Assessing the efficiency of science, technology and innovation using Data Envelopment Analysis (DEA): The case of Colombia

Daniel Aristizábal-Torres a, Mónica Yuleni Castro-Peña b, Paola Andrea Echeverri-Gutiérrez c & Jesús David Valencia-Salazar c

a Centro de Investigaciones, Universidad Libre Seccional Pereira, Pereira, Colombia. daristizabal@unilibrepereira.edu.co
b Universidad Católica de Pereira, Pereira, Colombia. monica.castro@ucp.edu.co
c Facultad de Ingeniería, Universidad Libre Seccional Pereira, Pereira, Colombia. paecheverri@unilibrepereira.edu.co, jdvalencia@unilibrepereira.edu.co

Received: August 11th, 2016. Received in revised form: March 22th, 2017. Accepted: July 30th, 2017.

Abstract
This paper provides insights into how to approach the assessment of scientific production efficiency using Data Envelopment Analysis (DEA). The methods employed in this paper take into account both the CCR and Assurance Region (AR) models, which are based on science, technology and innovation (STI) indicators. The sample focuses on the departments in Colombia as well as the Capital District (C.D.). The research concludes that both models calculate the efficiency of the units of analysis. Furthermore, according to the country’s STI targets, AR is susceptible to hierarchies generated within the variables.

Keywords: assessment of scientific production; efficiency; science, technology and innovation indicators; Data Envelopment Analysis (DEA).

1. Introduction

Science and innovation indicators demonstrate the potential of the technical scientific information that is generated by a country [1]. Scientific journals are the fundamental channel of formal communication within the scientific community [2]. The number of patents applied for or granted and the number of papers published in journals that have a high impact factor are indicators that are used to measure the quality of research [3].

On the global scale, the Organization for Economic Co-operation and Development OECD has published handbooks to assess activities relating to science, technology and innovation, such as the Oslo Manual (1997), the Canberra Manual (1995), the Manual of Patents (1994), the Technology Balance of Payments (TBP) (1990) and the Manual of Frascati (1963) [4].

The Administrative Department of Science, Technology and Innovation (Colciencias) in Colombia frequently monitors research groups. Colciencias has implemented a pattern to classify research groups using the categories A1,
A, B, C, D and recognized (Rec), which follows the research group indicator. The coefficient is assigned according to the categorization of four scientific elements: development of innovative knowledge, technological development, social appropriation of the knowledge and training of human resources. Furthermore, the pattern classifies researchers into three categories (Senior, Associated and Junior) according to their academic degree and their academic production [5].

The report entitled ‘State of Science in Colombia’ was written in 2015 and detailed the production of research in the different administrative divisions (departments) according to the four product categories. The indicators take into account investments in science, the number of scholarships granted, research centers and research groups, recognized researchers and registered products [5].

Some scholars have encouraged researchers to measure scientific production, which includes: measuring university research [6], the influence of the size of the institution, and the number of researchers producing scientific research in the centers [7]; scientific production in biological and health sciences in twenty Brazilian universities [8]; measuring scientific production in the field of the psychology [9]; and the empirical study of scientific measurement [10].

These above areas of research evaluate the number of products developed by the units of analysis. Other scholars make measurements in terms of efficiency; however, this involves two serious challenges: First, the research needs to develop standard criteria to compare the data collected from reality. Second, some cases analyzed are designed with complex organizational structures with multiple inputs and outputs for endowments and products, which do not assess efficiency with suitable precision. Therefore, the researcher should allocate weighted values according to trade-offs and factor endowments [11]. The mathematical application Data Envelopment Analysis (DEA) is such an alternative solution.

Some research that has used the DEA model has measured efficiency in the following areas: research projects in two clinical centers [12], national innovation systems [13], innovation in each German region [14], international comparisons of efficiency among the national innovation systems [15], analysis of scientific and technological activity [16], technology and knowledge transfer in Spanish public research centers [17], research groups in Andalucía [1], innovation regional systems in 31 Chinese provinces [18], classification of research group in Colombia [19], R&D activities according to the number of patents submitted in the European Union [20] and research activity at Valladolid University [21].

