Accepted Manuscript

Smoothed bootstrap Malmquist Index based on DEA model to compute productivity of tax offices

Ramón Fuentes, Adelaida Lillo-Bañuls

PII: S0957-4174(14)00685-X
DOI: http://dx.doi.org/10.1016/j.eswa.2014.11.002
Reference: ESWA 9654

To appear in: *Expert Systems with Applications*

Smoothed bootstrap Malmquist Index based on DEA model to compute productivity of tax offices

Ramón Fuentes
Faculty of Economics. Department of Applied Economic Analysis. University of Alicante.
Email: rfuentes@ua.es
Telephone and fax number: +34 965 90 97 11
Postal address:
Faculty of Economics. Department of Applied Economic Analysis. University of Alicante. P.O. Box 99. E-03080. Spain.

Adelaida Lillo-Bañuls
Faculty of Economics. Department of Applied Economic Analysis. University of Alicante.
Email: alillo@ua.es
Telephone and fax number: +34 965 90 34 00 ext.2709
Postal address: Faculty of Economics. Department of Applied Economic Analysis. University of Alicante. P.O. Box 99. E-03080. Spain.

Smoothed bootstrap Malmquist Index based on DEA model to compute productivity of tax offices

1. Introduction

The widespread use of the neo-classical doctrines as practically the only paradigm for providing an economic framework within which to carry out the production activity of a country or region has inevitably generated an increased interest in analysing efficiency and productivity in recent years.

Public services have attracted particular attention in this respect based on the opinion that their scope is limited due to their hypothetical x-inefficiency compared to the private sector, leading to logical attempts to improve management models in a way that does not compromise the fulfilment the objectives established by the state (González-Páramo & Onrubia, 2003).

This is the environment within which the SUMA offices develop their activity in the province of Alicante (Spain). They were established in 1990 and are fundamentally engaged in providing the most efficient tax management service as possible based on the available resources which are becoming increasingly scarce.

In addition to the effect of globalisation, another factor to be considered when evaluating these tax offices is the type of activity that they develop. As service providers they offer a product with intrinsic characteristics (intangibility and heterogeneity) that complicate the evaluation of the efficiency and productivity of their production process (McLaughlin & Coffey, 1990; Parsons, 1997).

The aim of this paper is analyses the productivity growth of the 30 SUMA tax offices located in Alicante (Spain) evolved between 2004 and 2006 by using Data Envelopment Analysis (DEA) models and Malmquist Index. Additionally, a smoothed bootstrap technique is used to provide confidence intervals, and a Mann Whitney U test is used to study the possible effect that certain variables (in particular, population and the number of municipalities within the area of influence of each office) could have on the productivity of the offices.

This study is organized as follows: Section 2 presents a literature review of previous studies will be conducted in order to reflect the current state of research in the field and to support the subsequent selection of the analysis model and variables. Section 3 presents the statistical model will be justified and described. The data used in the study will be presented and the results obtained will be expounded and discussed in Section 4, and finally in Section 5, the conclusions will suggest the main ideas that could be implemented in order to improve the efficiency and productivity of the SUMA tax offices.

2. Literature review

The main aim of this section is to obtain the necessary information to determine the way in which the analysis will be undertaken. Firstly, a review of the studies carried out to date evaluating the efficiency of tax offices will be performed. After, conclusions will be drawn from the analysis model studied and the variables used which will be useful in selecting the most appropriate variables for the specific case of this study.

The earliest of those studies was the article written by González and Miles (2000), who analysed the technical efficiency of 15 Regional Inspection Units of the Spanish Central Tax Authority in 1995 using an input-oriented BCC DEA and bootstrap technique continuing the work of Simar and Wilson (1998) and Löthgren and Tambour (1999). The function of this Authority is the management of the Spanish central tax system and customs services which is fundamental in ensuring the collection of the funds calculated in the Spanish Government’s budget. In this context, the objective of the Regional Inspection Units of the Authority is to prevent tax evasion and fraud by large taxpayers. In order to conduct their study, the authors used the percentage of inspectors over the total personnel of each unit (work factor) as an input. They did not consider other resources such as investments or current expenditure due to the intensive nature of the work of the units. With respect to outputs, the authors used the number of inspections per taxpayer in the area (proxy of the volume of relative effort) and the volume of recovered fraudulent debt per open proceeding in relation to the GDP of the area of jurisdiction (proxy of the result obtained corrected by the level of wealth of the area). The principal results obtained based on DEA only identified four centres with an efficiency level higher than 75%. However, in an analysis carried out using a bootstrap technique the average levels of efficiency of all the units revealed levels between 80% and 90% with no significant differences in efficiency between them.

Two years later, Moesen and Persoons (2002) conducted an analysis of the productive efficiency of 289 regional tax offices belonging to the Finance Ministry in Belgium during 1991. This time, the study was performed using two alternative parametric methods: DEA (under variable and constant returns to scale) and FDH (Free Disposal Hull) whose results were subjected to a sensitivity analysis to the outliers (Belsley, Kuh & Welsch, 1980). The inputs employed in the study were the number of full-time equivalent employees and the outputs were the number of audited returns of category A (wage-earners) and B (independent professionals) and the number of audited returns of category A an B that lead to an increase in the tax base.

