

# IDENTIFICAÇÃO DE CLUSTERS INDUSTRIAIS NA COLÔMBIA BASEADOS NA TEORIA DE GRAFOS

JUAN C. DUQUE  
SERGIO J. REY  
DAIRO A. GÓMEZ\*

---

\*Os autores agradecem ao revisor anônimo da ESPE por seus comentários claros e úteis durante o processo de revisão. As conclusões e opiniões são aquelas dos autores.

Os autores são, em sua ordem: Investigação em Território Economia (RISE-grupo).Departamento de Economia. Universidade EAFIT; Investigação em Território Economia (RISE-grupo).Departamento de Economia. Universidade EAFIT; e Faculdade de Ciências Geográficas Arizona State University.

Correio eletrônico:  
jduquec1@eafit.edu.co  
srey@asu.edu  
dgomezca@eafit.edu.co

Documento recebido no dia 2 dezembro de 2008; versão final aceita no dia 27 abril de 2009.

Este paper apresenta uma nova forma de identificar e entender os clusters industriais na economia colombiana. Esta análise baseia-se numa metodologia proposta recentemente por Duque and Rey (2008) na qual se aplica a teoria de redes para simplificar os complexos vínculos comerciais entre indústrias presentes numa matriz insumo-produto. Em comparação com outras técnicas existentes na literatura, esta inovadora técnica permite não só classificar cada indústria dentro de um cluster, mas também entender como as indústrias estão relacionadas dentro de seu cluster. Esta metodologia oferece uma aproximação conciliadora entre dois pontos de vista radicalmente diferentes com relação à unidade econômica sobre a qual se devem desenhar as estratégias para a atribuição de recursos: a visão de Porter baseada no apoio aos clusters contra a visão de Hausmann baseada no apoio a indústrias estrategicamente selecionadas.

**Classificação JEL:** C67, D57, L22.

**Palavras chave:** clusters industriais, teoria de redes, insumo-produto, análise de impacto.

## IDENTIFICACIÓN DE *CLUSTERS* INDUSTRIALES EN COLOMBIA BASADOS EN LA TEORÍA DE GRAFOS

JUAN C. DUQUE  
SERGIO J. REY  
DAIRO A. GÓMEZ\*

Este artículo presenta una nueva forma de identificar y entender los *clusters* industriales en la economía colombiana. El análisis se basa en una metodología propuesta recientemente por Duque y Rey (2008), en la cual se aplica la teoría de redes para simplificar los complejos vínculos comerciales entre las industrias presentes en una matriz insumo-producto. En comparación con otras técnicas existentes en la literatura, esta novedosa técnica permite no sólo clasificar cada industria en un *cluster*, sino también entender cómo las industrias están relacionadas dentro de su *cluster*. Esta metodología ofrece una aproximación conciliadora entre dos puntos de vista radicalmente diferentes con respecto a la unidad económica sobre la que deben diseñarse las estrategias para la asignación de recursos: la visión de Porter basada en el apoyo a los *clusters*, en contraste con la perspectiva de Hausmann, quien apoya las industrias seleccionadas estratégicamente.

**Clasificación JEL:** C67, D57, L22.

**Palabras clave:** *clusters* industriales, teoría de redes, insumo-producto, análisis de impacto.

---

\*Los autores agradecen al evaluador anónimo de ESPE por sus claros y útiles comentarios durante el proceso de revisión. Los resultados y opiniones son responsabilidad de los autores.

Los autores son en su orden: Investigación en Economía Espacial (RISE-group). Departamento de Economía. Universidad EAFIT; Investigación en Economía Espacial (RISE-group). Departamento de Economía. Universidad EAFIT; y Escuela de Ciencias Geográficas Universidad Estatal de Arizona.

Correo electrónico:  
jduquec1@eafit.edu.co  
srey@asu.edu  
dgomezca@eafit.edu.co

Documento recibido:  
2 de diciembre de 2008;  
versión final aceptada:  
27 de abril de 2009.

# IDENTIFYING INDUSTRY CLUSTERS IN COLOMBIA BASED ON GRAPH THEORY

JUAN C. DUQUE  
SERGIO J. REY  
DAIRO A. GÓMEZ\*

---

\*The authors thank the anonymous ESPE reviewer for his insightful and helpful comments during the review process. The usual disclaimer applies.

The authors are, in order, Research in Spatial Economics (RISE-group), Department of Economics, EAFIT University; Research in Spatial Economics (RISE-group), Department of Economics, EAFIT University; and School of Geographical Sciences Arozoza State University.

E-mail:  
jduquec1@eafit.edu.co  
srey@asu.edu  
dgomezca@eafit.edu.co

Document received:  
2 December 2008;  
final version accepted:  
27 April 2009.

This paper presents a new way to identify and understand the industry clusters in the Colombian economy. The analysis relies on a recent methodology proposed by Duque and Rey (2008) where intricate input-output linkages between industries are simplified using network analysis. In addition to other techniques for cluster identification available in the literature, this novel methodology allows us not only to classify each industry in a given cluster, but also to understand how industries are related to each other within their clusters. This methodology offers a conciliatory approach to two radically different views about the economic unit from which policy makers should design their strategies for resource allocation: Porter's cluster strategy versus Hausmann's industrial targeting.

**JEL classification:** C67, D57, L22.

**Keywords:** industry clusters, graph theory, input-output, impact analysis.

## I. INTRODUCTION

Industry clusters have been a matter of investigation since the beginning of twentieth century. One of the first references dates back to 1890 when Marshall (1890) introduced the concept of “agglomeration economies” as the benefits derived from the synergy that industries can generate by locating near each other. While the concept of “agglomeration economies” focuses on the spatial allocation of industries, the term “industry cluster”, introduced by Porter (1990), defines the benefits derived from the vertical or horizontal relationships between the industries of a given economy.

Vertical clusters are those that gather industries characterised by buyer-supplier relations. While horizontal clusters include industries that share a common market for final goods, or use same the technology or employees, or need a similar natural resource (Porter, 2003, p. 205, translated).

Kaufman, Gittlell, Merenda, Naumesand and Wood (1994) stress that cluster analysis offers guidance to policy makers in the identification of a state’s competitive advantage. In a same way, Doeringer and Terkla (1995) state that by widening the focus of development policies, cluster analysis offers the possibility of integrating non-export as well as export-based industries into regional growth strategies.

Nowadays, industry cluster identification is still an important research topic in public, private and academic sectors. On the one hand, a great deal of cluster observatories have been created worldwide as an important resource for policy makers and planners who are concerned with the strategic and tactical deployment of resources

(Europe INNOVA, 2007). On the other hand, cluster identification is also beneficial for industries since it provides insights about potential clients and suppliers, alternative markets, and a way to clarify the role of the industry in the economy (High level advisory group on clusters, 2008).

The benefits of liberalized trade have increased the search for trade agreements between Colombia and other countries around the world. For example, the recently approval of a free trade agreement with the European Free Trade Association (EFTA), or the possibility of a trade agreement with the US, have increased the need for a deeper understanding of Colombia's productive structure. The identification of industry clusters in this country would help to detect strengths and weaknesses when facing the arrival of products and services from external competitors, and to implement strategies to improve the competitiveness and innovation level of local industries.

Colombia's central government is aware of this situation. Proof of this can be found in the fact that the development of *world-class clusters* is one of the five pillars of the national policy of competitiveness and productivity (Sistema Nacional de Competitividad, 2008).

Nevertheless, there is a clear lack of research on the identification of industry clusters at national level. Two studies on this topic have been commissioned by the government. The first study dates back to 1993 when the central government<sup>1</sup> commissioned a study to assess the competitiveness level of Colombia and to identify the most representative clusters to establish the bases of the Colombia's competitive policy (Monitor Company, Inc, 1993)<sup>2</sup>. In order to identify these clusters the authors combine statistical information to measure the international competitiveness of each sector<sup>3</sup>, as well as qualitative information obtained from countrywide surveys of public and private decision makers. The study concludes that the key Colombian horizontal clusters are: Petrochemical, Flowers, Leather, Textiles, Fruit Juices, Capital Goods, Graphic Arts and Tourism.