Based on data from Colciencias, this paper details the advantages and features of DEA application in order to measure scientific activity in Colombia.

2. Theoretical approach


Science, Technology and Innovation (STI) are core elements that should be used to develop sustainable knowledge in societies. STI has the potential, nationally, to be a significant engine for social and economic development.
is applied and is taken as the basis for the Assurance Region model in order to adjust the efficiency measurement by means of weight restrictions assigned to products. This is more in line with the Colciencias’ evaluation of each one of the types of products. The final phase refers to the interpretation of the efficiency levels determined by the two models applied.

3.1. Selection of variables

The research calculated the global results of 32 departments in Colombia. The Capital District (Bogota) was also observed due to its significant level of research production: the result was that a total of 33 DMU were obtained. The “State of Science in Colombia” exhibits information that shows the dynamics and general lines of Colombia’s science and technology (S&T), which are the variables used to apply the DEA model to each of the items in the report. They are classified as being either input and output variables and have a set of values for each unit of analysis. This operation (shown in Fig. 1) assumes that both monetary investments and human capital of every DMU are the inputs used to develop research products.

Fig. 1 shows 15 input variables, which include categories and subcategories, 33 DMU and four output variables. The number of variables was reduced because of the operationalization of the DEA model. Otherwise, the number of efficient DMUs would have been high, which would have allowed them a higher number of possible combinations and been a grade closer to the efficiency border. Nevertheless, the reduction in variables does not generate substantial loss of information for the application.

First, three variables related to ‘categorization of researchers’ were united into one variable. The same process was then applied for the six ‘categorization of research groups’ variables, and it was taken into account that these variables have a hierarchical order that is represented in particular categories relating to effort and research relevancy. The following paragraph explains the criteria used to unify the variables.

3.1.1. Selection of variables

The weighting scale with the values proposed in Colciencias’ last scheme of measurement was applied to assess research and research groups. This scale is coherent with the relevancy and the level of contributions to the science described within the classification (c.p Table 1). Therefore, the value of these variables is a weighted sum. The two variables presented a co-relation of 0.994. Overall, ‘research groups’ was the representative variable, taking into account that researchers’ production contributes to the classification of the research groups.

Moreover, Colciencias also observes the investment in ASTI and R&D as independent categories in order to differentiate them. However, this research grouped them into one variable as they are complementary [26].

Table 1. Weighting value for category of research groups in Colombia

<table>
<thead>
<tr>
<th>Groups</th>
<th>Category</th>
<th>Weighting value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>A</td>
<td>10</td>
</tr>
<tr>
<td>A2</td>
<td>B</td>
<td>8</td>
</tr>
<tr>
<td>A3</td>
<td>C</td>
<td>6</td>
</tr>
<tr>
<td>A4</td>
<td>D</td>
<td>4</td>
</tr>
<tr>
<td>A5</td>
<td>Rec.</td>
<td>2</td>
</tr>
</tbody>
</table>

Source: The authors.

Scholarships awarded for Master’s and PhD programs were not included in the project as input variables because the products of this investment will become reflected in the selected output variables in the medium-term after the students have finished their education. The following are used as the input variables:

- Research centers and technological development that are recognized and strengthened by Colciencias, according to information from the period 2009-2013.
- Research groups that were recognized and classified by their particular field of knowledge and for their scientific and technological production (results based on call for applications 693 of 2014).
- Approved STI projects financed by Colciencias, the values of which were calculated at constant 2013 prices and correspond to the period 2010 – 2013.
- Investment in science, technology and innovation (ASTI) and research and development (R&D) activities as a percentage of the GDP. This investment relates to the year 2013.

Four types of products were taken as output variables: new knowledge (NK), technological development (TD), social appropriation of the knowledge (SAK) and human training (HT). This information allows researchers to identify the target of each DMU’s effort.

