An empirical study of Iranian regional airports using robust data envelopment analysis

Emad Roghanian* and Amin Foroughi
aDepartment of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran
bIslamic Azad University of Naragh Branch, Naragh Iran

1. Introduction

Data Envelopment Analysis (DEA) has been one of the most important tools on measuring the relative efficiency of different similar units such as transportation systems using terminals, airports, etc. In this study, we perform an empirical analysis on Iranian airports based on DEA methods to measure the efficiencies of various airports. One of the primary issues on many traditional DEA methods is that the data are almost always contaminated with noise. We use a DEA method which could handle the uncertainty associated with input and output data. The results of this comprehensive study show that most of the active airlines are practically inefficient and the government could significantly increase the efficiencies of the airports by setting new regulations and rules.
used for external macroeconomics and price factors and they report that technical, scale and mix efficiencies are high among the major Asia Pacific airports. DEA also can be used as an efficient tool for measuring the total factor productivity (TFP). Yoshida and Fujimoto (2004) use DEA for measuring the relative efficiency of Japanese airports and report TFP using endogenous-weight TFP methods and discuss whether there is any over-investment in Japanese regional airports. There are also other survey based techniques for evaluating the performance of different airports. Yeh and Kuo (2003), for instance, use fuzzy multi-attribute decision making technique to evaluate passenger service quality of 14 major Asia-Pacific international airports using different surveys. The primary concern on the method is the existence of noise on the data and this could impact the results of the relative efficiencies. Traditional DEA is normally formulated as linear programming problem and it can be solved for optimality using a direct implementation of Simplex method. The nature of uncertainty often is not clear and there is not much information about the behavior of the noise. One simple assumption may often hold with uncertainty is that the data has symmetric unknown distribution. This is the basic assumption of robust optimization techniques developed by many scientists. See Bertsimas and Sim (2004) to learn more about uncertainty and robust optimization but one important issue is the effects of uncertainty on the output results on a regular linear programming problem. Ben-Tal and Nemirovski show that a small change on some benchmark linear programming problems could generate some results which are neither feasible nor optimal. They also propose a new technique based on recent robust optimization methodologies which could immune the output results from the infeasibility by losing some of the optimality. The robust method of Ben-Tal and Nemirovski is based on having a counterpart which is mapped inside the feasible region. Therefore, the robust technique changes the structure of an ordinary linear problem into a nonlinear problem. Bertsimas and Sim (2004) propose a new robust optimization technique where the structure of the new robust problem is the same as the original one, e.g. in the case of DEA, the robust optimization model is formulated as a linear programming problem. Note that the robust solutions for Bertsimas and Sim's method are usually less conservative than Bertsimas method. Sadjadi and Omrani (2009, 2010) are believed to be the first who use the idea of robust optimization on the context of DEA. They examine different robust optimization techniques on traditional DEA and report some promising results on various electricity and telecommunication regularities. In this paper, we perform an empirical analysis on RDEA for Iranian regional airports and discuss the results in different situations. This paper is organized as follows. We first present the problem statement of the RDEA. The implementation of our RDEA is presented in the next section and we carefully analyze the results using DEA and RDEA and discuss the differences using some statistical tests.

2. Problem statement

Let $x_{ij}$ be the inputs for a decision unit with $i=1,\ldots,m$ and $y_{ij}$ be the outputs with $r=1,\ldots,s$ and $j=1,\ldots,n$. Let $u_i$ and $v_j$ be the dual variables associated with $x_i$ and $y_j$, respectively. The constant to scale DEA model is formulated as follows,

$$
\max \quad z = \frac{\sum_{r=1}^{s} u_r y_r}{\sum_{i=1}^{m} v_i x_i}.
$$

subject to

$$
\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1, \quad u_i, v_j \geq 0, \quad j = 1,\ldots,n
$$

Model (1) is the basis of traditional DEA and it is solved $j$ times to determine the relative efficiencies of various units. However, since (1) is nonlinear in structure, Charles et al. (1983) suggest a simple modification of the objective function to simplify the structure of the resulted problem as follows,
max \quad z = \sum_{r=1}^{s} u_r y_r.

subject to
\begin{align*}
\sum_{r=1}^{s} u_r y_{rj} & \leq 1. \\
\sum_{j=1}^{m} v_j x_{ij} & = 1 \\
u_r, v_j & \geq 0, \quad j = 1, \ldots, n
\end{align*}

(Note that the first constraint also becomes linear using a simple manipulation. Problem (2) has been widely used for the past three decades and the results are commonly accepted as a tool to measure the relative efficiency of different units. However, when there is uncertainty with the inputs and the outputs, one may use different techniques to make sure that a small change on input/output data does not change the output rankings.

