SOME RANKING METHODS IN THE DATA ENVELOPMENT ANALYSIS TO EVALUATE SPANISH AIRPORTS

Juan Carlos Martín♦ **and Concepción Román**
Department of Applied Economic Analysis. University of Las Palmas de Gran Canaria. Las Palmas G.C. 35017 Spain

Abstract

In this paper we will use the flexibility of Data Envelopment Analysis (DEA) methodology to analyze the relative performance of each individual Spanish airport, and to fully rank both efficient, as well as inefficient airports. Most airports have previously compared their efficiency according to the results of some partial productivity ratios. However, this approach does not provide a good understanding of the overall performance of the airports. In this paper, we will use four different approaches of super-efficiency DEA in order to fully rank the performance of Spanish airports. The first approach involves the evaluation of a cross-efficiency matrix, in which each airport is evaluated according to its weights and their rivals' weights. The second and third approaches rank the airports' performance through the exclusion of the airport being analyzed. And finally the fourth approach is based in the analysis of the virtual airport and the comparison of the rest of the airports with this "champion performer". We will compare each approach, concluding that no one methodology can be prescribed here as the ideal solution to the question of ranking.

Keywords: DEA; Airport performance; Super-efficiency; Ranking; Benchmarking
Topic Area: A3 Airports and Aviation

1. Introduction

During the last years, a conceptual change in relation to the provision of transport infrastructure has been common in most parts of the world. In the transport sector, some liberalization measures have been accompanied by a process of partial or total privatization of services and infrastructures. This current has brought to the air industry, besides its relative youngness, huge panoply of packages of the aforementioned measures. Airports, air traffic control facilities, and government airlines were increasingly being put in a more commercial orientation, and in many cases they have been partially or fully privatized.

The genesis of the reform was produced in 1978 when the United States deregulated its domestic market. Liberalization policies introduced in the air industry in the United States had an overall positive impact on the sector. Productive and allocative efficiency were improved, load factors raised and traffic grew substantially more than it would have done in the absence of the deregulation process (e.g. see Caves et al., 1987; Morrison and Winston, 1995;and Baltagi et al., 1995).

The process initiated in the United States had some demonstration effect for the rest of the world. Thus, soon after the United States deregulation, new and more liberal bilateral agreements between the United States and some European countries (UK, Netherlands, Belgium, Ireland, Germany) were signed. Domestic markets were deregulated to some

♦ Author for correspondence. e-mail: jcmartin@daea.ulpgc.es. This paper has been written with the financial support of the Research Program of the University of Las Palmas.

degree in Canada and Mexico during the 80s and the European Union took three steps on a gradual process to liberalize its air industry[1].

The liberalization processes in the distinct parts of the world have provoked that productivity measures, used to compare the operational efficiency, effectiveness and the relative competitive position of airports and airlines, had become very popular since these days. Airports' performance has been usually assessed based on some financial efficiency or operational efficiency throughout a big number of partial indicators. Doganis and Graham (1987) found that most airports only use partial financial indicators to study their performance.

The basic indicators that can be constructed are related to the following variables: inputs, outputs and outcomes (demand). These variables can be expressed as a physical quantity or monetary value. Usually, it is possible to construct an ample variety of performance ratios. Ratios of productive efficiency are related to links between outputs and inputs. Ratios between demand and inputs are usually referred as economical efficiency or social productivity. Finally, the ratios between outputs and demand are usually referred as service effectiveness. The basic unit of measure in airport studies is based around the definition of a work load unit (WLU): one passenger or 100 kg of freight serviced. This measure was adopted to provide a single measure of output for airport business. Some of the most common partial ratios studied are: total cost per WLU; operating cost per WLU; capital cost per WLU; labor cost per WLU; WLU per employee; WLU per unit asset value; total revenue per WLU and aeronautical revenue per WLU.

However, these partial performance ratios that compare one or more basic variables have evident shortcomings because they can only be used to obtain a first glance, and robust consequences based on this comparability cannot be usually extracted. For example, some financial measures can be misleading indicators, as a consequence of the relative market power that can exist. Monopolistic airports might be able to make substantial profits even if they were inefficient.

Efficiency has several dimensions, two of the most important that have been profusely quoted are economic efficiency and technological efficiency. Economic efficiency is achieved when an airport is using resources in such proportions that the total cost for some level of output is minimum. Technological efficiency means that the airport cannot obtain more level of output for some combination of inputs.

However, under the new commercial context for airports, many airports have developed performance measures beyond the WLU ratios. Ultimately, the facilities of airports have very different demands and rewards for managers, reflecting new opportunities in retailing activities. As a result of this, airport managers are paying more attention to non-aeronautical activities as an optimal strategy to maximize its revenues. For this reason, indicators such as commercial income per square meter of concession space, concession income per passenger, non-aeronautical income per passenger, among others, have emerged. Francis et al. (2001) study a survey of the worlds 200 busiest passenger airports, and provide an insight into the prevalence and perceived usefulness of an ample set of performance measures.

In the last years, research on airport performance has grown considerably due to the changes of the context in which airports operate that has been observed. There exist different methodologies that have been developed and applied. Excellent revisions on this topic exist (Francis et al., 2002; Humphreys and Francis, 2002; Oum et al., 2003). Some studies separate the activities of the multi-product nature of the airports, focusing on some

---

[1] For details of the various “Packages” of reforms, see Button et al. (1998); Vincent and Stasinopoulos (1990); Stasinopoulos (1992, 1993).

operations, e.g. Gillen and Lall (1997) and Pels et al. (2001) differentiate landside and airside operations, and evaluate the productive efficiency of both sides, respectively.

Some quantitative methodologies such as data envelopment analysis (DEA), total factor productivity (TFP) and stochastic frontiers (SF) have been applied to airports in order to measure their performance, using different inputs and outputs that are usually constrained for the lack of the data. DEA has become increasingly popular to study productivity analysis in diverse industries, and airport industry is not an exception to this rule. Parker (1997) applied DEA to study airports with one runway. Parker (1999) used DEA to study the relative performance of British Airport Authority, before and after privatization. Sarkis (2000) employed DEA to examine the productivity of US airports. Adler and Berechman (2001) used DEA to measure the quality of airports from air carriers' point of view, using principal component analysis to reduce the dimensionality of the space of inputs and outputs. Martin and Roman (2001) applied DEA to evaluate the performance of Spanish airports. Chin and Siong (2001) used DEA to compare the relative performance of Changi airport and four airports in the metropolitan area of New York. Abbott and Wu (2002) investigated the efficiency and productivity of 12 Australian airports using DEA. Bazargan and Vasigh (2003) studied the relationship between size and efficiency of some US airports using DEA. Pacheco and Fernandes (2003) used DEA to analyze the efficiency of Brazilian airports taking into consideration two different dimensions: financial and physical.

The primary objective of these studies is to find out the behavior of two different sets: those airports that are efficient and define the Pareto optimal frontier and those that are inefficient. All the studies employ DEA, a non-parametric multiple input-output efficiency technique that is able to handle this multidimensionality and eliminate the difficult task of weight estimation that exist in other multiple-criteria decision making (MCDM) methods, such as Analytical Hierarchic Process (AHP) or Total Factor Productivity (TFP). However, almost none of the studies have extended or adapted the DEA models in the field of ranking the activity of the airports[2].

This paper attempts to extend Bazargan and Vasigh approach using other theoretical models to rank the airports. In order to rank all the airports we will use four different methodologies. The first method involves the evaluation of a cross-efficiency matrix, in which each airport is evaluated according to its weights and their rivals' weights. The second and third methods rank the airports' performance through the exclusion of the airport being analyzed (super-efficiency). And finally the fourth approach is based in the analysis of the virtual airport and the comparison of the rest of the airports with this "champion performer" (Bazargan and Vasigh, 2003). We will compare each approach, trying to examine the main differences and similarities that exist.

The paper is organized as follows: section 2 explains the theoretical framework. Section 3 analyzes the organization of Spanish Airport System and presents a brief summary of the sample of the airports employed in the empirical exercise. Section 4 shows the basic results obtained and section 5 concludes.