3.2. Application of the model

This research used two DEA models as a reference. The first is the Assurance Region (AR), which maintains a hierarchy within the output variables. In this research, this refers to the different types of products developed by the research in Colombia; the second is the CCR-O model. Unlike the previous model, CCR-O does not provide any type of restriction when allocating weighted values for both input variables and output variables.

The AR model uses the CCR-O model as a baseline, which results in a constant scale. Specifically, in the Colombian case, every type of product counts on a different significance level, which depends on the research groups in Colciencias’ scheme of assessment. The calculated weighted values can be found in Table 2.
According to the weighted values (Table 2), each variable has a restriction when qualifying the efficiency of the DMUs, which is related to the proportion of weight assigned by DEA. Both the best and the worst stages were calculated for each of the possible relations between the products in order to establish the limits of the restrictions. Those limits can be found at the top and the bottom of data. Table 3 includes the limits of restrictions between product relations.

The Assurance Region model allows us to identify the weighting values assigned to each of the output variables for the DMUs. Considering their restrictions, they maintain the levels of relevance that are assigned to the types of research products by Colciencias, as shown in Fig. 2.

The CCR-O model does not take into account any kind of restriction. CCR-O freely allocates the most convenient weighted values to each DMU, according to the levels of the input variables and output variables.

4. Main findings

The application of the CCR-O model and AR model identified the DMU, which appears to be both efficient and inefficient depending on the levels that are described for the input variables and output variables. Furthermore, the research indicates the group of reference for the inefficient units. In addition, the results allow us to find the variables which need to be improved in order to be efficient. Finally, the research looks at the weighting assigned to the variables.

Four DMU were eliminated from 33 departments when the models were applied (Guainía, Putumayo, Vaupés and Vichada) as they offered values equal to zero in all their output variables. They were, therefore, disqualified from the analysis. Both models were output-oriented and were solved by the Software DEA Solver because the aim of this research is to determine the efficiency of the departments in Colombia and the Capital District in terms of their scientific production. The frequency of efficiencies for the DMUs were obtained and can be seen in Fig. 3.

Fig. 3 offers the DMUs according to the different ranges of efficiency levels for each model. The CCR-O model identified a higher quantity of technically efficient DMUs than the AR model. This difference is explained due to the AR model being more rigorous with restrictions that limiting the weighting values allocation. Neither the CCR- O nor the AR models described were Pareto-efficient DMUs because at least one of their variables was assigned with a zero weighting value.

The CCR-O model developed a border of constant efficiency that is shaped by the DMUs, which appear to be efficient. This takes into account that the aim of the projections in the model is to lead the inefficient units to the border. Based on this, efficient DMUs are converted into references that do not appear in the same situation. Fig. 4 shows how many times a DMU was referenced in each of the models, and a ranking is generated within the efficient DMUs since a higher frequency as a reference, implies a better performance of that unit compared to other efficient units [1].

Overall, Boyacá stood out as a global leader as it was efficient in both models as it had the biggest indicator such as modality of inefficient units. Amazonas and Atlántico came second and third in the number of DMUs. Cesar, Córdoba, Capital District, Caldas and Magdalena, which has a ‘zero’ frequency in the AR model, stopped being efficient after they were included in the restrictions of hierarchical structuring of the level of importance that was awarded to the research products.
DEA shapes the groups of reference in order to propose improvements to the inefficient DMUs, according to the levels presented by them. The goal is to achieve efficiency through identifying the variables and quantities which must be modified.

The main improvement is proposed for output variables because the CCR model and AR model are output-oriented. Despite this, some DMUs need the input variables to be further modified in order to become efficient. In this case, the models coincided with the type of variables that need to be changed in either a major or less manner so that the analysis units become efficient. The following variables should be adjusted with a descending relevance: products of technological development, social appropriation, human training and products of new knowledge.

Despite the models being output-oriented, both proposed modifications for some of the input variables. Investment in ASTI and research centers developed major amounts of change. This indicates that some departments are not developing enough products that justify the resources used as there are DMUs which achieved major production with less resources for those variables.