3. Robust optimization

Consider a linear programming problem of the following form,

\[
\begin{align*}
\text{min} \quad c'w \\
\text{subject to} \quad Aw = b, \\
w & \geq 0,
\end{align*}
\]

where \( w \in R^{n \times 1} \) is the vector of unknown variables, \( A \in R^{m \times n} \) and \( b \in R^{m \times 1} \) and \( c \in R^{n \times 1} \). Let \( A \) and \( c \) are subject to uncertainty. Therefore (3) can be reformulated as follows,

\[
\begin{align*}
\text{min} \quad \tilde{c}'w \\
\text{subject to} \quad \tilde{A}w = b, \\
w & \geq 0,
\end{align*}
\]

where \( \sim \) denotes the uncertainty with \( \tilde{A} = [\tilde{a}_{ij}] \). The robust optimization approach presented by Bertsimas and Sim (2004) converts (4) into the following problem,

\[
\begin{align*}
\text{min} \quad c'w \\
\text{subject to} \quad a_i'w - \Gamma_i p_i - \sum_{j \neq j} q_{ij} & \geq 0 \quad \forall i, \\
p_i + q_{ij} & \geq e a_{ij} y_j \quad \forall i, j, \\
y_j - w_j & \leq y_j \quad \forall j, \\
p_i, q_{ij} & \geq 0, \quad \forall j
\end{align*}
\]

where \( \Gamma_i \) determines the uncertainty associated with each input parameter. When \( \Gamma_i = 0 \) there is no uncertainty. As \( \Gamma_i \) increases, the uncertainty also increases. The \( e \) is also the vector of uncertain values. The DEA model originally developed by Charnes et al. (1983) is as follows,
\[
\text{(CCR)} \quad \max \quad z = \sum_{r=1}^{s} u_r \tilde{y}_r.
\]
\[
\text{subject to} \quad \sum_{r=1}^{s} u_r \tilde{y}_{rj} \leq 1. \tag{6}
\]
\[
\sum_{i=1}^{m} v_i \tilde{x}_{ij} = 1 \quad u_r, v_j \geq 0, \quad j = 1, \ldots, n
\]

where \( \tilde{x} \) and \( \tilde{y} \) are the uncertain inputs and outputs which are associated with \( x \) and \( y \), respectively. In (6) each uncertain parameter lies in an interval of uncertainty. Applying (5) to (6) yields,

\[
\text{(RCCR)} \quad \max \quad Z
\]
\[
\text{subject to} \quad \sum_{i=1}^{m} v_i x_i = 1,
\]
\[
\sum_{r=1}^{s} u_r y_r - z - \Gamma p - \sum_{j=a} q_r \geq 0,
\]
\[
\sum_{i=1}^{m} u_i y_{ij} - \sum_{r=1}^{s} u_r y_{rj} - \Gamma p_j - \sum_{j=a} q_{rj} \geq 0, j = 1, \ldots, n \tag{7}
\]
\[
p_j q_{rf} \geq e y_{rf} x_r \quad \forall r, j
\]
\[
-z_r \leq u_r \leq z_r \quad \forall r
\]
\[
p_j q_{rf} \geq 0, u_r, v_i \geq 0.
\]

Problem (7) is linear programming problem where \( e \) is a vector of uncertain values, \( \Gamma \) is the budget of uncertainty, \( p \) and \( q \) are new dummy non-negative variables associated with uncertain parameters in (6). As we explained earlier, there are two advantages associated with Bertsimas and Sim's robust model. First, the robust DEA is still linear in the structure although we need to add some additional auxiliary variables. Second, \( \Gamma \) adjusts the uncertainty associated with all parameters. Next section, we examine two models (2) and (7) and compare their results using some statistical technique.

4. Experimental Results

In this section, we present the implementation of the proposed RDEA. In Iran, all domestic and international airports are managed by the government and privatization has not implemented to this industry yet. Although some of the management of small parts of the airports are privatized but the actual privatization of the airports has not appeared. There are many reasons behind the government’s decision and one of the most important one is the key element is the security. For the past three decades, the Iranian airports have never had any high-jack experience. Although security plays a key role on airport industry, having an efficient airport is also a goal for the government. The government subside the airport industry significantly by providing low price fuel and other services. However, the new regulation has requested all Iranian industries to increase their TFP significantly. In fact, the new regulation implies to increase the growth domestic product by eight percent per year and one third of this growth must come from the increase in productivity. One practical way to increase the productivity is to increase the efficiency of airport industry. In Iran, there are over 57 different airports but most of them are almost inactive and there are very few regular flights. Therefore, we decided to choose 21 active airports to provide a meaningful comparison. Table 1 shows the inputs and the outputs used for our DEA implementation.
Table 1
The inputs and the outputs of RDEA model