In addition to the efficiency analysis, Moesen and Persoons (2002) also performed a Tobit censored regression to try to explain the differences between the results of the different offices. These disparities were explained by circumstances such as the presence of a central tax office that automatically handles aspects related to tax files, the position of a highly qualified manager, the daily zeal of the office and finally, its scale.

In the same line of research, Barros (2005) analysed the efficiency levels of 41 tax offices in Lisbon (Portugal) from 1999 to 2002 by using a Cobb-Douglas cost frontier model. Here the author chose a parametric model for the study instead of a non-parametric model. He used the price of work (average wage earned per assimilated full-time workers), total personal taxes divided by the population in the office area and the ratio of rents of the offices in the area as inputs. The outputs used were the total tax collected at the constant 1999 price and the total clear-up rate of actions (mainly legal and administrative disputes and executive actions). The main findings reveal that in each year the average efficiency level was approximately 80% (diminishing slightly towards the end of the period) and that the proximity of the office to an urban area had a positive impact on its efficiency level. Furthermore, the author identifies factors drawn from the literature as possible causes of inefficiency that may be applicable to the units analysed, such as the rigidity of the labour market when job tenure is not linked to performance, the difficulty of controlling managers that act on behalf of the government, the existence of asymmetric information among different offices and, finally, the x-inefficiency typical of public sectors; due to a lack of professional incentives, scarce competitive pressure and a tendency to prioritise the personal satisfaction of workers over that of the company (Albi, 1992; Leibenstein, 1966)

After Spain, Belgium and Portugal it was Norway's turn to have the efficiency of its tax offices analysed. Forsund, Kittelsen, Lindseth and Edvarsen (2006) studied the performance of 98 local tax offices in Norway from 2002 to 2004 by applying an output-oriented BCC DEA, Malmquist index and smoothed bootstrap. The input variables considered were: the cost of resources (such as manpower, offices and current expenses) adjusted for compensating special circumstances (for example, rent and travel costs) in order to control non-discretionary variables such as population, price levels, area or population density. The outputs used were: people relocated each year, false registrations, employees and pensioners' tax returns, complaints regarding tax assessments, returns from non-incorporated business and corporate taxes. Based on their findings the authors recommend that the small offices could improve their productivity by increasing their size and the government should be concerned with at least one third of the units. However, they do not provide details regarding the way to achieve this objective, leaving it open for further research.

Barely two years after completing his first study on the efficiency of the Portuguese offices, Barros (2007) complemented his original work by conducting an analysis of the technical and allocative efficiency of the same units through non-parametric models. More precisely, he used input-oriented CCR and BBC DEA models and, in the second stage, a Tobit regression to identify the possible explanatory variables involved. As inputs of the DEA models he used the number of employees, rents paid by the premises and the population, average salary, rents on premises divided by the area of those premises and personal taxes per capita (with all the monetary variables at the constant 2000 price). As outputs the author used the total amount of income tax collected, VAT, value of inheritance, donations and other taxes and the clear-up rate of contested cases.
In terms of efficiency levels, the offices produced good results given that in the worst of cases the average efficiency level was 81.6%. The author found that the factors that could help to improve those levels that were statistically and positively significant were the urban location of the offices, the municipal expenditure and the level of GDP of their area. On the other hand, the factor that had a negative effect was the level of salaries as a ratio of total costs.

Subsequently, the field of study returned to Spain although this time the units analysed were different to the previous one. Fuentes (2008) studied the behaviour of the efficiency and productivity of 32 SUMA tax offices in Alicante (Spain) between 2004 and 2006 using output-oriented BCC and CCR DEA, Malmquist Productivity and modified quasi-Malmquist indices. The inputs used in the analysis were the area of each unit and its number of employees. The outputs were the number of tax returns and number of taxpayers dealt with. The most significant conclusions drawn from the results focused on the idea that efficiency could be improved by adopting measures to generate team spirit and encourage responsibility and professionalism. Likewise, improving employees' skills, increasing their interest in taking part in business improvement and involving better technology could also be effective.

Two years later, Katharaki and Tsakas (2010) evaluated the efficiency of 27 tax offices in Greece during the period 2001-2006 by using output-oriented CCR and BBC DEA, DEA window analysis and also Tobit regression in order to explain non-discretionary factors. As inputs the authors considered the number of employees and computers, and the number of people and legal entities paying taxes. The outputs used were the taxation funds related to the number of people and legal entities. The authors' main findings revealed a good level of scale of the majority of the units. However, human resources, the technical infrastructure and the increase in the level of taxation funds were the main factors that were identified as needing improvement. In addition, from a dynamic point of view, the windows analysis indicated a high level of stability in the results as well as their slight improvement. Finally, the population of the area, its predominance of services, the relatively low importance of legal entities and a higher level of GDP in the region were factors that would improve the level of tax offices' efficiency.

In January 2013 once again Spain's tax authority was the focus of attention. This time, Barrilao and Villar (2013) analysed 14 of the 17 Special Tax Offices in Spain, including the regional offices of the Autonomous Communities into which the tax administration is divided, in order to make it more accessible to taxpayers. The model used was a DEA oriented to scale output and variable performance, while the total level of efficiency was broken down into pure technical efficiency (PTE) and scale efficiency (SE). The variables used were the settlement, acts, income and output and goods and service expenditure, with number of declarations and number of employees as inputs. As a result it was found that the Castilla-La Mancha tax office was the most efficient and that it differed significantly from the rest, with the total inefficiency level 31.5%.