---

1 Ministry of Development, the Instituto de Fomento Industrial, Confecámaras and Bancoldex.

2 It is worth to mentioning that in 1990 Colombia experienced its economic openness process.

3 The authors utilised variables such as the nation's share of world exports for each sector, degree and type of foreign ownership, number of employees, amongst others.

The second study is an advisory service contracted by the National Planning Department in 2007 to identify the key industries at national and regional level that could lead growth in Colombia (Hausmann and Klinger, 2007). In this study, the authors argue that when a government is deciding where to allocate resources, those resources do not necessarily have to go to the biggest industries (“picking winners strategy”). Instead, they propose a method “to help better direct public resources to support the process of structural transformation and export growth” (Hausmann and Klinger, 2007, pp. 7). Thus, the key industries are selected based on the following criteria: i) industries with a high potential of export growth, and ii) industries that offer the possibility of creating new products with the current capabilities. The selected industries are: Medicaments; Fabricated metal products; Yarn, wool & fabrics; Shipbuilding & repairing; Motor vehicles and Electrical appliances and housewares.

This paper presents the results of the identification of industry clusters and inter-industry networks based on 2005 input-output tables for Colombia. The methodology applied in this paper has been recently developed by Duque and Rey (2008) in which an algorithm, based on network theory, is proposed to identify the most representative vertical clusters in a given economy<sup>4</sup>.

Duque and Rey’s methodology can be seen as a conciliatory approach to the studies presented above, since it recognises the importance of clusters as an engine of development. But, at the same time, takes into account that there are constraints that force governments to strategically allocate their scarce resources. In this sense, this methodology indicates the key industry within each cluster, where the resources can generate the biggest impact in the cluster and, therefore, in the whole economy.

The outline of this paper is as follows. Section II describes the methodology that is applied to identify industry clusters. Section III offers the main results of applying this methodology using information from 2005 input-output tables for Colombia. Finally, Section IV summarizes the main findings and provides recommendations for future work.

---

<sup>4</sup> For overviews on regional cluster research see Steiner (1998); Bergman and Feser (1999). See also Rey and Mattheis (2000) for a taxonomy of industry clusters identification methods.

## II. METHODOLOGY

The algorithm applied in this paper is known as the network-based industry clusters (NBIC). This algorithm was recently developed by Duque and Rey (2008) and it is designed to identify vertical industry clusters using I/O tables<sup>5</sup>.

Czamanski and Ablas (1979) provide a very clear definition of vertical industrial clusters: "... 'cluster' means a subset of industries of the economy connected by flows of goods and services stronger than those linking them to the other sectors of the national economy" (p. 62).

Graph 1 shows six graphs illustrating the main steps of Duque and Rey's methodology. It starts by representing the transactions between industries as a dense directed network in which each node represents an industry and the links joining each pair of nodes represent a transaction between two sectors (see Graph 1, Figure I).

From this initial representation, the NBIC algorithm starts a simplification process in order to transform the initial network into a network type known as a "tree network" in which each pair of nodes is connected by a unique path. This simplification process is possible given a set of assumptions associated to inter-industry transactions that allow the network reduction without losing valuable information.

The first assumption declares that the transactions between two industries flow mainly in one direction<sup>6</sup>. This assumption allows for the first reduction in the network which consists of deleting the link representing the smallest transaction between each pair of industries (see Graph 1 Figure II). Thus, having  $Z_{ij}$  representing sales from industry  $i$  to  $j$ , and  $Z_{ji}$  representing sales from industry  $j$  to  $i$ , the NBIC algorithm only keeps the largest of those two values, i.e.:

$$Z_{i,j} = \max(Z_{i,j}, Z_{j,i}) \tag{1}$$

---

5 More thorough treatments of regional input-output modeling can be found in Roepke, Adams and Wiseman(1974); Miller and Blair(1985); Hewings and Jensen (1986); Lahr (1993); Isard, Azis, Drennan, Miller, Saltzman and Thorbecke(1998).

6 Empirical evidence from San Diego California shows that 72.2% of the relative differences between opposite flows are greater than 90%.

Up to this point each link represents a transaction between industries. The next step transforms the directed network into an undirected network (see Graph 1 Figure III) by applying the following expression:

$$MRW_{i,j} = \max(RW_{i,j}^{out}, RW_{i,j}^{inp}) \quad (2)$$

$$\text{Where, } RW_{i,j}^{out} = \frac{z_{i,j}}{\sum_{j=1}^n z_{i,j}} \text{ and } RW_{i,j}^{inp} = \frac{z_{i,j}}{\sum_{i=1}^n z_{i,j}}$$

With this transformation, the link between industries  $i$  and  $j$  has no longer a direction. This link is now a weight ( $MRW$  - Maximum Relative Weight-) that represents how important is this link for either industry  $i$  or industry  $j$ .

In the next step, Duque and Rey apply an algorithm known as the Kruskal's algorithm (Kruskal, 1956) to reduce the network from a dense network to a sparse network type known as a "tree network" in which each pair of nodes is connected by a unique path. The links that remain in the network are the ones that better summarize the most important relationships between the industries in the economy (see Graph 1 Figure IV).

One important characteristic in the tree network is that the removal of a link "breaks" the tree network into two disconnected subnetworks. More generally, the removal on  $k$  links breaks the tree network into  $k - 1$  subnetworks, with each subnetwork representing a cluster.

Thus, the next step is to break the network into subnetworks or clusters. At this point, the logical question is: How many subnetworks?

In order to decide the number of clusters, the NBIC algorithm scores each industry according to three different criteria:

- Transaction volume ( $TV_i$ ): Measures the share of I-O transactions accounted for industry  $i$ , with respect the total I-O transactions in the economy.

$$TV_i = \frac{\sum_{j=1}^n z_{i,j} + \sum_{j=1}^n z_{j,i}}{\sum_{i,j} z_{i,j}} \quad (3)$$



- Adjusted transaction volume ( $TVa_i$ ): This criterion seeks to give a higher score to those industries which are related to many other industries in the economy.

$$TVa_i = \frac{2 \cdot TV_i}{(GINI_i^{out} + GINI_i^{inp})} \quad (4)$$

Where  $GINI_i^{out}$  measures how dispersed the outputs of industry  $i$  are; and  $GINI_i^{inp}$  measures how dispersed the inputs of industry  $i$  are.

- Market power ( $MP$ ): Consists of an iterative procedure that measures how important each industry is for its direct buyers and suppliers. Thus, those industries whose outputs (inputs) represent a high percentage of its client's inputs (supplier's outputs) will receive a higher  $MP$  score.

Finally, the NBIC algorithm uses factor analysis to merge the three criteria into one single vector value by extracting the first factor (referenced to as *factor score*). This factor score makes it possible to sort the industries by their level of importance for the economy<sup>7</sup>.

The last step in NBIC algorithm incorporates an iterative procedure that breaks the tree network into different number of clusters (subnetworks) such that each cluster contains one core industry (see Graph 1 Figure V and VI). Thus, if the number of clusters is set to four, then the core industries will be the first four industries with the highest factor scores. The optimal number of clusters  $k$  is the highest value of  $k$  such that the proportion of "weak" clusters in the economy does not exceed 50%.