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
<th>Mean</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees</td>
<td>Sum of the people who work in the airport</td>
<td>151.43</td>
<td>148.08</td>
</tr>
<tr>
<td>Terminal area</td>
<td>Area of terminal</td>
<td>16885.24</td>
<td>21736.94</td>
</tr>
<tr>
<td>Length of runway</td>
<td>Surface of the asphalt road</td>
<td>5175.29</td>
<td>2023.61</td>
</tr>
<tr>
<td>Number of movements</td>
<td>Flights of domestic &amp; international</td>
<td>11332.71</td>
<td>19813.35</td>
</tr>
<tr>
<td>Number of Passengers</td>
<td>passengers</td>
<td>1389699.8</td>
<td>2453794.8</td>
</tr>
<tr>
<td>Amount of Cargo</td>
<td>Cargo</td>
<td>14104.29</td>
<td>25241.36</td>
</tr>
</tbody>
</table>

We first use the regular DEA model where there is no uncertainty associated with all input parameters. As usual, there are some units with the efficiency of one and we use the relaxation method given by Anderson (2004) to rank the most efficient units. The results indicate that eighteen units use their terminal space very efficiently. The other observation is that none of inefficient airport transits could use their band length optimally. For instance, Shiraz airport could reduce its band length up to 4490 meters. In terms of the output parameters, among three outputs, any increase on the number of flights could significantly affect the efficiencies of all non-productive airports. Next, we use the RDEA model to study the effects of the uncertainty on ranking of different airports. Table 2 summarizes the results of the implementation of our crisp and robust DEA. The first column of the table introduces the names of different airports; the second column shows the efficiencies of various units. As we can observe Mehrabad, Imam Khomeini and Ahvaz are three best performers among all 21 units. Note that Mehrabad was responsible for all domestic and international flights and during the past few years, all international flights from Tehran have been executed from Imam Khomeini airport. In order to rank the first three airports, we have used the Anderson and Peterson’s method (2004). The results of our implementation are demonstrated in the third column of table 2 where Mehrabad shows better performance among two other best performers. In order to see the effects of uncertainty on the efficiencies of the 21 units, we have used the proposed RDEA method. The results are examined in two different scenarios of e=0.05 and e=0.1. When the number of uncertain parameters is large, one may use the following,

\[
\Gamma = 1 + \Phi^{-1}_{(1-\xi)} \sqrt{\xi}
\]

where \(\xi\) is the number of uncertain parameters and \(\Phi\) is the CDF of a Gaussian distribution. However, for the case of our proposed method since there are only 3 inputs and outputs uncertain parameters we choose \(\Gamma = 3\) as recommended by Sadjadi and Omrani (2009). The fourth and the fifth columns of the table (2) summarize the results of the efficiencies of all units. Obviously, as the level of uncertainty increases, all units’ efficiencies decrease. For instance, Ahvaz airport is known as one of the most actives units in crisp model but in our robust model the efficiency is reduced to about 82% when the level of uncertainty for all input parameters is 0.1. One important point of this research is to study the effects of the overall ranking between the nominal and robust DEA methods. We have set up a null hypothesis of having the same ranking for both methods and perform a statistical test based on Spearman Pearson (Maritz (1981)) as follows,

\[
d_{\text{spearman}} = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2-1)},
\]
where \( n \) is the number of units and \( d_i \) is calculated as a difference between the rankings of two methods. Using on the results of table (2) yields \( d_{\text{speaman}} = 0.999984 \) which means, statistically, there is no significant difference between the DEA and RDEA ranking.