In the same year as the previous work Ryu and Lee (2013), evaluated the efficiency of 14 tax jurisdictions in Korea from 1998-2011, using windows-DEA oriented to input. As variables they included the total amount of tax collected as output (differentiating between indirect and direct taxes) and as inputs the authors included the number of regional tax payers (distinguishing between the number of direct and indirect tax payers) and the regional gross domestic product of each jurisdiction. The results indicated an average inefficiency level of 38% while an increase of inefficiency was noted over the whole period with the Seoul and Gwangiu jurisdictions being the least inefficient.

Continuing their 2010 study, Tsakas and Katharaki (2014) examined the performance of tax authorities in Greece, based on data obtained from a sample of 35 tax offices in the period from 2001-2006. Performance was assessed using the DEA method (BBC output-oriented model) and bootstrap. As outputs they considered the incoming taxation revenues related to natural persons and the incoming revenues related to legal entities, and as inputs the number of employees in each office, the number of computers, the number of persons and legal entities paying taxes. A Censored Tobit model was used to analyse the effect of specific environmental variables which might affect the efficiency and performance of the offices. As a result it was found that the greater the number of legal entities the lower the efficiency levels of the tax authorities studied due to their taking advantage of loopholes in the existing legal framework which enabled them to evade paying tax. As a general conclusion it was recommended that the structure within the operating framework of these offices needed to be improved in order to increase the efficiency of their organization, while indicating the need to implement measures to reduce tax evasion.

Finally, Alm and Ducan (2014) present an international comparison of the efficiency of tax offices in 28 different OECD (Organisation Of Economic Cooperation and Development) countries from 2007-2011.
The authors used a three stage method combining both input-oriented DEA and stochastic frontier analysis (SFA). As inputs, the authors used the wage level technology administrative cost and as outputs, el total tax revenues, corporate income tax (CIT), the level of personal income tax (PIT) and value added tax revenues (VAT). As a result the authors found that in 13 of the 28 countries the agencies are relatively efficient with a high average performance level (from 0.838 to 0.904). However, in a subsequent phase, when non OCDE member countreis were added to the analysis, this level was reduced.

The review of the studies performed to date provides information about the different methods and variables which have been used in the past in order to conduct efficiency analyses of tax offices. In this sense, the conclusions that may be obtained from this review will be useful when determining the type of statistical model and variables to use. However, it is not only necessary to take into account the models and variables already used in similar studies but it is also fundamental to consider the availability of the data and the opinion of the experts who are in daily contact with the sector (Barros, 2007).

It is observed that the models used have been eminently non-parametric as a stochastic frontier analysis was used on only two occasions (Barros, 2005; Alm & Duncan, 2014) and the kind of efficiency that has been assessed has mainly been technical efficiency vs allocative or overall efficiency. With regards to the variables involved, those that have been most used as inputs are labour and capital, represented by the best available data that fit. As for outputs, the variables have tried to include the necessary information regarding the results of each tax office related to the amount of taxation funds obtained and/or the people dealt with. When interviewed, senior managers agreed that the same types of variables were appropriate.

These conclusions about the model and variables based on the literature review, professional experience and data availability will be taken into account in the following sections in order to specify the way in which the study of the SUMA tax offices will be undertaken.

3. Methodology

Given that the majority of studies in this field until now have used DEA in order to analyse the evolution of the productivity of tax offices it was considered appropriate to use the same type of method for the current study.

Therefore, the productivity analysis of the agencies has been undertaken using DEA and the calculation of the DEA-based Malmquist Productivity Index (Malmquist, 1953).

DEA allows the units analysed to be organised into a hierarchy in terms of efficiency levels, whilst the Malmquist index allows changes in productivity to be estimated dynamically.

In terms of the output-oriented DEA evaluation process, a decision-making unit (DMU) is considered to be efficient when it obtains the maximum output empirically observable from any examined DMU given its input vector (Charnes, Cooper, & Rhodes, 1981). In other words, a DMU is inefficient when it cannot generate maximum output levels with minimal input consumption (Cooper, Sieford, & Zhu, 2004).

In contrast to parametric methods which are based on the estimation of the hyperplane best adjusted to the set of observations, DEA is a non-parametric method based on the estimation of an efficiency frontier pursuant to the Pareto criterion with the advantage that it avoids imposing a specific form of the production function (Charnes, Cooper, Lewin, & Seiford, 1997). Furthermore, as DEA is not a stochastic method either, it does not assume that the non-calculated efficiency follows some kind of probabilistic distribution (Charnes et al., 1997). In addition, DEA may be used to evaluate efficiency levels in sectors using diverse inputs and outputs in their productive process which do not require information on the importance of each variable in the evaluation process and may also involve the use of non-discretionary variables (Banker & Morey, 1986; Wöber & Fesenmaier, 2004). Finally, DEA also has advantages in situations where the price of resources and products are not known or are difficult to calculate, as is the case of the public sector, in which frequently outputs do not have a market price, as this knowledge is not required to estimate efficiency levels (Charnes et al., 1997).

