying
2	Food, beverages & tobacco
3	Textiles, apparel & leather
4	Wood and paper
5	Chemicals, drugs, oil and plastics
6	Minerals and metals
7	Electrical and non-elect. equipment
8	Transport equipment
9	Other manufacturing
10	Electricity, gas & water
11	Construction
12	Wholesale & retail trade
13	Restaurants & hotels
14	Transport & storage
15	Communication
16	Finance & insurance
17	Other sectors

Appendix 3

OECD Input-Output database coverage

	Early-70's	Mid-70's	Early-80's	Mid-80's	Early-90's
Australia	1968	1974		1986	1989
Canada	1971	1976	1981	1986	1990
Denmark	1972	1977	1980	1985	1990
France	1972	1977	1980	1985	1990
Germany		1978		1986	1990
Italy				1985	
Japan	1970	1975	1980	1985	1990
Netherlands	1972	1977	1981	1986	
United Kingdom	1968	1979		1984	1990
United States	1972	1977	1982	1985	1990