The proportion of weak clusters is calculated in two different ways: i) based on internal linkages; and ii) based on external linkages:

- Internal linkages: It classifies each cluster according to how important the cluster is to the industries belonging to it. Two coefficients are calculated in order to carry out this classification. First, the intra-clusters Purchase Share coefficient ( $PS$ ) that measures the share of inter-industry purchases made by the cluster industries that are supplied by other industries within the cluster. The second measure, intra-cluster Sales Share coefficient ( $SS$ ) that measures the share of inter-industry sales made by the cluster industries that are purchased by other industries within the cluster:

---

<sup>7</sup> The factor analysis technique has been also applied by Czamanski (1971) within the context of industry clusters based on I-O tables.

$$PS_c = \frac{\sum_{i \in c}^n \sum_{j \in c}^n Z_{i,j}}{\sum_{i=1}^n \sum_{j \in c}^n Z_{i,j}}, \text{ with } Z_{i,i} = 0 \quad (5)$$

$$SS_c = \frac{\sum_{i \in c}^n \sum_{j \in c}^n Z_{i,j}}{\sum_{i \in c}^n \sum_{j=1}^n Z_{i,j}}, \text{ with } Z_{i,i} = 0 \quad (6)$$

Where  $Z_{i,j}$  represents the inter-industry deliveries from industry  $i$  to industry  $j$ .  $i \in c$  indicates that industry  $i$  is a member of cluster  $c$ . Thus,

- if  $PS_c > \overline{PS}$  and  $SS_c < \overline{SS}$ ,  $\Rightarrow c$  is a *purchase oriented cluster*;
- if  $PS_c < \overline{PS}$  and  $SS_c > \overline{SS}$ ,  $\Rightarrow c$  is a *sale oriented cluster*;
- if  $PS_c > \overline{PS}$  and  $SS_c > \overline{SS}$ ,  $\Rightarrow c$  is a *strong cluster*;
- if  $PS_c < \overline{PS}$  and  $SS_c < \overline{SS}$ ,  $\Rightarrow c$  is a *weak cluster*;

- External linkages: It classifies each cluster according to how important the cluster is to the larger economy. Two coefficients are calculated in order to carry out this classification, backward linkages ( $BL$ ), and forward linkages ( $FL$ ):

$$BL_c = \sum_{i \in c} \frac{\frac{1}{n} \sum_{j=1}^n l_{ji} X_i}{\frac{\sum_{i=1}^n \sum_{j=1}^n l_{ij}}{n^2} \sum_{j \in c} X_j} \quad (7)$$

$$FL_j = \frac{\frac{1}{n} \sum_{i=1}^n l_{ji}}{\frac{\sum_{i=1}^n \sum_{j=1}^n l_{ij}}{n^2}} \quad (8)$$

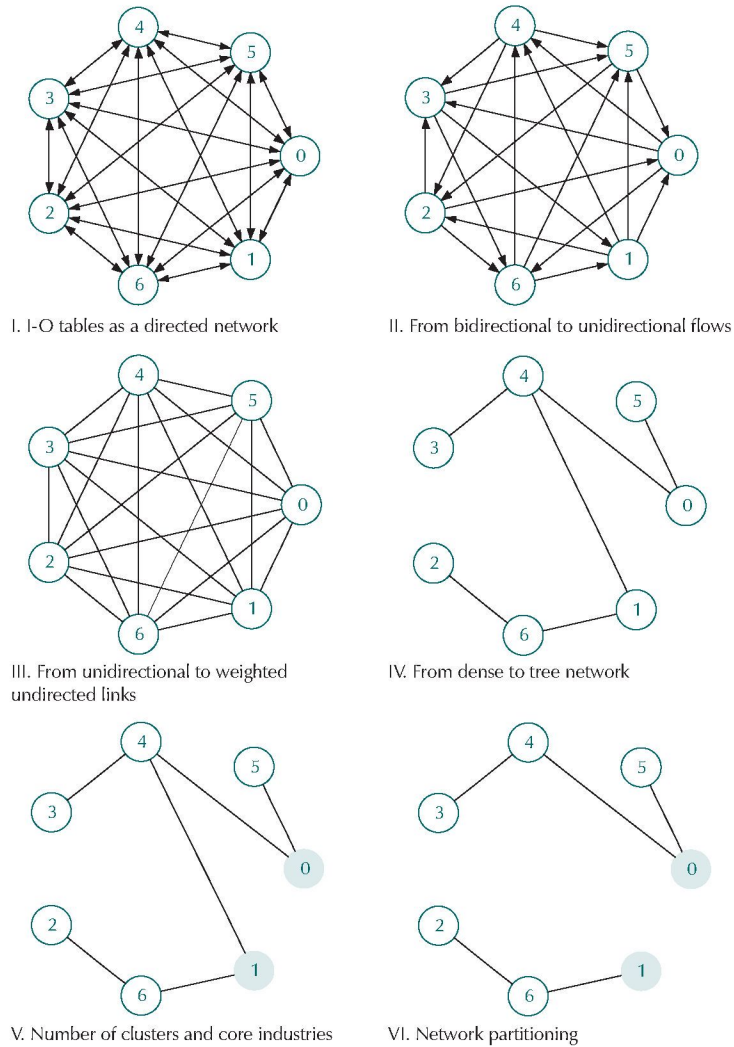
Where  $l_{ij}$  is each element of the the Leontief inverse for the regional input-output matrix.

- if  $BL_c > 1$  and  $FL_c > 1$ ,  $\Rightarrow c$  is a *key cluster*;
- if  $BL_c > 1$  and  $FL_c < 1$ ,  $\Rightarrow c$  is a *driver cluster*;
- if  $BL_c < 1$  and  $FL_c > 1$ ,  $\Rightarrow c$  is a *enabler cluster*;
- if  $BL_c < 1$  and  $FL_c < 1$ ,  $\Rightarrow c$  is a *weak cluster*;

The NBIC algorithm is currently being used by the San Diego East County Economic Development Council as a part of its project *Connectory.com* whose primary objective is to “link California businesses to each other and to provide information about the industrial/technology base of the nation’s largest economy”. (Connectory, 2008).

It is also being used by the San Diego Association of Governments (SANDAG) which serves as the “forum for regional decision-making” (Sandag, 2008)<sup>8</sup>.

Graph 1  
 Main Steps of Duque and Rey’s Network-based Industry Clusters Algorithm



Source: own elaboration.

<sup>8</sup> For additional information on this project see Rey, Duque, Schmidt and Li (2007) and Regal (2008).

### III. EMPIRICAL APPLICATION

#### A. DATA

The product utilization matrix used in this application was obtained from the National Administrative Department of Statistic (DANE). Since 1951, this institution has been responsible for planning, collecting, processing, analysing, and disseminating the official statistics of Colombia. In this particular case, we used the projected 2005 product utilization matrix. This matrix contains the input-output transactions (in millions of COPs) between 60 industries<sup>9</sup>. Table 1 presents a brief description of 2005p I-O table for Colombia.

Table 1  
Total Inter-industry Output and Input per Industry (millions of COP)

Code	Industry	Output	Input
1	Unroasted coffee	572,760	962,872
2	Other farming products	7,117,830	3,674,472
3	Animals and products derived from animals	12,440,416	5,854,613
4	Forestry products and wood extraction	620,008	85,699
5	Fish and other products from fishing	1,298,434	181,906
6	Lignite and peat	314,705	1,488,147
7	Oil, natural gas, uranium and thorium minerals	4,858,510	3,525,375
8	Metallic minerals	131,385	501,656
9	Other non-metallic minerals	2,308,560	275,007
10	Electricity and city gas	15,103,078	9,933,208
11	Water and sewerage services	1,014,610	1,097,957
12	Fish and meat	1,932,724	9,950,050
13	Animal and vegetable oil, fur and cakes	1,618,824	2,077,625
14	Dairy products	722,773	3,084,200
15	Mill products, starch and its products	4,643,820	7,048,511
16	Sugar	1,073,156	1,339,273
17	Processed coffee	291,699	578,128
18	Cacao, chocolate and sugar products	196,035	1,230,154
19	Other food products	768,981	2,006,509
20	Beverages	2,330,148	2,754,415
21	Tobacco products	8,769	198,090
22	Threads and textile fibre weaved	3,176,414	2,196,738

<sup>9</sup> This data is available at: <http://www.dane.gov.co/files/investigaciones/pib/anuales/anuales.zip>. (Accessed October, 2008).