Notwithstanding this fact, the method is not without disadvantages. Firstly, the institutions analysed need to be homogeneous, that is, that they use the same type of inputs to generate the same class of products and the context in which they carry out their activity should also be similar (Cooper, Seiford & Tone, 2007, chaps. 1 & 4) and, in addition, the reliability of results depends on the number of variables included and the amount of units considered in the study. In this respect Cooper et al., (2007, chaps. 1 & 4) recommend that the number of units should be at least the max $\{(s\cdot r), 3\cdot(s+r)\}$ (where s and r are the number of outputs and inputs, respectively), otherwise the hierarchisation based on the levels of efficiency of the units to be evaluated could be questionable due to the inadequate number of degrees of freedom of the model. Apart from the foregoing, the use of DEA requires special care when selecting the variables to be included as there are no adequate tests for estimating whether the results of the analysis are stable or if they would significantly vary with the use of other types of variables. This requires a careful choice of variables making an exhaustive preliminary review of the existing literature on the theme (Barros, 2005). Finally, the DEA estimates are based on finite samples, which means that the results may vary if the samples are changed (Simar & Wilson, 1998).

All these disadvantages were considered and resolved in this work by choosing perfectly homogenous units, complying with the recommendation to relate the number of units with that of the variables (max {(s·r), 3·(s+ r)}, carrying out an exhaustive review of the earlier literature and applying a procedure for preventing the results from varying with the sample (smoothed bootstrap). This last method, moreover, also eliminates the disadvantage being unable to offer information on the uncertainty of the estimations given that DEA is not stochastic (Löthgren and Tambour, 1999), as will be addressed below in this section.

Taking into account the foregoing, a DEA-based measure of any change in the unit's productivity over time will be calculated using the Malmquist Productivity Index (M) (Malmquist, 1953), in accordance with Färe, Grosskopf, and Lovell (1994).

The Malmquist index has several advantages over other indices frequently used to calculate productivity (such as the Törnqvist or Fisher indices). Firstly, it is calculated only on the basis of the quantitative data from inputs and outputs without the need for pricing information. Secondly, there is no need to assume an approach for maximising output or minimising input. Finally it is able to provide a breakdown of the change in productivity and therefore of the different sources which have led to the change (Grifell-Tatjé & Lovell, 1996). Other advantages include the fact that it does not use fixed weighting when adding inputs and outputs and it does not require standardised measurement units for the variables involved in its calculation (Färe, Grosskopf & Russell, 1998)**.**

However, it does have some disadvantages in that it cannot be calculated for an isolated unit as panel data is required for its calculation (Krüger, 2003; Coelli, Rao, O´Donnell & Battese, 2005) and, in addition, it involves calculation of distance function values (Sufian, 2007). However, these disadvantages will be avoided by using a panel database to obtain the results and the distance functions will be calculated using DEA as explained below.

A generic output-oriented distance function can be defined as:

$$D_t(Xt, Yt) = \sup\{\theta \in R : (Xt, Yt * \theta) \in Pt\}$$

where θ represents the highest factor by which the output vector in year t can be increased when the input vector and the technology for year t is utilised; Y is a vector of outputs; X is a vector of inputs; and Pt represents the feasible production set given the technology in period t, which is defined as:

$$Pt = \{Yt : Xt \text{ can produce } Yt\}$$

With regard to the above, the output-oriented Malmquist index (M) between time periods t and t+1 would be defined as:

$$M_{t,\,t+1}(X_{t+1}, Y_{t+1,}\, X_t, Y_t) = \left[\frac{D_t(X_{t+1}, Y_{t+1})}{D_t(X_t, Y_t)}\,\frac{D_{t+1}(X_{t+1}, Y_{t+1})}{D_{t+1}(X_t, Y_t)}\right]^{(1/2)} \quad (1)$$

As previously mentioned, this index can initially be broken down into two components: technological change (T) and technical efficiency change (E). The breakdown is as follows:

$$M_{t,t+1}(X_{t+1},Y_{t+1},X_t,Y_t)=\underbrace{\left[\frac{D_{t+1}(X_{t+1},Y_{t+1})}{D_t(X_t,Y_t)}\right]}_{E}\cdot\underbrace{\left[\frac{D_t(X_{t+1},Y_{t+1})}{D_{t+1}(X_{t+1},Y_{t+1})}\frac{D_t(X_t,Y_t)}{D_{t+1}(X_t,Y_t)}\right]^{(1/2)}}_{TC} \quad (2)$$

The first ratio (E) represents changes in technical efficiency between two periods (t and t+1). The second ratio (TC) is a measure of technological progress between the same evaluated periods.

The four different distances shown in equation (2) can be achieved using mathematical programming. In particular, an input oriented approach is used to estimate the Malmquist Productivity Index because the principal objective of the SUMA offices in Alicante is to obtain the maximum level of outputs given a specific input vector and technology.

$$\begin{aligned}
\left[D_{0t}(X_{t+1},Y_{t+1})\right]^{-1} &= \max_{\theta\lambda}\ \theta \\
\text{s.a.}\ \sum_{k=1}^{K}\lambda^{k,t}\cdot X_r^{k,t} &\le X_r^{k',t+1}, \forall r \\
\sum_{k=1}^{K}\lambda^{k,t}\cdot Y_s^{k,t} &\ge \theta\ Y_s^{k',t+1}, \forall s \\
\lambda^{k,t} &\ge 0, \forall k
\end{aligned} \quad (3)$$

where θ denotes an efficiency score for a particular DMU ($DMU_{k'}$ with k:1…K - the sub-index k´ shall be used to name the DMU under analysis-); $X_r^{k,t}$, represents the rth input respectively observed at DMUk in year t (with t:1…T); $Y_s^{k,t}$, is the sth output respectively observed at DMUk in year t; and $\lambda^{k,t}$, is a coefficient that shows the proportion of DMU_k used to evaluate $DMU_{k'}$ in year t.