2. Ranking methods and DEA

Charnes et al. (1978), in their seminal paper, described the DEA methodology as a "mathematical programming model applied to observed data that provides a new way of obtaining empirical estimates of extremal relationships such as the production functions and/or efficiency production possibility surfaces that are the cornerstones of modern

[2] To our knowledge there exists only one exception (Bazargan and Vasigh, 2003). In this paper, the authors achieve a full ranking of all airports considered in their sample, introducing a virtual super efficient airport with the rest of all airports. This method ensures that there is only one efficient airport with the rest being ineffICIEN.

economics". Since then, numerous applications employing the DEA methodology have been presented and involve a wide area of contexts: education, health care, banking, armed forces, sports, transportation, agriculture, retail stores and electricity suppliers. Originally designed to evaluate decision making units (DMUs), which use multiple inputs to produce multiple outputs, without a clear identification of the relation between them, DEA has progressed throughout a variety of formulations and uses to other kind of industries[3]. Seiford (1994) cited more than 400 articles in a comprehensive bibliography and stated that DEA methodology is an important analytical tool whose acceptance is no longer in doubt. Emorouznejad and Thanassolis (1997) provide a comprehensive list of more than 1500 applications of DEA.

We do not intend to cover the basic aspects of DEA models. A good introduction to DEA notation, formulation and geometric interpretation can be consulted in Charnes et al. (1994), Ali and Seiford (1993), Coelli et al. (1998) and Cooper et al. (2000). As discussed therein, a model can be described by the envelopment surface, orientation of the model, invariance of units, and efficiency measurement. There are three basic DEA models: variable returns to scale (VRS), constant returns to scale (CRS) and additive model. These can be used to seek which ones of the n DMUs determine the frontier of the envelopment surface, and are deemed efficient. The units that do not lie on the frontier are inefficient and the measurement of the grade of inefficiency is determined by the selection of the model.

The choice of a DEA model depends on some assumptions regarding the data set to be employed and in some prior results about the industry to be studied. The data set has to describe the activities of the units in the better possible way. It is especially important to have some idea about the hypothetical returns to scale that exist in the industry. This knowledge is going to determine the envelopment surface –constant return to scale CRS or variable return to scale VRS[4]) of the model.

Once that the selection of envelopment surface has been chosen, an orientation of the model to determine the measurement of the efficiency is needed. There are three basic orientations: input, output and equal. An input orientation focuses on proportional decrease of the input vector, the output orientation adjusts the proportional increase of the output vector and the equal orientation do not discriminate the importance or the possible increase of output or decrease of input. The units involved in the study determine the selection of the orientation. It is very important to have in mind what the real possibilities of the managers are, i.e. in Harvard tradition the investigator must try to establish what the conduct of agents and the structure of the market are in order to contemplate a possible orientation.

In DEA analysis, it is generally assumed that there are n production units to be evaluated, using amounts of m different inputs to produce quantities of s different outputs. Specifically, the *o'th* production unit consumes x_{io} units of input i $(i= 1 \text{ to } m)$ and produces y_{ro} units of output r $(r=1 \text{ to } s)$. The *o'th* production unit can now be described more compactly with the vector (X_o,Y_o), which denote, respectively, the vectors of input and output values for DMU_o.

Next, we consider the dominance comparisons for this production unit using the data set as a reference. DEA consider the dominance of the linear combinations of the n

---

[3] DEA can be applied to scenarios where the data cannot be strictly interpreted as inputs or outputs or there is no direct functional relationship between the variables. In such situations, a general guideline to the classification of the variables is that variables for which lower levels are better are considered inputs, while outputs are those variables for which higher amounts are better.

[4] CCR and BCC acronyms are sometimes used in reference to CRS and VRS models. The acronyms come from the initial of the authors of the papers that employed these two different envelopment surfaces (Charnes et al., 1978 and Banker et al., 1984).

production units, i.e. $(\sum_k \lambda_k X_k, \sum_k \lambda_k Y_k)$, with the scalar restricted to be non-negative[5]. The production unit *o* is dominated, in terms of inputs, if at least one linear combination of production units shows that some input can be decreased without worsening off the rest of inputs and outputs. The production unit *o* is dominated in terms of outputs if at least one linear combination of production units shows that some output can be increased without worsening off the rest of inputs and outputs[6].

Thus, the method serves to partition a set of production units into two subsets: the efficient production units and the inefficient ones. The method also serves to calculate the level of inefficiency of a given inefficient production unit.

Airport managers can affect the efficiency of their airport using their inputs, runways, terminal buildings, in different manners. In this paper, ranking methods are going to be based on output orientation. We think that once an airport has invested in the building of new runways or new facilities in the terminal buildings, it is difficult for the managers to initiate processes of disinvestment more close to an input orientation. In this sense, it is more credible to use the airport as much as it is demanded due to the fact that the factors of production usually do not change year to year. Of course, authors are well aware that some factors that affect airport performance are not directly under the control of airport manager, e.g. airline inefficiency (low load factors) appears to contribute significantly to airport inefficiency in terms of air passenger movements (Pels et al., 2003).

Formally, the multiplier-DEA VRS output efficiency for the unit *o* is calculated through the following linear programming problem:

$$\begin{aligned} &\min_{\nu\mu} \frac{\sum_{i=1}^{m} \nu_i x_{io} + \nu_o}{\sum_{r=1}^{s} \mu_r y_{ro}} \\ &s.t. \\ &\frac{\sum_{i=1}^{m} \nu_i x_{ij} + \nu_o}{\sum_{r=1}^{s} \mu_r y_{rj}} \geq 1 \quad (j = 1 \cdots n), \\ &where\ \nu_i, \mu_r \geq 0, \nu_o\ \ free \end{aligned} \tag{1}$$

The set of constraints requires that the same weights, when applied to all the airports, do not provide any airport with efficiency lower than one. The solution to this minimization problem is not unique. It can be shown that if there exists a solution (ν,μ) to the above problem, then there exist an infinite number of solutions because $((\phi\nu,\phi\mu),\phi \geq 0)$ is also a solution to the problem (Coelli, 1996). Since, there are an infinite number of solutions for the dual variables (multipliers), it is necessary to formulate an equivalent linear programming program which avoids this problem. In this sense, the following problem is resolved for each airport:

---

[5] The different assumptions about the scalar produce distinct envelopment surfaces: VRS, CRS or extensions of these basic models.

[6] This discussion is very close to the definition of Pareto-Koopmans efficiency. The unit o is considered fully efficient if and only if the performance of other DMUs does not provide evidence that some of the inputs or outputs of the unit o could have been improved without worsening off some of its other inputs or outputs. This definition of relative performance has its origin in Farrell (1957).

$$\min_{v\mu} \sum_{i=1}^{m} v_i x_{io} + v_o$$
$$s.t.$$
$$\sum_{i=1}^{m} v_i x_{ij} + v_o - \sum_{r=1}^{s} \mu_r y_{rj} \geq 0 \quad (j = 1 \cdots n), \tag{2}$$
$$\sum_{r=1}^{s} \mu_r y_{ro} = 1$$
$$where\ v_i, \mu_r \geq 0, v_o \ free$$

An airport is in the frontier if and only if $\sum_{i=1}^{m} v_i x_{io} + v = 1$ in optimality. The constraint $\sum_{r=1}^{s} \mu_r y_{ro} = 1$ is known as a normalization constraint, and the weighted input and output are called virtual input and virtual output, respectively. See Seiford and Thrall (1990) for a detailed discussion of these models. The efficiency ratio ranges from 1 to infinity. Thus, each airport will choose weights so as to minimize self-efficiency, given the constraints.