Table 1 (continued)  
Total Inter-industry Output and Input per Industry (millions of COP)

Code	Industry	Output	Input
23	Textile items (except clothes)	-	-
24	Clothes	194,365	3,742,095
25	Leather and shoes	591,121	1,149,032
26	Wood products, cork, straw and plait materials	1,358,315	633,666
27	Cardboard and paper	5,290,013	3,395,399
28	Printing and similar goods	1,772,532	2,031,022
29	Refined oil products, nuclear combustibles and coke furnace products	11,104,862	6,548,943
30	Basic and elaborated chemical products (except plastic and rubber products)	18,947,064	9,517,910
31	Plastic and rubber products	7,468,413	3,559,427
32	Glass, glass products and other non-metallic products	6,662,406	3,877,940
33	Furniture and other transportable goods	634,985	2,283,662
34	Waste products	520,799	-
35	Common metals and metallic products, except machines and equipment	14,273,878	6,775,892
36	Special and general use machines	3,629,288	1,869,796
37	Other machines and electrical supply	5,260,476	1,797,704
38	Transport equipment	5,305,930	5,236,637
39	Construction	2,298,579	13,080,363
40	Civil engineering works	2,008,051	5,797,583
41	Wholesale and retail	-	12,993,976
42	Repair services of engines and domestic appliances	4,493,140	3,146,614
43	Hotel and restaurant services	1,803,366	6,667,451
44	Road transport services	5,682,332	11,782,006
45	Fluvial transport services	288,055	573,974
46	Aerial transport services	2,294,848	2,388,190
47	Complementary road transport services	3,665,573	1,485,860
48	Post office services and telecommunications	7,206,901	3,488,016
49	Financial intermediation services, computer and related services	19,324,778	6,226,284
50	Real-estate services and house renting	1,943,976	808,499
51	Services to enterprises, except financial and real-estate services	11,501,014	2,072,623
52	Domestic services	-	-
53	No market education services	605,174	574,856
54	Social service and market health services	3,970,503	1,222,560
55	Leisure services and other market services	1,623,293	1,103,878
56	Government administration services and other services for the community	-	9,076,643
57	Market education services	-	915,259
58	Social service and no market health services	-	5,524,693
59	Leisure services and other no market services	-	493,315
60	Financial intermediation services indirectly measured	-	12,381,296

Source: own elaboration based on data from DANE.

According to Duque and Rey (2008), before applying the NBIC algorithm it is necessary to assess whether or not a network reduction can lead to a considerable loss of information. In this context, Table 2 describes the distribution of transactions across deciles of links. As can be seen, 10% of the I-O links account for 76.11% of all the inter-industry transactions in Colombia's economy<sup>10</sup>. This level of concentration in inter-industry transactions is a first step to guarantee that the network reduction will not be too damaging for the analysis. The second step in assessing the impact of this reduction is to calculate the relative difference between opposite flows. The results in Table 3 show that in 74.18% of the pairwise industry relationships, the difference between opposite flows is greater than 90%<sup>11</sup>. This difference also guarantees that the reduction described in Equation 1 is valid for the Colombian economy.

Table 2  
Analysis by Deciles of Transaction Volume

Decile	Transaction Volume
10th	76.11%
20th	87.93%
30th	93.68%
40th	96.71%
50th	98.37%
60th	99.27%
70th	99.72%
80th	99.93%
90th	99.99%
100th	100.00%

Source: own elaboration.

## B. EVALUATION MEASURES

In this section we report on the results of applying the NBIC algorithm using the 2005p I-O table for Colombia. Following the procedure for determining the optimal number of clusters, the industries can be aggregated into 12 clusters. At this level of

<sup>10</sup> These results are similar to the ones obtained by Duque and Rey for San Diego's economy.

<sup>11</sup> Duque and Rey report 72.2% of relative difference between opposite flows for San Diego's economy.

aggregation the proportion of weak clusters do not exceed 50% for either the classification based on internal linkages nor the classification based on external linkages. Table 4 presents several characteristics of the twelve industry clusters.

Table 3  
Relative Difference Between Pairwise Opposite Flows

Relative difference	% of transaction
0-10%	2.11%
10-20%	2.26%
20-30%	2.50%
30-40%	2.03%
40-50%	1.79%
50-60%	2.89%
60-70%	3.59%
70-80%	3.74%
80-90%	4.91%
90-100%	74.18%

Source: own elaboration.

Table 4  
Summary of the Clusters

Name	Core*	Size	Total Output (million COP)	Output Share
Construction	35	13	106,380,839.00	18.20
Petrochemical	30	6	75,244,833.00	12.87
Transport	44	6	61,860,666.00	10.58
Food	3	7	56,094,322.00	9.59
Wholesale and retail	41	5	51,931,533.00	8.88
Educational Services	56	3	47,953,590.00	8.20
Restaurants	43	6	39,058,335.00	6.68
Energy for industries	29	3	36,727,719.00	6.28
Banking	49	2	35,045,780.00	5.99
Farming	2	3	29,587,632.00	5.06
Energy for residencial and commercial use	10	2	26,369,770.00	4.51
Public utilities	51	2	18,390,011.00	3.15
Total		58	584,645,030.00	100.00

\* Code of the core industry in each cluster.  
Source: own elaboration.

The size of the clusters, measured as the number of industry members, varies from two to 13 industries. *Construction* is the largest cluster in the economy with 13 industries representing a 22.41% of the industries included in this study. *Energy for residential and commercial use*, *Banking*, and *Public Utilities* are the smallest clusters in the economy. Each of these clusters contain two industry members.

The total output (production) of a given industry corresponds to the sum of all inter-industry sales and its sales to final demand, which includes purchases by consumers, government, and sales to other activities of investment goods. The total output of a cluster is then the sum of the total output of each industry assigned to the cluster; and the total output share is estimated as the ratio between the total output of each cluster and the total output generated by all the industries in the economy. The average total output share per cluster is 8.33%, with values ranging from 3.15% to 18.20%. The largest cluster in terms of output share is *Construction*, producing over 106,380,839 million COP, representing 18.20% of the total output. The second largest cluster is *Petrochemical* producing over 75,244,833 million (12.87% of the total output).

	Backward Linkages	Forward Linkages	Type	Purchase Share Coefficient	Sales Share Coefficient	Type
	1.27	1.06	Key	0.45	0.37	Strong
	1.12	1.10	Key	0.41	0.31	Purchase
	1.23	0.99	Driver	0.27	0.35	Sales
	1.30	1.02	Key	0.52	0.65	Strong
	1.11	0.85	Driver	0.19	0.33	Weak
	1.31	0.66	Driver	0.03	0.47	Sales
	1.09	0.77	Driver	0.21	0.49	Sales
	1.03	1.17	Key	0.34	0.23	Purchase
	1.17	1.08	Key	0.67	0.64	Strong
	0.89	0.92	Weak	0.12	0.08	Weak
	1.20	1.41	Key	0.02	0.02	Weak
	0.80	1.36	Enabler	0.11	0.03	Weak
				0.28	0.33	←average

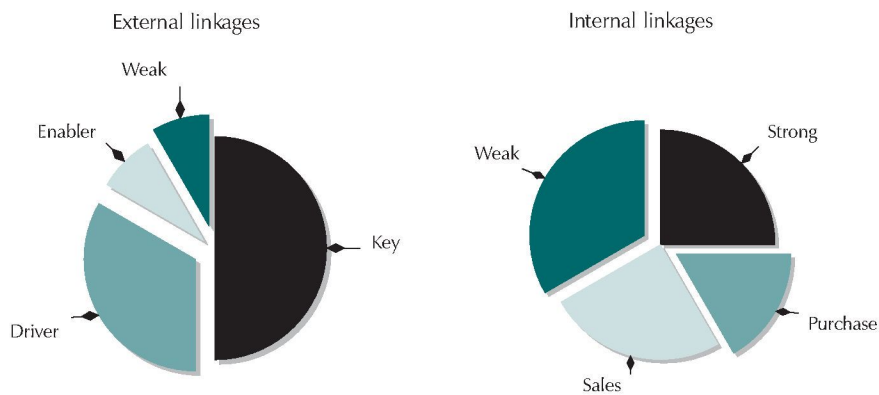


Graph 2 shows the distribution of the 12 clusters based on internal and external linkages. The classification based on internal linkages (right figure) reflects the importance of the cluster as an internal sales and/or purchase market for the industries assigned to the same cluster (industry members). Internal linkages can also be understood as a measure of interactivity between the industries within the cluster.