However, as mentioned above, the fact that the non-parametric DEA estimators are based on a finite sample of observations, which renders them susceptible to variations in the sample values, necessitates the use of a method capable of analysing the sensitivity of the productivity results in accordance with changes in the data (Simar & Wilson, 1998). Furthermore, as DEA does not incorporate any randomness in the process, it cannot offer any information with respect to the uncertainty in the estimates of the efficiency of each unit (Löthgren and Tambour, 1999). The bootstrap is a statistical procedure capable of eliminating these two inconveniences of DEA.

This technique was introduced by Efron (1979) and is based on the idea of simulating the data-generating process (DGP) in order to obtain a new estimate of each simulated sample. In this way, the estimates obtained would mimic the distribution of the real population estimator (Simar and Wilson, 1998). For example, it is possible to obtain confidence intervals for the estimates of the efficiency parameters enabling us to determine whether the efficiency levels of the DMUs initially obtained by DEA are statistically significant (Tortosa-Ausina, Grifell-Tatjé, Armero & Conesa, 2008; Fuentes, 2011; Fuentes & Álvarez-Suárez, 2011).

This study will follow the method described by Simar and Wilson (1999) (smoothed bootstrap). This method improves the estimates obtained when we resample directly from the original data, as this procedure (naive bootstrap) provides a poor estimate of the DGP. Furthermore, it incorporates the reflection method described by Silverman (1986), which avoids estimate problems derived from the fact that in the input-oriented model, the efficiency parameters have an upper limit equal to one.

For the DEA approach, the smoothed bootstrap algorithm follows the steps described below (Simar & Wilson, 1999):

1. Compute the Malmquist productivity index $\widehat{M}_{t,\,t+1}(X_{t+1}, Y_{t+1}, X_t, Y_t)$ for each DMU by solving the linear programming models (3) to obtain each of the necessary factors that are shown in (2) ($\widehat{E}$ and $\widehat{T}$).
2. Obtain a pseudo dataset (X^*_t, Y^*_t) for each DMU and t to construct the reference bootstrap technology by using bivariate kernel density estimation and the reflection method.
3. Calculate the bootstrap estimate of the Malmquist index for each DMU $\widehat{M}^{*kb}_{t,t+1}(X_{t+1}, Y_{t+1}, X_t, Y_t)$ by using the sample obtained in step 2.
4. Repeat steps 2–3 B times to obtain a set of estimates $\widehat{M}^{*kb}_{t,t+1}(X_{t+1}, Y_{t+1}, X_t, Y_t)$. Simar and Wilson (2000) recommend a value of B = 2000.

5. Obtain confidence intervals for the Malmquist index and its components after the first 2000 estimates have been obtained from the pseudo-samples generated.

Lastly, as well as the Mamquist and smoothed bootstrap techniques, another statistical method will be used in order to analyse the influence of specific context variables on the productivity of the offices.

This method, the Mann-Whitney U-Test, is based on the idea that the relationship that could potentially exist between two variables may be revealed when their values are organised in increasing order as in this way their values may offer information regarding the relationship between their populations (Gibbons & Chakraborti, 1992, chap. 7).

The use of this method has been chosen for this study for two reasons. Firstly, as with the Malmquist technique, it is non-parametric. And secondly, as there is no reason to assume the existence of any type of underlying probability distribution in the efficiency levels or the variables whose potential relationship is being analysed, the conditions are ideal for this model to be effective in analysing the hypothesis (Sheskin, 2000). Furthermore, this test is more powerful than other non-parametric alternatives such as the Sign Test (Conover, 1999, chap. 5) and it has also been used in previous studies with similar objectives (Fuentes, 2011; Fuentes & Álverez-Suárez, 2010 or Köksal & Aksu, 2007).

The test is based on the calculation of a statistic usually called U which is:

$$Ui = R - \frac{n\,(n+1)}{2} \qquad (4)$$

where R is the sum of the ranges of sample 1 and n is the size of this sample. The existence of a relationship between the samples analysed is rejected with a level of significance α when the value of U is lower than its percentile α/2 or when it is higher than its percentile 1- α/2. It is accepted in all other cases (Conover, 1999).

However, this test is not without limitations:

a) When a set of data is transformed into a range, part of the information is lost. However, if the original data are important as with an ordinal comparison this is not a problem, which is the case with Malmquist estimates (Conover, 1999).
b) Non-parametric tests are only designed to test statistical hypotheses, not to estimate parameters. Nevertheless, this point is not a problem because the aim is to elucidate the existence of a relationship between the productivity of the agencies and the variables studied.

In any of these cases, this work attempts to improve the previously used methods, in that it is the first to estimate productivity levels using bootstrapping DEA-based Malmquist index, while at the same time evaluating the influence of context variables on those levels using non parametric methods, thus maintaining consistency in the use of DEA.