5. Discussion and conclusions

There is no evidence in Latin-American scientific reports regarding assessing the efficiency of the scientific activity on a regional level. 59% of research used CCR and 40% of research used BCC from all the projects that used DEA. The remaining 1% of projects used Super Efficiency, Crossed Efficiency and a proper guideline. None of the cases used the Assurance Region model. Therefore, there were no restrictions considered with incidence for the share of the weighting values assigned to the variables by DEA.

More than 95% of the research analyzed used the human and financial resources as input variables. Moreover, 70% of the research used patents and publications (papers and books) as output variables. The share of efficiency for the units of analysis ranged between 20% and 66%. For this exercise with CCR, the proportion of efficient DMU was 38%, and when the AR was adjusted, the proportion was lowered to 21%. This share makes more rigorous estimation possible because it includes restrictions that refer to levels of importance regarding the output variables. In the both cases, there were no Pareto-efficient DMUs as, at least in one of the variables, a weighting value of zero was allocated. Therefore, they were technically efficient.

Broadly speaking, the both the CCR model and the AR model allow researchers to undertake a measurement of STI approach. The AR model is suitable for hierarchical structures developed in the variables by the particular entity that manages the STI in each of the countries.

Researchers are concerned about efficiently measuring scientific up-dates. The main patterns used to be able to evaluate scientific activity are the high human resource skills and the relation they have with proper, public and private funds. These patterns allow research centers to generate products such as new knowledge, technological development, social appropriation and human training.

Furthermore, there is an opportunity to carry out a comparative study in other countries using the AR model, which allows researchers to determine hierarchical structures within the output variables, according to the CTI targets in the country. A multi-centric model can also be applied to compare the DMU case study.

Aknowledgements

We thank the Universidad Libre Seccional Pereira for their support during the research process.

References


Facultad de Ingeniería de la Universidad de Antioquia, [Online]. 42, colombianos aplicando análisis envolvente de datos. Revista de la 

Tong, L. and Liping, C., Research on the evaluation of innovation 

dissertation. Universidad Rey Juan Carlos de Madrid, Madrid, 

Gao, W. and Yuan, L., Analysis of science and technology activity 

Crespi, G. and Geuna, A., An empirical study of scientific production:

Kotsemir, M., Measuring national innovation systems efficiency – A 

Broekel, T., Rogge, N. and Brenner, T., The innovation efficiency of 

Gao, M., La eficiencia de los centros públicos de investigación en el 

Castro-Peña, P.A. Echeverri-Gutiérrez, received his BSc. in Commercial Eng. in 2004, and her MSc. in Economic and Financial Management in 2010. She works in programs and projects relating to business management and entrepreneurship with emphasis on business competitiveness for the Universidad Libre Seccional Pereira. She is a researcher in the Faculty of Engineering, Universidad Libre, Seccional Pereira, Colombia. 

ORCID: 0000-0002-4605-2741

J.D. Valencia-Salazar received his BSc. in Commercial Eng. in 2003, his post-graduate diploma in education and human development in 2009, and his MBA in 2006. He is PhD Engineering student in the University Nacional de Colombia. He works in programs and projects relating to the science, technology and innovation with emphasis on knowledge management at the Universidad Libre Seccional Pereira. He is a researcher in the Faculty of Engineering, Universidad Libre, Seccional Pereira, Colombia. 

ORCID: 0000-0002-8170-1954

Universidad Nacional de Colombia 
SEDE MEDELLÍN
FACULTAD DE MINAS

Área Curricular de Ingeniería Administrativa e 
Ingeniería Industrial

Oferta de Posgrados

Especialización en Gestión Empresarial 
Especialización en Ingeniería Financiera 
Maestría en Ingeniería Administrativa 
Maestría en Ingeniería Industrial 
Doctorado en Ingeniería - Industria y Organizaciones

Mayor información:
E-mail: acia_med@unal.edu.co 
Teléfono: (57-4) 425 52 02