In Colombia, the average purchase share ( $\overline{PS}$ ) and the average sales share ( $\overline{SS}$ ) are 0.28 and 0.33 respectively. This means that, on average, the clusters purchased 28% of their intermediate products from industries within the cluster, and 33% of the intermediate sales occur within the clusters. According to the results 33.33% of the clusters are important for industry members as both sales and purchase market (strong clusters); 25.00% are an important sale market (sales oriented clusters); 16.67% are an important purchase market for the industry members (purchased oriented clusters); and 33.33% of the clusters are classified as weak clusters.

The figure on the left (Graph 2) presents the distribution of the clusters according to the external linkages classification. External linkages seek to measure how important each cluster is to the larger economy. On one hand, clusters with backward linkages greater than one ( $BL_c > 1$ ), indicate that the cluster creates an above average increase in activity for the regional economy when the cluster experiences a marginal increase in its final demand.

Graph 2  
 Distribution of the Clusters Based on Internal and External Linkages



Source: own elaboration.

These clusters can be viewed as output drivers for the regional economy due to their reliance on locally produced inputs. On the other hand, a cluster with a forward linkage coefficient greater than one ( $FL_c > 1$ ) shows an above average increase to marginal increase in other industries' final demand. This is indicative of the sector playing a strategic enabling role as a core supplier of inputs to other industries (Rey and Mattheis, 2000).

In Colombia, 50.00% of the clusters were classified as “key” clusters ( $BL_c > 1$  and  $FL_c > 1$ ), 8.33% as “enabler” clusters ( $BL_c < 1$  and  $FL_c > 1$ ), 33.33% as “driver” clusters ( $BL_c > 1$  and  $FL_c < 1$ ), and 8.33% as “weak” clusters ( $BL_c < 1$  and  $FL_c < 1$ ).

### C. COMPOSITION OF THE CLUSTERS AND IMPACT ANALYSIS

This section provides an in-depth analysis of each cluster. This analysis includes information about the role of each cluster within the larger economy, as well as the internal composition and dynamics of each cluster.

Graph 3 shows a graphical representation of each cluster. This representation is a useful tool to get an initial idea about the dynamics within each cluster. For example, through these graphics it is possible to recognize the role of each core industry (represented as squared shaped nodes). Thus, there are some clusters where the core industry is an important supplier for the other industry members (e.g. *Farming, Petrochemical*). In some other clusters, the core industry stands out as an important customer for the industry members (e.g. *Restaurants, Transport, and Wholesale and retail*). And finally, there are clusters where the core industry is a “connector” that purchases from some industry members and sells its intermediate products to other industry members (e.g. *Construction, Food, and Energy for industry*). According to Duque and Rey (2008) very few algorithms for industry cluster identification provide a clear way to understand the relationships between the industries assigned to the same cluster.

Table 5 describes the internal composition of each cluster and provides information to assess how important the cluster is for its industry members (Output and Input) and for the larger economy (Multiplier). This assessment is carried out with three measures:

- **Output:** Percentage of the industry output that goes to the other industries in the cluster. This value is a ratio between the industry output flowing

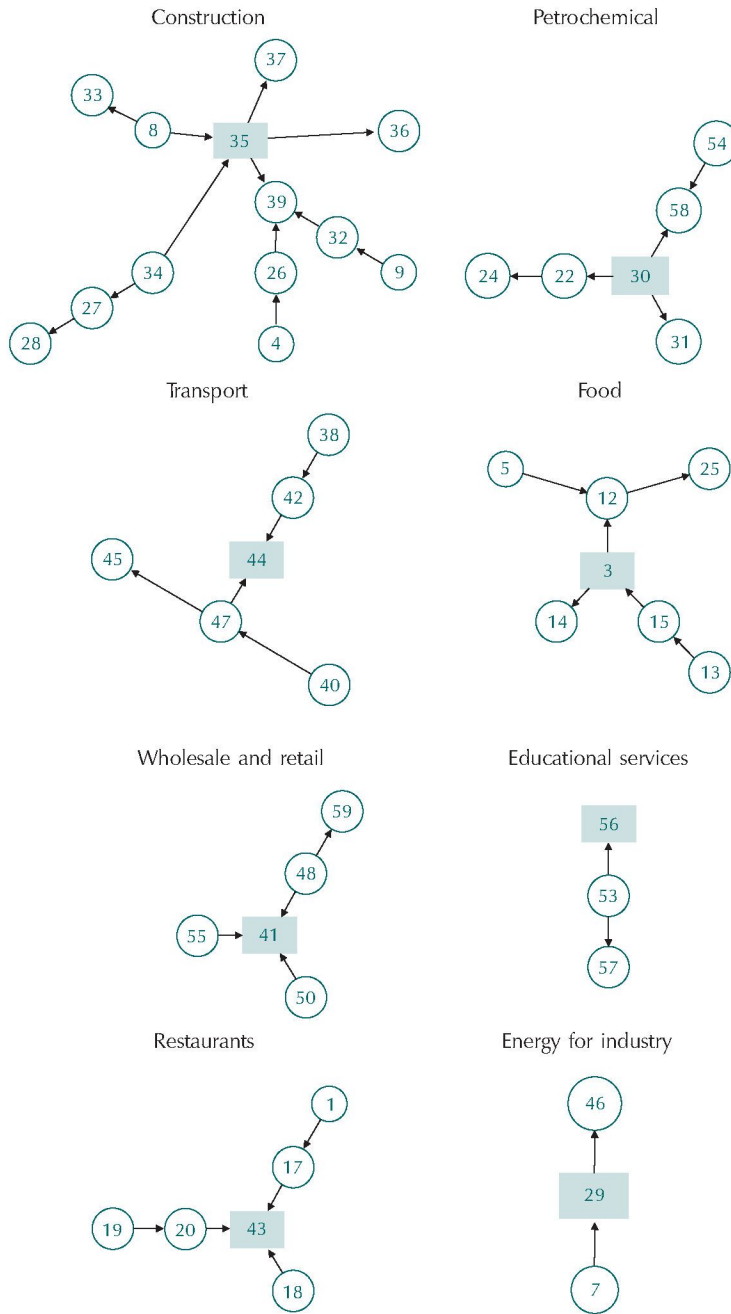
within the cluster, and the total output of the industry. On average, 20% of the industries' output flows within the clusters. This value varies from cluster to cluster, and it strongly depends on the main activity of the cluster. Thus, there are clusters, like *farming* and *public utilities*, whose products and services are purchased by a wide range of industries, leading in low values of this index, 1.54% and 3.68% respectively. On the other hand, clusters like *food* and *banking* have an important portion of their sales flowing within the cluster, 31.7% and 30.84% respectively.

- **Input:** Percentage of the industry purchases that come from other industries in the cluster. This value is a ratio between the industry's input flowing within the cluster, and the total input of the industry. On average, 17.13% of the industries' purchases are supplied by other industries within the clusters. The clusters *farming* and *public utilities* present the smallest values of this index, 7.64% and 6.81% respectively; and *banking* and *food* have the highest values, 53.53% and 27.76% respectively.
- **Multiplier:** The multipliers estimated in this study are known in the literature as simple output multipliers, which seek to measure the change in the gross output of the local economy when there is a COP's worth change of final demand for a given industry, cluster or local economy (Leontief, 1953).