2.1. Cross-efficiency DEA model

Sexton et al. (1986) were the first to develop the cross-efficiency evaluation matrix, initiating the subject of ranking in DEA. Doyle and Green (1994) validated this method, saying that decision makers do not always have a reasonable prior knowledge from which to estimate assurance regions for multipliers, and thus they recommended the cross-efficiency evaluation matrix for ranking units. The cross-efficiency evaluation method simply calculates the efficient score for each airport n times, using the virtual multipliers obtained in each of the n linear programming programs resolved before. The results of all the DEA cross-efficiency scores can be summarized in a cross-efficiency matrix as following:

$$h_{kj} = \frac{\sum_{i=1}^{m} v_{ik} x_{ij} + v_k}{\sum_{r=1}^{s} \mu_{rk} y_{rj}}, (k = 1, \cdots, n,\ j = 1 \cdots n) \tag{3}$$

Thus, h_{kk} represents the score given to airport j in the DEA run of airport k, i.e. the performance of airport j is evaluated by the weights of airport k. Note that all the elements in the matrix are in the range one and infinity, and the elements in the diagonal, h_{kk}, represent the standard DEA efficiency score (the elements in the diagonal equal 1 for efficient airports and greater than 1 for inefficient airports, accordingly to a conventional DEA methodology). Sexton et al. (1986) established a set of secondary goals for either aggressive or benevolent DMUs. In this context, a DMU could be considered aggressive if it minimizes self-efficiency and at a secondary level maximizes the other DMUs cross-efficiency scores. The benevolent secondary objective would be to equally minimize all DMUs cross-efficiency scores.

The cross-efficiency ranking method in this DEA context employs the results of the cross-efficiency matrix h_{kj} in order to rank all Spanish airports. There are different synthetic indexes that can be used in order to rank the performance of the airports. In this paper, we will use the average cross-efficiency score given to airport j defined as: $\bar{h}_j = \frac{1}{n} \sum_{k=1}^{n} h_{kj}$. However, averages are not the only possibility, as there are other standard univariate summaries, such as, median, variance or some other quantile point that could also be applied. It has been commented that any central measure of the cross-efficiency matrix, such as the average, represents better the performance of airports than h_{jj}, the standard DEA efficiency score. This is based on the fact that all the elements of the cross-efficiency matrix have been considered, meanwhile h_{jj} only includes the elements of the diagonal. Furthermore, all the airports are evaluated with the same sets of weight vectors. The minimum value of cross-efficiency is 1, which occurs if airport j is efficient in all the runs, i.e. all the airports evaluate unit j as efficient. In order to rank the units, we can simply assign the airport with the lowest score a rank of one and the unit with the highest score a rank of n. While the DEA scores h_{jj} are non-comparable, since each element uses different weights, the $\bar{h}_j$ score is comparable because it utilizes the weights of all the units equally. However, this is also the drawback of the technique, since the evaluation subsequently loses its connection to the multiplier weights (Adler et al., 2002).

2.2. Super-efficiency ranking methods

Andersen and Petersen (1993) developed a new procedure for ranking efficient units. The methodology enables an extreme efficient unit o to achieve an efficiency score lower than one by removing the oth constraint in the primal formulation, as shown below

$$\begin{aligned}
&\min_{\nu\mu} \sum_{i=1}^{m} \nu_i x_{io} + \nu_o \\
&s.t. \\
&\sum_{i=1}^{m} \nu_i x_{ij} + \nu_o - \sum_{r=1}^{s} \mu_r y_{rj} \geq 0 \quad (j = 1 \cdots n, j \neq o), \qquad (4) \\
&\sum_{r=1}^{s} \mu_r y_{ro} = 1 \\
&\text{where } \nu_i, \mu_r \geq 0, \nu_o \text{ free}
\end{aligned}$$

When a DMU under evaluation is not included in the reference set of the envelopment models, the resulting DEA models are called super-efficiency DEA models. Charnes et al. (1992) use a super-efficiency model to study the sensitivity of the efficiency classification. Zhu (1996) and Seiford and Zhu (1998) develop a number of super-efficiency models to determine the efficiency stability regions. Seiford and Zhu (1999) studied the problems associated with possible infeasibilities that may appear in some super-efficiency models. Wilson (1993) used the super-efficiency DEA models to detect influential observations. Looking at equation 4, we see that the difference between the super-efficiency and the envelopment models is that the DMU_o under evaluation is excluded from the reference set in the super-efficiency models, i.e., the super-efficiency DEA models are based on a reference technology constructed from all other DMUs.

However, some potential problems may appear with this methodology. First, it is necessary to have in mind that even if we use the DEA objective function value as a rank score for all units, each unit is evaluated according to different weights. This value in fact explains the proportion of the efficiency score that each unit o attained with its chosen weights in relation to a virtual unit closest to it on the frontier. Second, the super-efficiency methodology can give ‘‘specialized’’ DMUs an excessively high ranking. To avoid this biased problem, Sueyoshi (1999) introduced specific bounds on the weights in a super-efficient ranking model. And third, sometimes the super-efficiency model is infeasible for some efficient DMUs. When this problem appears, this technique does not provide a complete ranking of all the DMUs. Zhu (1996) shows that the input-oriented CRS super-efficiency model is infeasible if and only if a certain pattern of zero data occurs in the inputs and outputs elements. Despite these drawbacks, possibly because of the simplicity of the concept, many published papers have used these super-efficiency models.

2.3. Virtual super-efficiency model

We have already discussed that basic DEA models classify the DMUs into two sets, those that are efficient and define the Pareto frontier and those that are inefficient. Many empirical papers have adapted the models to deal with problems that have occurred in practice. However, these models need a major adaptation in the field of ranking DMUs, because practitioners cannot compare the units that lie on the frontier. It is often quite common that decision-makers are interested in a complete ranking, beyond the dichotomized classification of efficient and inefficient units, in order to understand the overall performance of the units. One problem that has been frequently discussed in the literature has been the lack of discrimination in some DEA applications, in particular when there are insufficient DMUs or the number of inputs and outputs is too high relative to the number of units. According to Cooper et al. (2000) a good rule of thumb for the number of DMUs in applying DEA is $n \geq \max\left(m \times s, 3(m+s)\right)$, where n is the number of DMUs, m is the number of inputs and s in the number of outputs. This is an additional reason for the growing interest in complete ranking techniques. Furthermore, fully ranking units is an established approach in the social sciences.

To achieve a full ranking of all airports, a virtual super efficient airport is introduced and included with the rest of existing airports. This will ensures that there is only one efficient airport (efficiency value of 1) with all the real airports being inefficient. The efficient frontier, based on this model, therefore consists of only this virtual super-efficient airport. All other airports are inefficient and are penalized for not operating at the same scale of efficiency. This approach serves to rank all the airports, and has been employed in previous studies (Bazargan and Vasigh, 2003). This ranking is justified because the same virtual airport is used for all airports as the reference set.

The input and output for this virtual super-efficient airport are:

$$X_v = \min_j \left\{ X_j \right\},$$

$$Y_v = \min_j \left\{ Y_j \right\},$$

where X_v and Y_v are the input and output vectors of the virtual super-efficient airport and X_j and Y_j are the input–output vectors of the jth airport. In other words, the virtual airport has the lowest input and the highest output among the airports considered in the study.

The DEA model is run with the inclusion of this new virtual airport and the efficiency scores are used to fully rank the airports because as expected, the virtual airport is the only efficient airport.

3. The data and sample of the airports

Our sample includes 34 Spanish airports that have different size and form part of AENA. We used data of the Spanish airports for the year 1997 to evaluate the efficiency of the airports estimating the DEA models described above. The performance of airports is going to be clearly linked with the selection of the outputs that an airport produces and the inputs airport use in producing these outputs. In the revision of the literature, the most commonly used output measure for airports is the number of passengers processed, as the most important function of an airport is to serve as an interface between land and air transport. Another important output for airports is air cargo. Air cargo can be served with passenger planes or with dedicated freighters. Some Spanish airports have expanded its facilities to accommodate new cargo terminals, because air cargo is becoming more important for some high-value goods. Passengers and cargo handling can be considered as final outputs of an airport, and these are associated with passenger and cargo terminal buildings that conform the landside operations of an airport. On the other hand, air traffic movements are considered as an intermediate output that is associated with the airside operations of an airport, runways, taxiways, aprons and other elements. We measured output with three variables: the air traffic movements, the number of passengers and the number of tons of cargo transported in the airport. The input variables were introduced as expenditures and were divided according to the following classification: labor, capital and materials[7].