**Table 2**
The efficiencies of different airport units

<table>
<thead>
<tr>
<th>(DMU) AIRPORT</th>
<th>CCR</th>
<th>CCR A&amp;P</th>
<th>CCR e=0.05</th>
<th>CCR e=0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mehrabad</td>
<td>1</td>
<td>2.2719</td>
<td>0.904762</td>
<td>0.818182</td>
</tr>
<tr>
<td>Imam Khomeini</td>
<td>1</td>
<td>2.1976</td>
<td>0.904762</td>
<td>0.818182</td>
</tr>
<tr>
<td>Ahvaz</td>
<td>1</td>
<td>1.6699</td>
<td>0.904762</td>
<td>0.818182</td>
</tr>
<tr>
<td>Mashad</td>
<td>0.9959</td>
<td>-</td>
<td>0.90109</td>
<td>0.814861</td>
</tr>
<tr>
<td>Shiraz</td>
<td>0.8119883</td>
<td>-</td>
<td>0.734656</td>
<td>0.664354</td>
</tr>
<tr>
<td>Isfahan</td>
<td>0.5732527</td>
<td>-</td>
<td>0.518657</td>
<td>0.469025</td>
</tr>
<tr>
<td>Rasht</td>
<td>0.5611192</td>
<td>-</td>
<td>0.507679</td>
<td>0.459098</td>
</tr>
<tr>
<td>Birjand</td>
<td>0.5471406</td>
<td>-</td>
<td>0.495032</td>
<td>0.447661</td>
</tr>
<tr>
<td>Bandarabas</td>
<td>0.4805213</td>
<td>-</td>
<td>0.434757</td>
<td>0.393154</td>
</tr>
<tr>
<td>Kermanshah</td>
<td>0.4390564</td>
<td>-</td>
<td>0.397242</td>
<td>0.359228</td>
</tr>
<tr>
<td>Kerman</td>
<td>0.4389432</td>
<td>-</td>
<td>0.397139</td>
<td>0.359135</td>
</tr>
<tr>
<td>Gorgan</td>
<td>0.4324138</td>
<td>-</td>
<td>0.391232</td>
<td>0.353793</td>
</tr>
<tr>
<td>Ardabil</td>
<td>0.428176</td>
<td>-</td>
<td>0.387397</td>
<td>0.350326</td>
</tr>
<tr>
<td>Tabriz</td>
<td>0.4086879</td>
<td>-</td>
<td>0.369765</td>
<td>0.334381</td>
</tr>
<tr>
<td>Zahedan</td>
<td>0.4059032</td>
<td>-</td>
<td>0.367246</td>
<td>0.332103</td>
</tr>
<tr>
<td>Larestan</td>
<td>0.3186695</td>
<td>-</td>
<td>0.28832</td>
<td>0.26073</td>
</tr>
<tr>
<td>Boshehr</td>
<td>0.2683667</td>
<td>-</td>
<td>0.242808</td>
<td>0.219573</td>
</tr>
<tr>
<td>Oromieh</td>
<td>0.231795</td>
<td>-</td>
<td>0.209719</td>
<td>0.18965</td>
</tr>
<tr>
<td>Sari</td>
<td>0.2295115</td>
<td>-</td>
<td>0.207653</td>
<td>0.187782</td>
</tr>
<tr>
<td>Yazd</td>
<td>0.224646</td>
<td>-</td>
<td>0.203251</td>
<td>0.183801</td>
</tr>
<tr>
<td>Abadan</td>
<td>0.1851345</td>
<td>-</td>
<td>0.167503</td>
<td>0.151474</td>
</tr>
</tbody>
</table>

One of the most important issues on the implementations of DEA is to study the effects of a particular input or output on overall efficiencies of each unit. In other word, one single input may have significant influence of a unit but this factor may not necessarily play an important role on other units. We have run DEA method by deleting a single input or output each time. Since there are three inputs, three outputs and 21 units, we need to run 126 sub-problems. We have also studied the effects of all six the inputs and the outputs on efficiencies of the efficiencies. One important observation is that, among six factors, only cargo was the most influencing parameter for Mehrabad unit which could reduce the efficiency from one to 0.7051 which is a significant
reduction of 0.2949. The influence of the other factors was not as significant as the cargo one and is not reported here.

5. Conclusions

In this paper, we have examined data envelopment analysis for major Iranian airports to measure the relative efficiencies of various airports. The results of our proposed method indicate that there any only a few airports running efficiently and most of the Iranian airports are economically inefficient. The two major airports, Imam Khomeini and Mehrabad, located near the capital city of Iran are determined to be the most efficient units. However, the relative efficiencies of most of the other airports located in different regions of Iran are less than 50 percent. Since the input/output data used in this study may be subject to uncertainty, we have used a robust technique to measure the relative efficiencies. The statistical tests indicate that the results of ranking of the robust optimization method are the same as the crisp model.

Acknowledgment

The authors would like to thank the officials of the Iranian airports company and the transportation research institute of ministry of road and transportation of Iran for providing the necessary information needed for this project. Also, the authors would like to thank the anonymous referees for their comments on the earlier version of this work.

References

performance of Taiwan's major airport, Journal of Air Transport Management 10, 353-60.
Wu, T. H., Chen, M-S, Yeh, J-Y. (2010). Measuring the performance of police forces in Taiwan using data
Yeh, C. H., Kuo, Y. L. (2003). Evaluating passenger services of Asia-Pacific international airports,
methods: testing the criticism of over-investment in Japanese regional airports, Transportation Research
part E 40, 533–546.