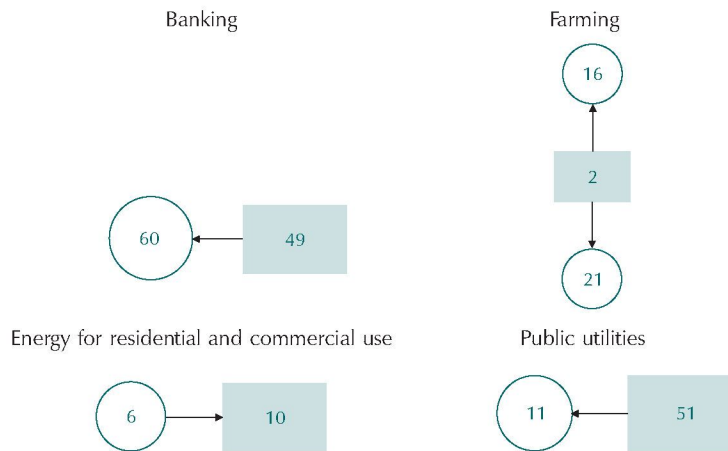
An output multiplier for a given industry  $i$  is defined as the total value of production in all the industries of the local economy that is necessary in order to satisfy an additional dollar worth of final demand for industry  $i$ 's output. The average output multiplier for Colombia is 1.55. Thus, a change of one COP's worth of final demand for Colombia will generate a change of 1.55 COP in the gross output of Colombia. At industry level *construction* is the industry with the highest output multiplier (3.61) and *banking* is the cluster with the highest output multiplier (1.84).

Two important results are derived from the figures in the Table 5: i) seven out of 12 industries selected as "core industries", using the NBIC algorithm, have output multipliers ranked in the top ten; ii) only two clusters contain more than one industry ranked in the top ten highest output multipliers. This result may suggest an easier way to define core industries within the NBIC algorithm.

Graph 3  
Graphical Representation of Each Cluster



Graph 3 (continued)  
Graphical Representation of Each Cluster



Source: own elaboration.

## 1. Construction

*Construction* is the biggest cluster in Colombia with 13 industries that account for a 18.20% of the gross output of Colombia. This cluster includes four complementary activities (branches) related to the construction industry: industries 9 and 32 representing the branch of glass; industries 26 and 4 representing the branch of wood; industries 8, 33 and 35 representing the branch of iron and steel; industries 27 and 28 representing the branch of recycled products.

According to the backward and forward linkages, this cluster is classified as a “key” cluster, meaning that it plays an important role as both driver and enabler of the economy. It also represents a wide market of resources and clients: 45% of the purchases of the cluster come from its industry members, and 37% of the sales of the cluster are directed toward industries in the cluster. At industry level, four industry members sell more than 50% of their output to industries within the cluster. It is also important to note the key role played by the industry *Construction* which purchases 44.47% of its input from other industry members, and reports the highest output multiplier in the Colombian economy (3.61).

## 2. Petrochemical

This is a medium sized cluster with six industries belonging to it. However, this cluster contributes a significant share of the Colombia's gross output, 12.87%, ranking as the second highest cluster in output contribution.

The core industry *Basic and elaborated chemical products* plays an important role in the cluster. This industry supplies chemical products for three different usages: i) textile and clothing; ii) plastic and rubber products; and iii) health services related products. From the arrange of the industries within the cluster and their relationships, it is easy to predict that this is a purchase oriented cluster. In fact, 41% of the purchases of the industry members remain within the cluster.

## 3. Transport

Six industries compose the third biggest cluster in term of output share, representing 10.58% of Colombia's gross output. This cluster covers road and fluvial transport and other related services such as transport equipment, road transport and repair services, and civil engineering works.

The transport cluster is classified as a driver cluster of the economy. With backward linkages greater than one (1.23), this cluster creates an above average increase in activity for the regional economy when the cluster experiences a unit increase in its final demand. It is also an important sale market for its industry members. In fact, industries such as *Complementary road transport services*, *Repair services of engines and domestic stuff*, and *Transport equipment* have an important portion of their sales in the transport cluster.

The transport cluster is also one of the two clusters that contain two industries with output multipliers ranked within the top ten highest multipliers, they are *Road transport services* (2.65), and *Civil engineering works* (1.97).

## 4. Food

The seven industries belonging to this cluster represent the production of food from livestock and its derivatives. This cluster produces 9.59% of the total output of the Colombian economy and it is considered to be a key cluster with forward and backward linkages greater than one. The industries within the cluster are strongly related

to each other, with purchase and sale shares above the average, 0.52 and 0.65 respectively. The cluster presents the second highest multiplier in the economy with 1.69.

This cluster is the main market for the core industry *Animals and products derived from animals*, representing 80% of its intermediate sales. Moreover, the cluster is the main supplier of the industry *Fish and meat* providing 62.61% of its inputs, mainly coming from *Fish and other products from fishing* and *Animals and products derived from animals*. *Fish and meat*, has also the highest industry multiplier (2.72) of this cluster.

## 5. Wholesale and Retail

*Wholesale and retail* is a medium sized cluster in term of industries. The five industries of this cluster represent 8.88% of Colombia's gross output. This cluster is classified as a driver cluster in terms of its external linkages, meaning that its reliance on locally produced inputs creates an above average increase in activity for the regional economy when the cluster experiences a unit increase in its final demand.

Although the cluster does not have a particularly high output multiplier, its core industry *Wholesale and retail* has the second highest output multiplier in Colombia (3.10).

This cluster is an important source of inputs for the industries *Wholesale and retail* and *Leisure services and other no market services*, which obtain 23.35% and 20.97% of their inputs from other industry members.

## 6. Educational Services

The educational services industry includes a variety of institutions that offer vocational, career or technical instruction, and other educational and training services to millions of students each year. Three industries compose this cluster. They produce 8.20% of the total output of the economy. The multiplier for the cluster is 1.58, slightly higher than the Colombian one, however the core industry, *Government administration services and other services for the community*, has a relatively high multiplier ranking among the top five highest multipliers.

*Educational services* is one of the four driver clusters in the economy. It has the highest backward linkage 1.31. This is mainly due to the fact that its core industry is highly linked to the government sector, and government expenses are a great engine for the economy.

## 7. Restaurants

Three branches converge to the core industry *Hotels and restaurants services*: i) coffee production; ii) cacao products; and iii) beverages and other food products. Its core industry has the sixth highest output multiplier (2.29).

This cluster accounts for 6.68% of Colombia's gross output, and it is classified as a driver cluster in terms of its external linkages, which indicates that the backward linkages reflect purchases of intermediate goods and services by this cluster that are necessary to meet the demand. Regarding the internal linkages, the cluster is classified as sales oriented, suggesting that the cluster is an important sale market for its industry members.

*Unroasted coffee* is one of the main industries in the Colombian economy. In 2005, this industry exported 88.7% of its total production, having as its main customers the USA, Germany and Japan.

## 8. Energy for Industries

The three industries included in this cluster represent the production of energy for powering industries and air transport. It contributes 6.28% of Colombia's gross output. With forward and backward linkages greater than one, this cluster is classified as a key cluster for the Colombian economy.

The industries in this cluster purchase 34% of their intermediate products from other industries within the cluster. This figure places the cluster as one of the two purchase oriented clusters in this analysis.

When analysing the graphical representation in Graph 3 a natural relationship can be observed. The industry *Oil, natural gas, uranium and thorium minerals* extracts raw oil and sells 26.33% of its intermediate products to the other industry members. Next, the industry *Refined oil products, nuclear combustibles and coke furnace products* provides 15.52% of the inputs required by *Aerial transport services*.

## 9. Banking

Although the banking cluster does not have important output share in the economy, 5.99% of gross output, it is classified as a key/strong cluster in terms of its external/



internal linkages. The two industries belonging to this cluster are responsible for moving the monetary resources through all of the economic sectors. The relevance of this cluster in the economy is also reflected in its output multiplier, 1.84, the highest multiplier at cluster level.