These variables have a clear meaning and the fact that the source of the data is AENA, clearly helps in reducing the problems of comparability. This is especially true in reference to the capital cost. Some differences in accounting practices usually difficult the comparison of these variables in the studies of airport with distinct nationalities or type of ownership.

4. Empirical results

As discussed earlier, we will use four different approaches to rank the overall performance of the Spanish airports for the year 1997. Table 1 shows four different super-efficiency scores for the airports included in the sample. The first column expresses the cross-efficiency ratio measure, the second and third columns are the super-efficiency VRS and NIRS output orientation scores, respectively. And finally, the fourth column displays the virtual super-efficiency score[8].

An examination of the table 1 reveals that, according to the cross-efficiency score, Lanzarote, Barcelona, Madrid, Tenerife norte, Ibiza, Gran Canaria and Tenerife sur are the most efficient airports in the Spanish System. It is interesting to remark that these airports are located in the main cities of Spain, Madrid and Barcelona, or in tourist island cities, the rest of the airports. Looking at the opposite direction, it can be observed that Jerez, Santander, San Sebastian, Vitoria, Girona, San Javier and Hierro are the less efficient airports in the Spanish System. In this respect, we would like to remark that only the airport in the island Hierro could be sustained by public service obligation, due to their insular characteristic of the population that use its facilities. However, the rest of the airports are not far from other airports in their respective regions. For example, San Sebastian, Santander and Vitoria are near the airport of Bilbao.

Comparison between third and fourth columns of table 1 shows that there are no big differences between these two approaches to measure super-efficiency. In fact, only the airports operating in the area of increasing returns to scale show some difference. It can be seen that there is only one airport, Valladolid, which presents the unfeasibility problem. However, this problem is not present with the output-oriented NIRS super-efficiency model. If we rank the airport with the super-efficiency NIRS, it can be seen that Madrid, Lanzarote, Melilla, Vitoria, Barcelona and Mallorca are the most efficient airports. It is remarkable the position change that some airports exhibit, e.g. the airport of Vitoria changes from the group of less efficient to the group of most efficient. However some airports, like Madrid, Barcelona and Lanzarote, present a more stable ranking behavior. On the other hand, it can be seen that San Javier, Almeria, Hierro, Sevilla, Granada and Girona conform the group of the less efficient airports.

---

[7] The authors are conscious that some other measures of input, such as, number of runways, number of gates, terminal area and number of employees would have made the experiment more realistic but lack of available data preclude us from using these kind of variables. Gillen and Lall (1997) applied DEA to the airport sector using real input variables and measuring the efficiency of two different productive processes: terminal services and movements. The envelopment surfaces are estimated according to variable returns to scale and constant returns to scale, respectively.

[8] We note here that all the cross-efficiency measures are greater than 1. The cross-efficiency evaluation matrix has been calculated according to the formulation of DEA-LP programs described by equation 3.

Table 1. Super Efficiency Measures of Spanish Airports during 1997

Airport	Cross_Eff.	NIRS_Sup_Eff	VRS_Sup_Eff	Virt_Eff
Alicante	2.005	1.409	1.409	5.257
Almeria	3.693	2.924	2.924	32.107
Asturias	3.579	2.242	2.199	30.992
Barcelona	1.126	0.717	0.717	1.214
Bilbao	1.749	1.377	1.346	8.806
Coruña	3.960	2.392	2.316	43.710
Fuerteventura	1.839	1.316	1.162	9.476
Girona	7.416	3.862	3.632	45.606
Granada	4.512	3.299	3.160	43.373
Hierro	7.999	2.960	2.231	106.961
Ibiza	1.514	1.338	1.338	6.554
Jerez	4.797	2.624	2.558	41.904
Lanzarote	1.018	0.594	0.593	5.773
La Palma	2.673	1.925	1.895	23.792
Gran Canaria	1.576	1.193	1.193	2.917
Madrid	1.307	0.305	0.305	1.000
Malaga	2.148	1.384	1.384	3.216
Mallorca	1.709	0.768	0.768	1.406
Melilla	2.463	0.651	0.651	25.359
Menorca	1.686	1.494	1.494	9.739
Pamplona	3.314	1.270	1.212	29.877
Reus	3.775	2.310	1.426	44.637
Santander	5.233	2.104	2.040	53.143
Santiago	2.451	1.798	1.798	13.944
Sevilla	3.784	3.133	3.123	12.530
San Javier	7.974	2.629	2.461	90.185
San Sebastian	5.362	2.377	2.167	67.985
Tenerife norte	1.393	0.935	0.935	8.413
Tenerife sur	1.616	0.883	0.883	3.109
Valencia	2.008	1.551	1.551	7.466
Valladolid	3.846	1.133	infeasible	67.948
Vigo	2.952	1.949	1.863	32.071
Vitoria	6.061	0.678	0.678	8.446
Zaragoza	3.553	0.791	0.753	20.155

Apparently, there are no noticeable differences in performance rankings between cross-efficiency scores and virtual-efficiency values. In this sense, Madrid, Barcelona, Mallorca, Gran Canaria, Tenerife sur and Malaga appear to be the higher performers. On the other hand, Girona, Santander, Valladolid, San Sebastian, San Javier and El Hierro appear to be the lower performers. This measure is highly affected by the size of the airport.

It is remarkable that the ranking DEA performance measures introduced to treat the problem empirically often give conflicting indications of airport performance. Therefore we will also compute the Spearman rho rank correlation coefficients of the ranking DEA methods shown in table 1. The ranking DEA measures simply rank the overall performance of the airports, trying to make comparisons between the airports that lie on the frontier. Thus, we only need to pay attention to rank orderings according to the distinct methodologies. Spearman's rho correlation coefficient is really related to Pearson's correlation coefficient, because it is simply a Pearson correlation coefficient computed on the same data after converting them to ranks. Table 2 shows that different approaches can be considered complements instead of substitutes, because they present a different ranking perspective of airport performance. In the trivial case, in which all the approaches had given the same ranking of the airports, we would have obtained a matrix of 1. Table 2 shows that the cross-efficiency and virtual efficiency ranking methods are highly influenced by the size of the airports, in which we have used the number of passengers to approximate the concept of the size of airports. Other issues that we need to highlight are based on the strong relationship that exists between both super-efficiency methods. In fact, as we have already explained, only small differences are observed for airports operating in the area of increasing returns to scale. However, it is clear that the NIRS super-efficiency ranking method allows decision-makers to obtain a full ranking of all the airports, and VRS could present some infeasibilities that preclude DMUs from a full ranking.

Table 2. Spearman rank correlation coefficients of ranking DEA methods.

	CE	SE-NIRS	SE-VRS	V
Cross-Efficiency	1			
Super-Efficiency NIRS	0.732	1		
Super-Efficiency VRS	0.703	0.985	1	
Virtual Efficiency	0.863	0.768	0.728	1
Passengers	-0.874	-0.492	-0.461	-0.872

As it was explained in the methodological section, Sexton et al. (1986) were the first to study and introduce the concept of cross-efficiency methods, in which the airports in our case study, are both self and peer evaluated. Decision makers do not have obvious production functions to aggregate the data consistently, so relative efficiency of DEA ranking methods are a good substitute for this task. Now we are going to pay attention to the first of our methods: cross-efficiency evaluation matrix. Figure 1 shows the box-plot[9] of the input multipliers of the airports. It can be seen that input multipliers associated with labor and capital are zero for almost all the airports in the sample. In fact, there are 20 airports for which the labor input multiplier is zero, and Hierro's airport is the one that presents the highest labor input multiplier ($1.93\ 10^{-5}$). If we focus our attention to the capital input multiplier, it can be seen that 18 airport present a value of zero, and San Javier's airport is the one with the highest capital input multiplier ($2.61\ 10^{-5}$). There are ten airports for which both input multipliers labor and capital are zero. For these airports, relative efficiency is almost determined by the free variable v of the linear programming problem (the constant of returns to scale in which the airport operates) and the materials input multiplier. It can be seen, that the median of this input multiplier is distinct from zero, and there are only seven airports for which the materials input multiplier is zero, and Girona is the airport with the highest materials input multiplier ($3.24\ 10^{-5}$).