The results of implementing the three methods that have been explained in this section (output-oriented DEA Malmquist productivity indices, smoothed bootstrap and Mann Whitney U Test), are reported and discussed below.

4. Data and results

Given the fact that this study aims to continue and improve the previous analysis carried out by Fuentes (2008), the same set of data and period ('04 -'06) were the first references to be taken into account to conduct the analysis of each of the 30 offices that were available. Nevertheless, this decision was additionally related to the type of variables previously used by other studies in the field that are referred to in section 2. Furthermore, the final decision concerning which variables to use was also based on the experience and knowledge of the management teams at the regional government office and the pragmatic approach required when analysing these statistics (Barros, 2005).

Essentially, the inputs deemed most suitable for analysing the efficiency and productivity of the units were the same as were used in Fuentes (2008). Table 1 shows a statistical summary of the data: the number of workers per office and the area (square meters) of the offices (the latter as a proxy variable of each unit's fixed costs and provision of equipment). Furthermore, the chosen outputs were the number of tax returns and the number of taxpayers to whom services were provided at each office.

Table 1

By using the previous data, table 2 shows the results of the Malmquist productivity indices (M) for the '04-'05 period (fourth column) for each of the offices and the breakdown of the index into its components (other columns). In this period, 21 of the 30 offices have a Malmquist index that is higher than one which means that only nine of them did not improve their level of productivity. Moreover, the average M level for the period is 1.06396, which indicates an average productivity increase of 6.39% in this period.

Table 2

M can be broken down into two factors. The first one (E) is shown in the second column of Table 2 and is called the efficiency change index (E). The fact that the average level for the period reaches a value of 1.0160 implies that the average improvement in efficiency is 1.6 % for that period.

The third column in Table 2 refers to technological change (TC), which is the geometric mean of change in technology between 2004 and 2005. A value of TC that is greater than one implies a technological innovation process. The fact that the average value for the period was 1.0471 reflects that some measures have been taken by management in this respect (improvement of 4.71%). This is also reflected by the fact that the number of individual offices where technological innovation improved (TC > 1 = 23) was substantially higher than those whose efficiency change evolved positively (E > 1 = 12).

Table 3 contains the same information as above but in relation to the '05-'06 period. The first thing that can be noticed is that in the later period, the improvement in productivity was 5.09%, representing a lower value with respect to the previous period. However, it is still a quite high level of progress.

Table 3

On an aggregate level, we can also see that the other values of M's components have attained good levels of improvement. In fact, they are greater than one and, except in terms of technological change (TC), higher than they were in the previous period. Specifically, TC decreases by -3.12 % with respect to the '04-'05 period which means that there was a technological improvement between the evaluated years ('05-'06) but it was lower than before.

Additionally, the number of specific units that improved their productivity due to one effect or the other were similar (TC > 1 = 17 and E > 1 = 16) which indicates that TC was more widespread than E.

The individual results regarding the intensity and extension of the effects lead to the conclusion that the improvement in the levels of M was due largely to the evolution of E and therefore the final positive evolution of M could be mostly explained by an improvement in the specific aspects of management linked to the way in which inputs and outputs were used.

Moreover, it can be observed that the number of units with M values lower than or equal to one has not been a key factor in reducing the average improvement in the period '05-'06, as almost the same number of units (10 instead of 9) have equal or lower levels of productivity in both periods (Tables 2 and 3). In fact, again, the figures corroborate the previous idea that it was the way in which they performed (E effect) which caused the real difference because almost the same number of offices improved their productivity (20 instead of 21) and the E effect was the main reason to explain this evolution.

Nevertheless, from a global perspective (table 4), total average productivity increased by 5.73% with a simultaneous improvement in all components and TC was the main reason to explain the overall positive evolution of this productivity, in spite of the reduction of its value from one period to the other (TC = 1.0306 > E = 1.0259 - see table 4-).

Table 4

To sum up, during the first period (04-05) (Table 2), both the technological (TC) and the efficiency change (E) improved but TC was the most important factor to explain the evolution of the productivity because of its higher mean value (1.0471>1.016), together with the greater quantity of offices whose productivity increased due to its influence (23 vs. 12). However, during the second period (05-06) (table 3), although the number of offices whose productivity increased due to the effect of efficiency (E) vs. technological change (TC) was almost the same (16 vs. 17), the fact that the average value of E was considerably higher than the average value of TC (1.0359> 1.0144) illustrates that their productivity rose because of the better use of resources made by the offices (E effect). All in all, considering the whole period of time (2004-2006) (table 4), the highly increase in the level of TC in the first stage was large enough to compensate its lower influence in the second phase and, finally, TC turned out to be the most important factor to be taken into account (TC =1.0306 > E=1.0259).

In any event, the afore-mentioned results were based on output-oriented Malmquist indices calculated with information referring to the years 2004, 2005 and 2006 and, therefore, these results and their conclusions would be subject to variations in the samples. In order to eliminate this effect of sample variability, smooth bootstrap has been applied in this study (data were analysed with FEAR 1.15 package, running on R 2.12.2). In this way, the conclusions that may be drawn from the results would not be influenced by random sample variations, rendering them more consistent.