## 10. Farming

*Farming* is composed of three industries of the agricultural sector: *Other farming products*, *Sugar* and *Tobacco products*. This cluster generates 5.06% of the gross output of the Colombian economy, and it is classified as a weak cluster in terms of both, internal and external linkages.

The figures in Table 5 suggest that this cluster is not an important market or source of inputs for its industry members. This may be consequence of the export orientation of industries like *Sugar* and *Tobacco*.

## 11. Energy for Residential and Commercial Use

The cluster has two industry members, *Lignite and peat* and *Electricity and city gas*. This small cluster represents the production of energy for the households. Although the contribution of this cluster to the Colombian gross output is not outstanding (4.51%), this cluster is classified as a key cluster in terms of its external linkages.

According to the internal linkages, it is classified as a weak cluster. This is mainly due to the fact that the *Lignite and peat* industry exports almost the totality of its production to Europe and the USA.

## 12. Public Utilities

This cluster is composed of two service industries. Their contribution to the economy's gross output is 3.15%, the lowest among the clusters.

As expected, this cluster is an enabler cluster in the economy (the only enabler cluster). A cluster with a forward linkages coefficient greater than one is considered to have an above average sensitivity to unit changes in all sectors' final demands. This is indicative of the sector playing a strategic enabling role as a core supplier of inputs to other industries. Given the enabler role of this cluster, it is not surprising that its internal linkages are not very important for its industry members.

Table 6 shows the clusters and core industries obtained in this paper with the results found by (Monitor Company, Inc, 1993) and Hausmann and Klinger (2007). Since these three studies utilized different methodologies, input data, aggregation level and concepts of clusters, it is difficult to perform an objective comparison of the results.

Table 5  
Industry Clusters in the Colombian Economy

Code	Sector	Output	Input	Multiplier
<b>Construction</b>				<b>1.54</b>
35	Common metals and metallic products, except machines and equipment	53.19	23.64	2.04*
9	Other non-metallic minerals	68.85	1.73	1.03
34	Waste products	46.23	0.00	1.00
8	Metallic minerals	3.75	4.93	1.06
27	Cardboard and paper	26.41	22.51	1.73
36	Special and general use machines	7.48	9.86	1.23
4	Forestry products and wood extraction	45.48	1.49	1.03
26	Wood products, cork, straw and plait materials	63.42	33.38	1.41
32	Glass, glass products and other non-metallic products	50.28	21.40	1.76
39	Construction	3.29	44.47	3.61*
37	Other machines and electrical supplies	7.84	5.71	1.22
28	Printing and similar goods	9.80	22.85	1.30
33	Furniture and other transportable goods	2.16	13.45	1.58
<b>Petrochemical</b>				<b>1.56</b>
30	Basic and elaborated chemical products (except plastic and rubber products)	30.72	20.10	2.02*
31	Plastic and rubber products	12.44	21.76	1.36
22	Threads and textile fibre weaved	41.11	18.72	1.27
24	Clothes	0.65	22.03	1.65
54	Social service and market health services	61.98	8.60	1.18
58	Social service and no market health services	0.00	40.53	1.89
<b>Transport</b>				<b>1.62</b>
44	Road transport services	1.75	23.07	2.65*
38	Transport equipment	25.26	10.46	1.53
42	Repair services of engines and electrodomesticals	29.20	32.44	1.29
40	Civil engineering works	4.96	1.46	1.97*
47	Complementary road transport services	41.04	14.52	1.20
45	Fluvial transport services	0.60	43.03	1.10

Table 5 (continued)  
Industry Clusters in the Colombian Economy

Code	Sector	Output	Input	Multiplier
<b>Food</b>				<b>1.69</b>
3	Animals and products derived from animals	80.01	27.61	1.94
13	Animal and vegetable oil, fur and cakes	34.20	18.24	1.40
15	Mill products, starch and its products	30.22	17.56	1.96
5	Fish and other products from fishing	51.80	3.71	1.03
12	Fish and meat	6.15	62.61	2.72*
14	Dairy products	3.14	31.66	1.46
25	Leather and shoes	16.71	32.95	1.33
<b>Wholesale and retail</b>				<b>1.59</b>
41	Wholesale and retail	0.00	23.35	3.10*
50	Real-estate services and house renting	2.96	0.27	1.11
48	Post office services and telecommunications	18.98	1.66	1.50
59	Leisure services and other no market services	0.00	20.97	1.09
55	Leisure services and other market services	14.50	4.34	1.16
<b>Educational services</b>				<b>1.58</b>
56	Government administration services and other services for the community	0.00	0.65	2.51*
53	No market education services	3.90	0.02	1.09
57	Market education services	0.00	0.89	1.15
<b>Restaurants</b>				<b>1.43</b>
43	Hotel and restaurant services	0.49	15.94	2.29*
19	Other food products	10.16	5.59	1.36
18	Cacao, chocolate and sugar products	4.79	1.79	1.19
20	Beverages	15.61	4.38	1.50
1	Unroasted coffee	12.52	0.35	1.10
17	Processed coffee	20.42	38.29	1.14
<b>Energy for industries</b>				<b>1.61</b>
29	Refined oil products, nuclear combustibles and coke furnace products	7.41	23.38	1.76
7	Oil, natural gas, uranium and thorium minerals	26.33	1.90	1.46
46	Aerial transport services	1.88	15.52	1.38
<b>Banking</b>				<b>1.84</b>
49	Financial intermediation services, computer and related services	61.69	7.06	1.74
60	Financial intermediation services indirectly measured	0.00	100.00	1.95
<b>Farming</b>				<b>1.19</b>
2	Other farming products	3.71	1.19	1.43
16	Sugar	0.50	19.19	1.13
21	Tobacco products	0.43	2.54	1.03

Table 5 (continued)  
Industry Clusters in the Colombian Economy

Code	Sector	Output	Input	Multiplier
<b>Energy for residential and commercial use</b>				<b>1.62</b>
10	Electricity and city gas	32.12	32.25	2.06*
6	Lignite and peat	2.32	1.91	1.17
<b>Public utilities</b>				<b>1.19</b>
51	Services to enterprises, except financial and real-estate services	5.34	2.97	1.26
11	Water and sewerage services	2.02	10.64	1.12

\* Belongs to the top ten industries with higher output multiplier.  
Source: own elaboration.

Table 6  
Results from Previous Studies, and NBIC Results

<b>Monitor Company, Inc. (1993)</b>	<b>NBIC</b>
Clusters	Clusters
Capital goods	Construction
Petrochemical	Petrochemical
Tourism	Transport
Fruit juices	Food
Flowers	Farming
Textiles	Wholesale and retail
Graphic arts	Educational services
Leather	Restaurants
	Energy for industries
	Banking
	Energy for residential and commercial use
	Public utilities
<b>Hausmann and Klinger (2007)</b>	<b>NBIC</b>
Industries	Core industries
Fabricated metal products	Common metals and metallic products, except machines and equipment
Medicaments	Basic and elaborated chemical products (except plastic and rubber products)
Motor vehicles	Road transport services
Shipbuilding and repairing	Animals and products derived from animals
Yarn, wool and fabrics	Wholesale and retail
Electrical appliances and housewares	Electricity and city gas
	Government administration services and other services for the community
	Hotel and restaurant services

Table 6 (continued)  
 Results from Previous Studies, and NBIC Results

Hausmann and Klinger (2007)	NBIC
Industries	Core industries
	Refined oil products, nuclear combustibles and coke furnace products
	Financial intermediation services, computer and related services
	Other farming products
	Services to enterprises, except financial and real-estate services

Source: own elaboration.

#### IV. CONCLUSIONS

This paper presented an empirical application of a recent methodology proposed by Duque and Rey (2008) to identify industry clusters based on network analysis. Such application utilizes the projected 2005 Colombia’s product utilization matrix.