[9] Box-plots are a extraordinary tool to summarize a great deal of information very clearly. First, it is very good at showing extremes and/or outliers values. Clearly, this is an interesting task for the exercise we are doing. We can highlight the outliers that are present in both, the input multipliers and output multipliers. Thus, airport regulators can obtain an overall picture of the different behavior of the airports included in the sample. Now we would explain briefly the basic characteristics of this tool: the vertical line inside the box shows the median of the distribution; the left and right sides of the box show two important quantiles, the 25 and 75 percentiles, respectively (i.e. it gives the location of the middle 50 percent of observations). The vertical marks joined to the box by the dashed line (they usually are known as whiskers) show observations of the data whose distance with the sides of the box is less than 1.5 times the interquartile range (difference between the third and first quartile). Points outside these vertical lines are outliers and are drawn as small circles. Box-plots not only show the location and spread of the data but also give a good indication of skewness with the sizes of the left and right parts of the box.

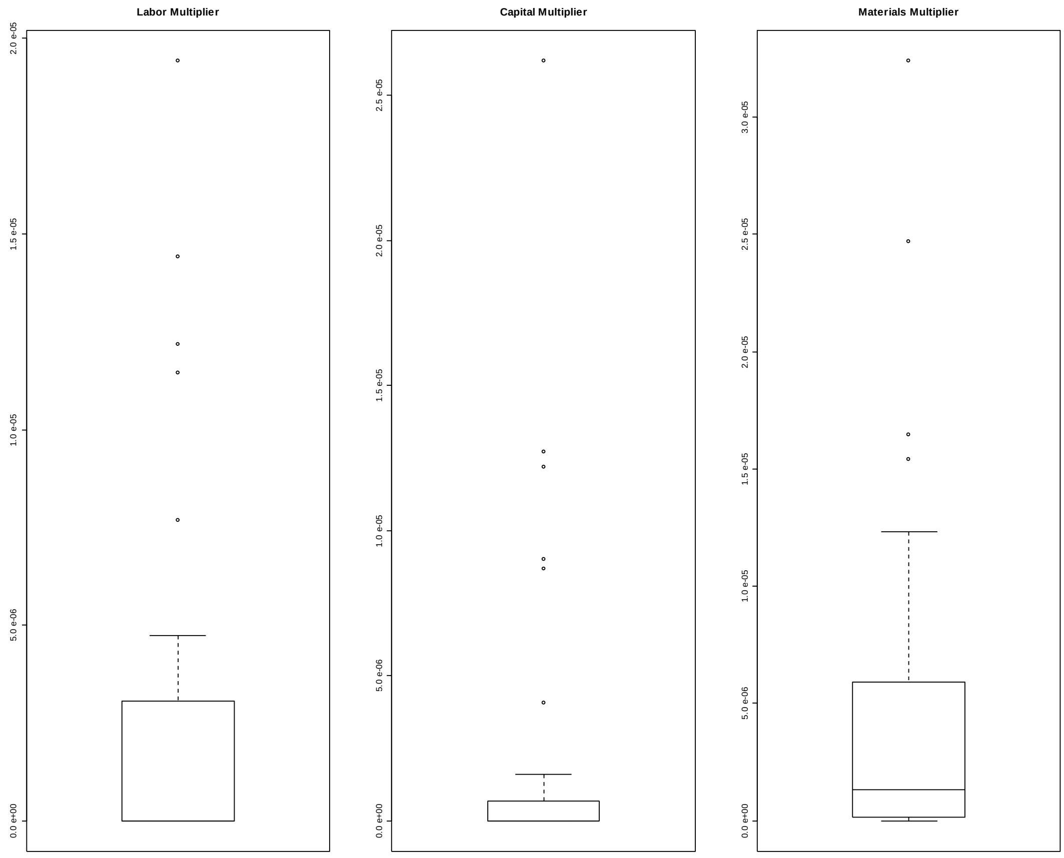

Figure 1. Box-Plot. Input Multipliers of Cross-Efficiency Ranking Method

Figure 2 shows the box-plot of the output multipliers and the constants of returns to scale of the airports in the sample. It can be seen that output multipliers associated with passengers and cargo are zero for many airports in the sample. In fact, there are 18 airports for which the passengers output multiplier is zero, and Girona and Reus are the airports that present the highest passengers output multiplier (0.0019). Similarly, analyzing the behavior of cargo output multiplier, it can be concluded that 19 airports present a value of zero, and Santander is the airport that present the highest cargo output multiplier (0.0003). It is also interesting to remark that the median of the ATMs output multiplier is distinct from zero, and there are only five airports for which this output multiplier is zero. Hierro's airport presents the highest value (0.0004) with respect to this multiplier. The values of these output multipliers are really consistent because air traffic movements are the primary output from airports' perspective. Passengers and cargo are transported in planes and landings and takes-off are necessary to produce air passengers and cargo. Analyzing airlines perspective, the primary outputs are passengers and cargo units. Of course, ATMs are essential but they can be considered an intermediate output. There exists a vague relationship between ATMs, load factors and size of the airplanes. Pels et al. (2003) study the relationship between ATMs, passengers, runway efficiency and terminal efficiency. They consider that load factors may influence terminal efficiency, because there exists a positive relationship between load factors and terminal efficiency, e.g. if load factors are high, airport would move more passengers with a fixed number of ATMs. This would imply that airports would be relatively more efficient. So terminal efficiency is influenced by load factors of airlines and this strategic variable is not usually under the control of airport managers.

Table 3. Cross-efficiency evaluation matrix of good and low performers

Airport	Lanz	Bar	Mad	Ten	Ibi	Gca	Sant	San S.	Vits	Gir	San J.	Hie
Lanzarote	1.000	1.000	1.000	1.000	1.487	1.386	3.860	5.149	1.187	7.346	7.115	7.883
Barcelona	1.011	1.000	1.000	1.000	1.501	1.393	3.854	5.150	1.177	7.491	7.103	7.875
Madrid	1.000	1.000	1.000	1.688	1.477	1.371	6.933	6.033	4.094	7.121	9.124	7.704
Tenerife norte	1.011	1.000	1.000	1.000	1.501	1.393	3.854	5.150	1.177	7.491	7.103	7.875
Ibiza	1.000	1.000	1.399	1.267	1.338	1.503	4.502	5.410	2.971	6.895	7.317	8.343
Gran Canaria	1.000	1.000	1.000	1.731	1.415	1.193	11.485	13.158	4.867	7.892	19.510	21.994
Santander	1.000	1.686	1.320	1.024	1.951	1.874	2.040	2.287	1.000	4.889	3.937	3.167
San Sebastian	1.000	1.000	1.356	1.297	1.400	1.656	2.642	2.167	2.875	5.870	3.410	2.241
Vitoria	1.083	1.093	1.000	1.000	1.724	1.564	2.608	2.856	1.000	7.380	4.339	3.551
Girona	1.000	2.716	3.013	2.064	1.825	2.642	3.712	2.820	30.157	3.632	5.274	4.181
San Javier	1.000	1.145	1.702	1.333	1.506	1.753	3.588	2.713	2.315	10.584	2.461	3.337
Hierro	1.025	1.000	1.365	1.309	1.426	1.682	2.649	2.172	2.832	6.085	3.410	2.231