The bootstrap methodology enables us to contrast the null hypothesis of no productivity change by calculating confidence intervals. The figures represented in bold in Tables 2 and 3 (95% confidence) indicate that the null hypothesis is rejected and therefore it is accepted that changes occurred. In other words, only the figures in bold are statistically significant and therefore, only those figures can show the existence of a real change in the level of productivity of the corresponding offices. In consequence, it will be necessary to take into account the fact that different conclusions could be derived from both types of results and therefore, both should be commented and compared.

Table 5 shows the mean values of the two types of effects in which the change in the level of productivity can be divided (E and TC) obtained from the statistically significant values (figures in bold) of tables 2 and 3.

Table 5

From the point of view of these values, in the first period, both the technical (TC) and efficiency change (E) generated similar results (1.0987 vs. 1.0865). However, during the second period, there was a reduction in the influence of both effects meaning that the productivity enhancement was lower. Even so, the decrease in the E effect was much lower than it was for the TC effect (- 0.98% vs - 6.96%).
On the other hand, with respect to the number of individual offices with improved effects that were statistically significant, the predominance of TC was very clear during the first period (TC > 1 for 18 units vs. E > 1 for only 8 - see bold figures in table 3 -). Nevertheless, the situation changed during the second period since the number of DMUs influenced by the TC effect was similar (TC > 1 for 10 vs. E >1 for only 9 (see bold figures in table 3)), due to the strong decrease in the offices' productivity altered by TC.

In short, according to the afore-mentioned results related to mean values and influenced DMUs, efficiency change (E) was the most constant in periods while technological change (TC) suffered a sharp decline, becoming the less important of the two effects during the second period.

The previous conclusion obtained from the aggregate data is accentuated by the behaviour of the significant results for the most efficient offices. Thus, during the first period (04-05) only 13 offices improved their productivity, with the Altea, Villena, Onil, Orihuela, Alfaz del Pí and Pilar de la Horadada units experiencing a greater increase although only the first four were able to maintain high and significant levels when M was divided into its two components E and TC (Table 2). Furthermore, during the period (05 -06) only 9 offices significantly improved productivity with the biggest improvements in the offices of Altea, Santa Pola, Ibi, Elda, Benidorm and Crevillente, which in general did not coincide with those in the previous period. Of these last, the first four also offered high and significant results in their E and TC levels with the last two offices (Benidorm and Crevillente) the only ones of those mentioned which did not offer significant results in any of their components. As may be seen, only the Altea office was common to both, and was thus identified as the best in the group.

Based on the individual behaviour of the offices with the greatest significant productivity in each period, it is possible to observe that it is the component E which has most effect on the final value M when considering the total period 04-06. In particular, among these units, this component also presented a greater specific weight at an individual level than that which it had at aggregated level for the whole group. Thus and for each period, the average value of the component E for the most efficient and significant was from 1.2936 and 1.1279, and for the TC component from 1.1100 and 1.0379 respectively (table 6) that is, they presented differences between both effects which were greater than those offered by the offices overall with significant values (table 5), which attests to the differential effort of management of resources made by the managers of the most efficient offices (E) aside from the effect caused by improvements in equipment (TC).

Table 6

From a contrary perspective, that is, taking into account the individual behaviour of the offices whose levels of productivity significantly worsened, the previous conclusion may also be seen to be valid. In this respect, Benisa during the period (04-05) and

Torrevieja-Mata, Villena and Denia during (05-06) (tables 2 and 3) showed a negative development in behaviour with respect to E, thereby reinforcing the idea that the way in which resources were managed was the differentiating factor affecting productivity in offices. Only two of the four also saw a deterioration in their TC component, but also in one of these two cases (Villena) the E component experienced a greater reduction and therefore, it was of greater importance in the negative development of M.

In addition to the productivity results obtained through the Malmquist and smooth bootstrap techniques, the Mann Whitney U test (IBM SPSS statistics 20 was used) was also carried out in order to establish whether a relationship existed between the productivity of the offices and other context variables such as the population and the number of municipalities that form part of the area in which they operate. In order to obtain reliability in the results, only statistically significant values of M for each period were used in the analysis, generating the results shown in Table 7.

Table 7

As the table above shows, neither the population nor the number of municipalities within the area of influence of each office affect the results of productivity in either of the periods analysed. In this way, on the contrary to what was initially believed, there is no statistical evidence to support that a larger population or a greater number of municipalities served by an office affect its proper functioning despite the fact that the higher the values of these variables the greater the potential workload that each office must undertake.

5. Conclusions

After evaluating the productivity levels of the tax offices and analysing the existence of a potential relationship between these levels and certain environmental variables, population and number of municipalities, this section will extract conclusions from the results obtained that will help to improve the management of these units which in turn will improve their productivity.

The results obtained from analysing the data using the output-oriented Malmquist, reveal an increase in productivity in both periods due to improvements in technology and resource management.

However, the improvement experienced by the technological factor (TC) decreased during the second period, whereas that referring to management (E) increased. This implies that in their attempts to improve productivity during that period, managers favoured the optimum use of resources over an improvement in technology.

On the other hand, the conclusion drawn from the results for the first period (Table 2) is that the main factor that affected the levels of productivity during 04-05 was the level of technology. Both, the number of offices (TC > 1 = 23 vs. E > 1 = 12) and the mean value of the period (T mean = 1.0471 > E mean = 1.016) support this idea.