This algorithm shows up as a novel methodology which conciliates Porter’s approach, who emphasizes the importance of creating industry clusters to enhance country development, and Hausmann’s research which offers guidance to policy makers in the identification and support of key industries. On one hand, the NBIC algorithm sorts the industries by their level of importance for the economy identifying the core industries; on the other hand, it sets the most representative vertical clusters in a given economy. Thus, the methodology allows policy makers design specific policy initiatives for both clusters and key industries.

The NBIC algorithm identified 12 industry clusters in Colombia’s economy. According to the cluster’s internal linkages, 66.67% of the clusters were classified as either purchase oriented, sales oriented or strong clusters. In addition, based on the external linkages analysis, 91.67% of the clusters were classified as either key, enabler or driver clusters.

Among the outstanding clusters are: the clusters of *Construction* (the largest cluster in terms of industry members), *Petrochemical* and *Transport* represent together a 41.65% of 2005 Colombia’s total inter-industry output; the clusters of *Banking* and *Food* are the largest clusters in terms of output multiplier, with 1.84 and 1.69 respectively.

The possibility of visualizing the way the industries are related to each other within the cluster is one of the main strengths of the NBIC algorithm. This visualization tool has been greatly appreciated by policy makers, since it offers useful information about the structure of the supply chain; it allows easy identification of the role that each industry plays in the cluster; and speeds up the interpretation process as well.

An interesting finding is the strong relationship between being a core industry and having a high simple output multiplier. This will be matter of further research since it can be useful for simplifying the process of core industries identification, which is the most complex step in the NBIC algorithm.

Among the weaknesses of this algorithm stands out. Firstly, its sensitivity to the Modifiable Area Unit Problem (Openshaw, 1984), since it only takes into account inter-industry transactions within a predefined geographical area, excluding those flows that cross its boundaries. Thus, based on previous analysis performed in California Rey *et al.* (2006, 2007), it is expected that when the analysis is carried out on a smaller geographical scale, the number of industries and the structure of the clusters variate. This situation does not mean that performing the analysis at a subnational scales invalidates the results, but that the results depend on the selected area of analysis. Secondly, the NBIC does not account for those flows toward final demand (consumption, investment, government and exports), which in some industries represent an important share of their outputs.

The possibility of including information from more than one period; relaxing the assumption of exclusivity, where an industry is forced to belong to one and only one cluster (Dridi and Hewings, 2003, 2002); and the inclusion of flows toward final demand, seem to be very fruitful areas for future research.

## REFERENCES

1. Bergman, E.; Feser, E. *Industrial and Regional Clusters: Concepts and Comparative Applications*, The Web Book of Regional Science, 1999.
2. Connectory. Connectory home page. 2008. <http://www.connectory.com> (accessed December, 2008).
3. Czamanski, S. "Some Empirical Evidence of the Strengths of Linkages Between Groups of Related Industries in Urban-regional Complexes", *Papers, Regional Science Association*, num. 27, pp. 137-150, 1971.
4. Czamanski, S.; Ablas, L. "Identification of Industrial Clusters and Complexes: A Comparison of Methods and Findings", *Urban Studies*, num. 16, pp. 61-80, 1979.
5. Doeringer, P. B.; Terkla, D. G. "Business Strategy and Cross-industry Clusters", *Economic Development Quarterly*, num. 9, pp. 225-237, 1995.
6. Dridi, C.; Hewings, J. "Sectors Associations and Similarities in Input-output Systems: An Application of Dual Scaling and Fuzzy Logic to Canada and the United States", *Annals of Regional Science*, num. 37, pp. 629-656, 2003.
7. Dridi, C.; Hewings, J. "Toward a Quantitative Analysis of Industrial Clusters I: Fuzzy Clusters vs. Crisp Clusters", *REAL working paper 02-t-1*, 2002.
8. Duque, J. C.; Rey, S. J. "Network Based Approach Towards Industry Clustering", in *The Economics of Regional Clusters: Networks, Technology and Policy*, Edwar Elgar Publishing, Northampton, 2008.
9. Europe INNOVA. *Europe INNOVA. Annual report 2007*. Technical report, European Commission, <http://www.europe-innova.org/servlet/Doc?cid=10392&lg=EN> (accessed October, 2008), 2007.
10. Hausmann, R.; Klinger, B. *Achieving Export-Led Growth in Colombia*. Technical report, Quantum Advisory Group, 2007.
11. Hewings, G. J. D.; Jensen, R. C. "Regional, Interregional and Multiregional Input-Output Analysis", in P. Nijkamp (ed.), *Handbook of Regional and Urban Economics: Volume 1 Regional Economics* (pp. 295-355), Amsterdam, North-Holland, 1986.
12. High level advisory group on clusters. *The European Cluster Memorandum. Promoting European Innovation Through Clusters: An Agenda for Policy Actions*. Technical report, Center for strategy and competitiveness, Stockholm School of Economics., <http://www.europe-innova.org/servlet/Doc?cid=8857&lg=EN> (accessed October, 2008), 2008.
13. *Industrial clusters in San Diego County*. Technical report, Regional Analysis Laboratory, San Diego State University, 2007.
14. Isard, W.; Azis, I. J.; Drennan, M. P.; Miller, R. E.; Saltzman, S.; Thorbecke, E. *Methods of Interregional and Regional Analysis*, Ashgate, Brookfield, 1998.
15. Kaufman, A.; Gittlell, R.; Merenda, M.; Naumesand, W.; Wood, C. "Porter's Model for Geographic Competitive Advantage: The Case of New Hampshire", *Economic Development Quarterly*, num. 8, pp. 43-66, 1994.
16. Kruskal, J. B. "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem", *Proceedings of the American Mathematical Society*, vol. 7, num. 1, pp. 48-50, 1956.
17. Lahr, M. L. "A Review of the Literature Supporting the Hybrid Approach to Constructing Regional Input-Output Models", *Economic Systems Research*, num. 5, pp. 277-293, 1993.
18. Leontief, W. W. "Interregional theory", in *Studies in the Structure of the American Economy*, Cambridge, M. A., Harvard University Press, 1953.

19. Marshall, A. *Principles of Economics*, Prometheus Books, London, 1890.
20. Miller, R. E.; Blair, P. D. *Input-Output Analysis: Foundations and Extensions*, Englewood Cliffs edition, Prentice-Hall, 1985.
21. Monitor Company, Inc. *Creating the Competitive Advantage of Colombia*. Technical report, Ministry of Development, Instituto de Fomento Industrial, Confecámaras and Bancoldex, 1993.
22. Openshaw, S. *The Modifiable Areal Unit Problem*. Norwich, Geo Books, 1984.
23. Porter, M. *Ser competitivo: nuevas aportaciones y conclusiones*, Ediciones Deusto, Barcelona, 2003.
24. Porter, M. *The Competitive Advantage of Nations*, Free Press, New York, 1990.
25. Regal. Regal home page. <http://regionalanalysislab.org/?n=Connectory> (accessed December, 2008).
26. Rey, S. J.; Duque, J. C.; Smirnov, O.; Folch, D.; Schmidt, C.; Stephens, P. *Industry Clusters in California*. Technical report, Regional Analysis Laboratory, San Diego State University, 2006.
27. Rey, S. J.; Mattheis, D. J. *Identifying Regional Industrial Clusters in California: Volume I Conceptual Design*. Technical report, Californian Employment Development Department, Sacramento, 2000.
28. Roepke, H.; Adams, D.; Wiseman, R. "A New Approach to the Identification of Industrial Complexes Using Input-Output Data", *Journal of Regional Science*, num. 14, pp. 15-29, 1974.
29. Sandag. Sandag home page. <http://www.sandag.org> (accessed December, 2008).
30. Sistema Nacional de Competitividad. *Política Nacional de Competitividad y Productividad*, Technical report, 3527, Departamento Nacional de Planeación, 2008.
31. Steiner, M. *Clusters and Regional Specialization: On Geography, Technology, and Networks*, London, Pion, 1998.