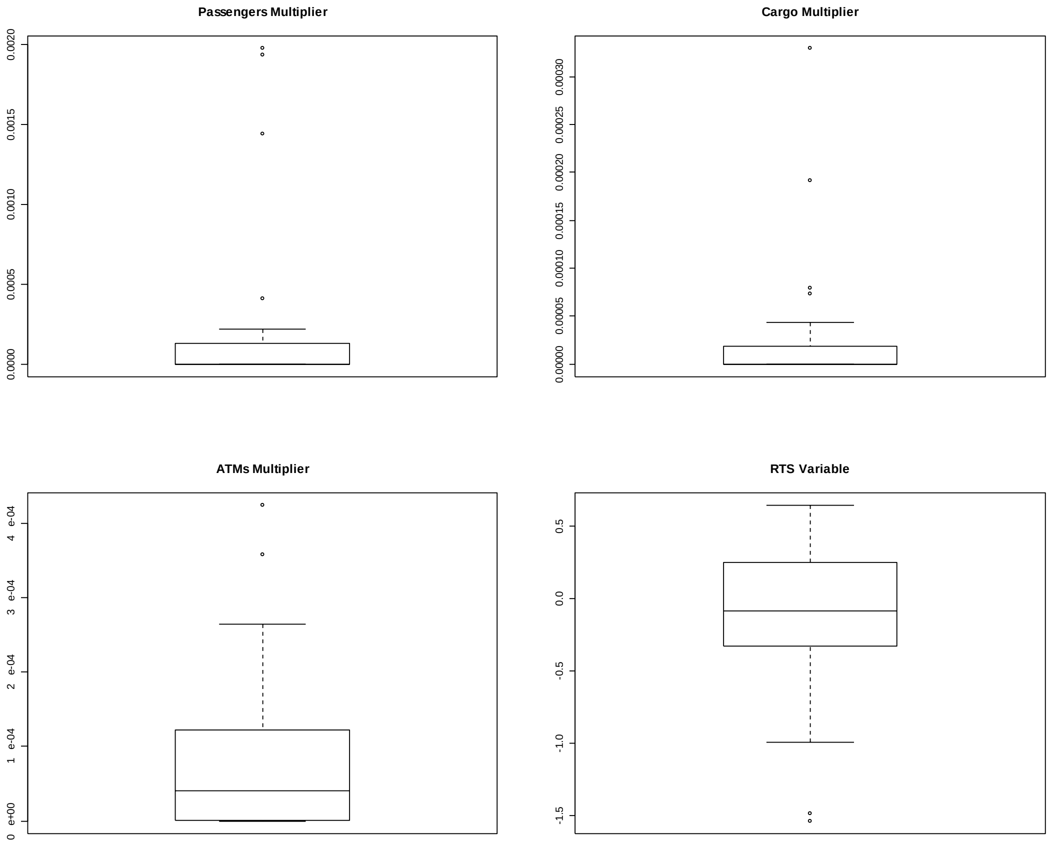

Figure 2. Box-Plot. Output Multipliers of Cross-Efficiency Ranking Method

Figure 2 shows the box-plot of the output multipliers and the constants of returns to scale of the airports in the sample. It can be seen that output multipliers associated with passengers and cargo are zero for many airports in the sample. In fact, there are 18 airports for which the passengers output multiplier is zero, and Girona and Reus are the airports that present the highest passengers output multiplier (0.0019). Similarly, analyzing the behavior of cargo output multiplier, it can be concluded that 19 airports present a value of zero, and Santander is the airport that present the highest cargo output multiplier (0.0003). It is also interesting to remark that the median of the ATMs output multiplier is distinct from zero, and there are only five airports for which this output multiplier is zero. Hierro's airport presents the highest value (0.0004) with respect to this multiplier. The values of these output multipliers are really consistent because air traffic movements are the primary output from airports' perspective. Passengers and cargo are transported in planes and landings and takes-off are necessary to produce air passengers and cargo. Analyzing airlines perspective, the primary outputs are passengers and cargo units. Of course, ATMs are essential but they can be considered an intermediate output. There exists a vague relationship between ATMs, load factors and size of the airplanes. Pels et al. (2003) study the relationship between ATMs, passengers, runway efficiency and terminal efficiency. They consider that load factors may influence terminal efficiency, because there exists a positive relationship between load factors and terminal efficiency, e.g. if load factors are high, airport would move more passengers with a fixed number of ATMs. This would imply that airports would be relatively more efficient. So terminal efficiency is influenced by load factors of airlines and this strategic variable is not usually under the control of airport managers.

The last box-plot shows the distribution of the constant associated to the economies of scale. It can be seen that the median of this variables is less than zero. This fact means that most of the airports in Spain are operating in the area of increasing returns to scale.

Finally, we would like to study the relationship between the cross-efficiency of the airports for the groups of low and high performers. We would analyze the impact that different objectives of the airports may produce on the self or peer evaluation, i.e. we would like to show the differences that can potentially exist between the scores obtained when an airport is evaluated according to the multipliers of its group (self evaluation) or the multipliers of the other group (peer evaluation). This exercise tries to shed some light in the controversial issue of the comparability of the airports. It is frequent for some airport managers to defend their poor performance, saying that airports in different regions have different objectives.

First, we have chosen twelve airports (6 good performers that will be denoted by G, and six low performers that will be denoted by L). Table 3 shows the cross-efficiency evaluation matrix for these twelve airports. They are ordered according to the first ranking method of table 1. Thus, the good performers appear in the six first rows and columns; and the low performers in the last six rows and columns. Once, this sub-matrix has been extracted we proceed to create the following associated factor GG, if the score analyzes the behavior of a good performer with the weights or multipliers of a good performer, e.g. the element of the matrix (a_{15}) is associated with GG because it is the evaluation of a good performer (Ibiza) peered by other good performer (Lanzarote). Similarly, we create factors associated with the rest of the values of the matrix as GL, LG or LL according to whether the airport is self or peered evaluated. Thus, we create the following factor-matrix:

$$B = (b_{ij}) = \begin{cases} GG & 1 \le i, j \le 6 \\ GL & 1 \le i \le 6 \quad 7 \le j \le 12 \\ LG & 7 \le i \le 12 \quad 1 \le j \le 6 \\ LL & 7 \le i, j \le 12 \end{cases} \qquad (5)$$

Second, we use one-way analysis of variance to test if the cross-efficiency score differs significantly across the factors that were created. The interest of the experiment is to show if there is any difference in cross-efficiency score between the different groups created according to equation 5. Results from anova[10] show that there is a statistical significant difference across the factor group. Table 4 shows the results of this anova, and it can be seen that the null hypothesis, i.e. the average performance of the airports is equal whether the airport is self or peered evaluated, may be rejected. The p-value, shown in the sixth column, casts doubt on the null hypothesis and suggests that at least the performance of one group of airports is significantly different than the other groups. The grand and factor means can be observed in table 4.

[10] Table 4 shows the standard anova table, which divides the variability of the airport cross-efficiency evaluation into two parts: variability due to the differences among the factor groups means (variability between groups); and variability due to the differences between the airport performance in each group and the group mean (variability within groups). The ANOVA table has seven columns: The first shows the source of the variability; the second shows the Sum of Squares (SS) due to each source; the third shows the degrees of freedom (df) associated with each source; the fourth shows the Mean Squares (MS) for each source, which is the ratio SS/df; the fifth shows the F statistic, which is the ratio of the MSs.; the sixth shows the p-value, which is derived from the cumulative distribution function of F. As F increases, the p-value decreases; and finally the seventh shows the significant code associated to the p-value. The choice of a critical p-value to determine whether the result is judged "statistically significant" is left to the researcher. It is common to declare a result significant if the p-value is less than 0.05 or 0.01.

Table 4. One-way analysis of variance. Cross-Efficiency Evaluation Matrix

	Df	SumSq	MeanSq	Fvalue	Pr(>F)	
Case	3	805.41	268.47	25.978	2.021e-13	***
Residuals	140	1446.82	10.33			

Signif. codes: 0 `***' 0.001 `' 0.01 `*' 0.05 `.' 0.1 ` ' 1**

Tables of means
Grand mean: 3.487153

Factor means

GG	GL	LG	LL
1.196	7.032	1.487	4.234

Tukey multiple comparisons of means
95% family-wise confidence level

	diff	lwr	upr
GL-GG	5.8355556	3.8653743	7.8057369
LG-GG	0.2905556	-1.6796257	2.2607369
LL-GG	3.0380556	1.0678743	5.0082369
LG-GL	-5.5450000	-7.5151813	-3.5748187
LL-GL	-2.7975000	-4.7676813	-0.8273187
LL-LG	2.7475000	0.7773187	4.7176813

In the previous anova, we have compared the performance of the groups of the airports and tested the hypothesis that the average cross-efficiency scores are all the same, against the general alternative that they are not all the same. However, as we have accepted the alternative hypothesis and it is too general, we would like to obtain more general information about which pairs of means are significantly different, and which are not. For this reason, we have also studied pair wise mean differences to assess in what sense a group can be characterized by its better or lower performance. To do this, we need to employ some multiple comparison procedure. In our case, we have decided to employ the Tukey-Kramer test in order to determine if average cross-efficiency evaluation of airport performance differences between distinct groups are statistically different from zero. As we want to compare all the four groups to one another, one can form 6 unique pairs of groups to obtain their mean differences (GL-GG, LG-GG, LL-GG, LG-GL, LL-GL and LL-LG). Differences and 95% confidence interval for these differences appear in table 4. For example, it can be seen that the difference between the groups GL and GG is 5.83 and a 95% confidence interval for the true mean is [3.86, 7.80]. In this example the confidence interval does not contain 0, so the difference is significant at the 0.05 level, and we can conclude that the performance of `lower performers' is worse than `good performers` when each group of airports is evaluated according to the multipliers of the group of `good performers`. If the confidence interval would have contained 0, the difference would not have been significant at the 0.05 level. (see for example the second row in the Tukey confidence intervals). In this case, we can conclude that the performance of the airports considered `good performers` is the same, independently of whether they have been self or peered evaluated. So we can conclude, that the performance of the group of `good performers` is quite robust to the peered evaluation. However, the fifth row shows that focusing on the performance of the group of `low performers`, it can be seen that the behavior of the airports is lower if they are peered evaluated. Another interesting result that can be extracted from table 4 is that `lower performers airports` are always dominated by the performance of the 'good performers airports', no matter what multipliers are used. In this sense, we obtain results that can be considered highly robust. Figure 3 shows the confidence intervals that appear in table 4.

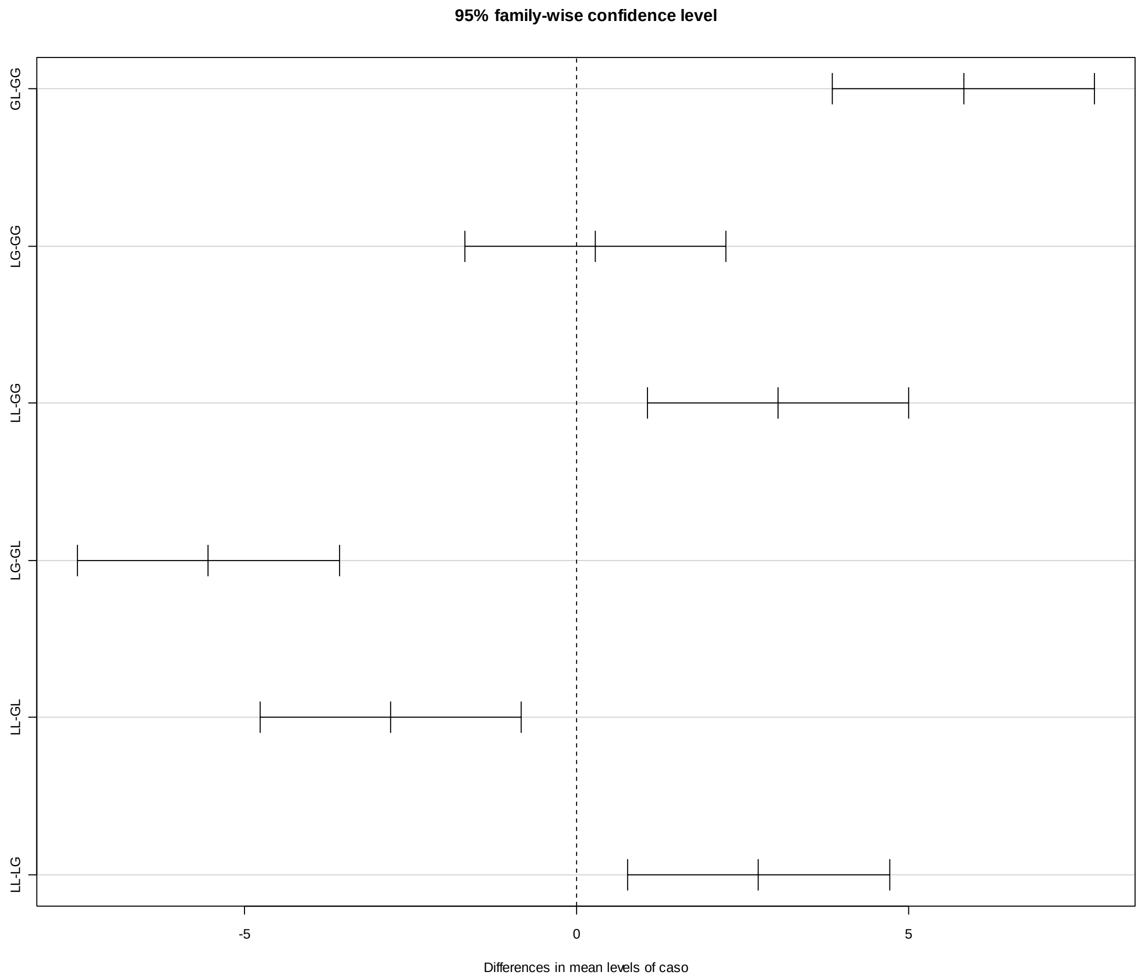

Figure 3. 95 % Confidence Intervals for means differences according to performance

5. Remarks and conclusions

The liberalization processes in the aviation sector have made popular the use of productivity measures, to compare the operational efficiency, effectiveness and the relative competitive position of airports and airlines. Airports' performance has been usually assessed based on some financial efficiency or operational efficiency throughout a big number of partial indicators. In this paper we present four different approaches, based on the flexibility of DEA, in order to fully rank the performance of Spanish airports. The first approach has involved the evaluation of a cross-efficiency matrix, in which each airport has been evaluated according to its weights and their rivals' weights. The second and third approaches have ranked the airports' performance through the exclusion of the airport being analyzed. And finally the fourth approach has been based in the analysis of the virtual airport and the comparison of the rest of the airports with this "champion performer". We have compared each approach, concluding that no one methodology can be prescribed here as the ideal solution to the question of ranking.

The cross-efficiency method has been recommended by some authors in order to adjust a reasonable mechanism from which to choose input and output multipliers to compare airport performance. In our case, this method reveals that some of the most efficient airports in Spain are located in the main cities and in tourist or island cities. However, the less efficient airports are located not far from other airports in their respective regions. For example, San Sebastian, Santander and Vitoria are near the airport of Bilbao, and some efficiency gains in the system can be plausibly obtained with some traffic diversion.

We also show that there are no big differences between the super-efficiency approaches. In fact, there are only some differences if the airports operate in the area of increasing returns to scale. We only obtained an airport with problems of unfeasibility with the output-oriented VRS super-efficiency model. In this case, the method cannot be used to fully rank the airports. For this reason, we also calculate the output-oriented NIRS super-efficiency model.

We have seen more similarities between the rankings obtained through the cross-efficiency and virtual-efficiency methods. However, we have also shown that the ranking DEA performance measures introduced to treat the problem empirically have produced conflicting indications of airport performance. Thus, we conclude that different approaches can be considered complements instead of substitutes, because they present a different ranking perspective of airport performance. To our knowledge, this is the first time that cross-efficiency has been applied to the airport industry, and whilst each method may be useful in explaining the overall performance of airports, none of the methods can be prescribed here as the best option to full rank the airports activity.

Finally, we have shown the relationship between the cross-efficiency of the airports for the groups of low and good performers, analyzing the potential differences that can exist according to the different objectives of the airports on the self or peer evaluation, i.e. the existing differences that exist between the scores obtained when an airport is evaluated according to the multipliers of its group (self evaluation) or the multipliers of the other group (peer evaluation). This exercise sheds some light in the controversial issue of the comparability of the airports, and defeats the defense that some airport managers use to justify their poor performance. We have concluded, that the performance of the group of `good performers` is quite robust to the peered evaluation and `lower performers airports` are always dominated by the performance of the 'good performers airports', no matter what multipliers are used.

References

Abbott, M. and S. Wu, 2002. Total factor productivity and efficiency of Australian airports, The Australian Economic Review 35, 244-260.

Adler, N., L. Friedman and Z. Sinuany-Stern, 2002. Review of ranking methods in the data envelopment analysis context, European Journal of Operational Research 140, 249-265.

Adler, N. and J. Berechman, 2001. Measuring airport quality from the airlines' viewpoint: an application of data envelopment analysis, Transport Policy 8, 171-181

Ali, A. and L.M. Seiford, 1993. The Mathematical Programming Approach to Efficiency Analysis. In: The Measurement of Productive Efficiency: Techniques and Applications, eds, Fried, H.O., Lovell, C.A.K and Schmidt, S.S. Oxford University Press, New York.

Andersen P. and N.C. Petersen, 1993. A procedure for ranking efficient units in data envelopment analysis, Management Science 39, 1261-1294.

Baltagi, B.H.; Griffin, J.M. and Rich, D.P., 1995. Airline Deregulation: The Cost Pieces of the Puzzle. International Economic Review 36 (1), 245-258.

Banker, R.D., A. Charnes and W.W. Cooper, 1984. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, Management Science 30, 1078-1092.

Bazargan M. and B. Vasigh, 2003. Size versus efficiency: a case study of US commercial airports, Journal of Air Transport Management 9, 187-193.

Button, K.J., Haynes, K. and Stough, R. 1998. Flying into the Future: Air Transport Policy in the European Union. Edward Elgar. Cheltenham.

Caves, D.W., Christensen, L.R., Tretheway, M.W. and Windle, R.J. 1987.An Assessment of the Efficiency Effects of US Airline Deregulation via an International Comparison. In Public Regulation: New Perspectives on Institutions and Policies, eds Bailey E. E. MIT Press, Cambridge Mass.

Charnes, A., Cooper, W., Lewin, A.Y., and Seiford, L.M., 1994. Data Envelopment Analysis. Theory, Methodology and Applications. Kluwer Academic. Boston.

Charnes, A., S. Haag, P. Jaska and J. Semple, 1992. Sensitivity of efficiency classification in the additive model of data envelopment analysis, International Journal of Systems Science 23, 789-798.

Charnes, A., Cooper, W.W. and Rhodes, E., 1978. Measuring the Efficiency of Decision Making Units. European Journal of Operational Research 2(6), 429-444.

Chin, A.T.H., Siong, L.E., 2001. Airport performance: a comparative study between Changi airport and airports in the New York-New Jersey metropolitan area. Paper presented at World Conference on Transportation Research, Seoul.

Coelli, T., Rao, D.S.P. and Battese, G.E., 1998. An Introduction to Efficiency and Productivity Analysis. Kluwer Academic. Boston.

Coelli, T.. 1996. A guide to DEAP version 2.1: a data envelopment analysis (computer) program. CEPA Working Paper 96/08. Centre for Efficiency and Productivity Analysis, University of New England, Armidale.

Cooper, W., Sieford, L., Tone, K., 2000. Date Envelopment Analysis. A Comprehensive Text with Models, Applications, Reference and DEA–Solver software. Kluwer Academic Publishers, Norwell.

Doganis, R.S. and Graham, A., 1987. Airport Management: The Role of Performance Indicators. Transport Studies Group, Polytechnic of Central London, London.

Doyle, J.R., Green, R., 1994. Efficiency and cross-efficiency in data envelopment analysis: Derivatives, meanings and uses. Journal of the Operational Research Society 45 (5), 567-578.

Emorouznejad, A., Thanassolis, E., 1997. An extensive bibliography of data envelopment analysis (DEA), Vol. 3, Supplement 1. Working Paper 258, Business School, University of Warwick.

Farrel, M.J., 1957. The Measurement of Productive Efficiency. Journal of the Royal Statistical Society A120, 253-290.

Francis, G., Humphreys, I., Fry, J., 2002. The benchmarking of airport performance. Journal of Air Transport Management 8, 239–247.

Francis, G.A.J., Fry, J., Humphreys, I., 2001. An International Survey of Performance Measurement in Airports, Open University Working Paper

Gillen, D. and Lall, A., 1997. Developing Measures of Airport Productivity and Performance: An Application of Data Envelopment Analysis. Transportation Research E 33(4), 261-273

Humphreys, I., Francis, G., 2002. Performance measurement: a review of airports. International Journal of Transport Management 1, 79-85.

Martin, J.C., Roman, C., 2001. An application of DEA to measure the efficiency of Spanish airports prior to privatization. Journal of Air Transport Management 7, 149-157.

Morrison, S. and Winston, C., 1995. The Evolution of the Airline Industry. Brookings Institution, Washington.

Oum, T.H., Yu, C. and Fu, X., 2003. A comparative analysis of productivity performance of the world's major airports: summary report of the ATRS global airport benchmarking research report-2002. Journal of Air Transport Management 9, 285-297.

Pacheco R.R. and E. Fernandes (2003), Managerial efficiency of Brazilian airports, Transportation Research 37A, 667-680.

Parker, D., 1997. The Performance of BAA under Privatisation and Regulation, Centre for the study of Regulated Industries, Occasional Paper no. 8.

Parker, D., 1999. The performance of BAA before and after privatization. Journal of Transport Economics and Policy 33, 133-145.

Pels, E., Nijkamp, P., Rietveld, P., 2001. Relative efficiency of European airports. Transport Policy 8, 183-192.

Pels, E., Nijkamp, P., Rietveld, P., 2003. Inefficiencies and scale economies of European airport operations, Transportation Research Part 39E, 341-361

Sarkis, J., 2000. An analysis of the operational efficiency of major airports in the United States. Journal of Operations Management 18, 335–353.

Seiford, L.M., Thrall, R.M., 1990. Recent developments in data envelopment analysis: The mathematical programming approach to frontier analysis. Journal of Econometrics 46, 7-38.

Seiford, L.M., Zhu, J., 1998. Stability regions for maintaining efficiency in data envelopment analysis, European Journal of Operational Research 108, 127-139.

Seiford, L.M., Zhu, J., 1999. Infeasibility of super-efficiency data envelopment analysis models. INFOR 37 (2), 174-187.

Seiford, L.M., 1994. A DEA Bibliography, 1978-1992. In Data Envelopment Analysis. Theory, Methodology and Applications, eds Charnes, A., Cooper, W., Lewin, A.Y., and Seiford, L.M., Kluwer Academic. Boston.

Sexton, T.R., Silkman, R.H., Hogan, A.J., 1986. Data envelopment analysis: Critique and extensions. In: Silkman, R.H. (Ed.), Measuring Efficiency: An Assessment of Data Envelopment Analysis. Jossey-Bass, San Francisco, 73-105.

Stasinopoulos, D., 1992. The Second Aviation Package of the European Community. Journal of Transport Economics and Policy 26, pp, 83–87.

Stasinopoulos, D., 1993. The Third Phase of Liberalization in Community Aviation and the Need for Supplementary Measures. Journal of Transport Economics and Policy 27, pp, 323–328.

Sueyoshi, T., 1999. Data envelopment analysis non-parametric ranking test and index measurement: Slack-adjusted DEA and an application to Japanese agriculture cooperatives. Omega International Journal of Management Science 27, 315-326.

Vincent, D. and Stasinopoulos, D., 1990. The Aviation Policy of the European Community, Journal of Transport Economics and Policy 24, 95–100.

Wilson, P.W., 1993. Detecting Outliers in Deterministic Nonparametric Frontier Models with Multiple Outputs, Journal of Business and Economic Statistics 11, 319-323.

Zhu, J., 1996. Robustness of the efficient DMUs in data envelopment analysis, European Journal of Operational Research 90, 451-460.