So, technology was the main factor in the first period and efficiency in the second and, logically, the best strategy should be to improve both effects by carrying out the necessary measures.

In fact, when using a smoothed bootstrap technique (values in bold type), the above conclusion is confirmed in spite of the fact that the results are different. Taking into account the number of offices affected, the factor which should be improved to a greater extent in both periods would be the more efficient use of resources (E) because in both periods the number of offices with an improvement due to TC is greater than the number that have improved due to E (18 as opposed to 8 in 04-05 and 10 as opposed to a 9 in 05-06). However, on average, the bootstrap TC and E improvements worsened from one period to the next but the decrease in E was lower.

Therefore, although more offices increased their productivity due to technological changes, on average, the improvement in management had a greater influence on their enhanced productivity. The same conclusion was reached when individually analysing the behaviour of the offices which were significantly the most or least productive in each period

From this point of view, the implementation of both extensive improvement processes in management and intensive improvement processes in technology would be key factors in increasing performance levels, as concluded from the bootstrap and non-bootstrap results. More specifically, a greater dedication of resources to comprehensive follow-up training in the efficient use of resources and a higher investment in technical equipment (software and hardware) would have a positive effect on productivity.

Once a general outline of the strategic actions to improve productivity had been made based on the analysis of the data, those offices with the highest levels of productivity in each period were interviewed about their operating practices in order to determine more specific strategies. The idea was to extrapolate their experiences to the rest in order to increase the overall results, as suggested by Norman and Stoker (1991).

The responses obtained in the interviews with the managers of these offices revealed that in all cases they highlighted the ideas obtained from the previously analysed information. In particular, suggestions for improving the productivity of the offices were related mainly to the advisability of investing in IT infrastructure that would enable faster connections with the different dependent government bodies and the continuous training of staff. More specifically, the need for greater investment in equipment and IT programs that would facilitate and speed up the tasks of the staff and continuous training courses directed at specialising workers in key positions as opposed to the existing general training were highlighted. In short, they recommended a specialised training of the workers and a segmentation of work in order to increase productivity.

Another idea suggested by the managers of the best-performing offices was the need to involve the workers when making decisions regarding the measures to adopt to improve the management of the office. They were not in favour of simply receiving orders to carry out their tasks but expected to be consulted and contribute their ideas to improving the processes based on their knowledge gained from working directly in the offices. A further suggestion consisted in involving staff in the actions by setting personalised economic targets based on individual performance and the difficulty and recycling of training programmes carried out.

To sum up, based on an analysis of the available information it was observed that improvement strategies should include a greater investment in technology and greater efforts should be taken to increase the efficiency in the use of resources. These suggestions were those that were specifically mentioned by the managers of the best-performing offices and focused on aspects such as the improvement of the IT infrastructure and specialised training, involving and motivating staff through personalised economic objectives.

Regarding the limitations of the study, firstly it should be mentioned that it was impossible to involve any other type of inputs which could have helped to shed light on the conclusions obtained based on the results. In particular, an attempt was made to obtain information on wage costs, the value of properties where the offices were located, and the computer equipment in each one. However, despite repeated requests this information was not provided. In addition, there were context variables, the effect of which it would have been interesting to include in the analysis (such as, for example, the average income level of the contributors in each office or the level of economic activity generated by the companies located in each municipal district in the area where each office operated) but which could not be taken into consideration as this data did not exist at a municipal level.

Finally, possible future lines of research may include the application of other statistical models which would provide new results that could be compared with those obtained in this study. For example, from a non-parametric perspective, the Lumberger productivity indicator or Quasi-Malmquist indices or even a Tobit regression linked to bootstrap (Simar & Wilson, 2007) could be considered. Additionally, other factors that could not be contemplated in this study might exist or become available in the future and be taken into account to obtain new conclusions (such as, local GNP or the economic activity of the local number of firms that may affect the level of efficiency of the offices).

References

Albi, E. (1992). Evaluación de la eficiencia pública. El control de la eficiencia del sector público. *Hacienda Pública Española*, 120(1), 299-319.

Alm, J., & Duncan, D. (2014). Estimating tax agency efficiency. *Public Budgeting & Finance*, 34(3), 92-110.

Banker, R. D., & Morey, R. C. (1986). Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research*, 34(4), 513–521.

Barros, C. P. (2005). Performance measurement in tax offices with a stochastic frontier model. *Journal of Economic Studies*, 32(6), 497-510.

Barros, C. P. (2007). Technical and allocative efficiency of tax offices: a case study. *Int. J. Public Sector Performance Management*,1(1),41-61.

Barrilao, P. E., & Villar, E. (2013). The efficiency of the regional management centres of the tax administration in Spain. *Journal of US-China Public Administration*, 10(1), 49-56.

Belsley, D., Ktth: E. and Welsch, R., 1980, *Regression diagnostics: Identifying influential data and sources of collinearity,* (John Wiley, New York).

Charnes, A., Cooper, W. W. &, Rhodes, E. L. (1981). Evaluating program and managerial efficiency: an application of data envelopment analysis to Program Follow Through. *Management Science*, 27(6), 668-97.

Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1997). *Data Envelopment Analysis: Theory, Methodology and Applications*. New York, Kluwer Academic Publishers, Second edition.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). New York: NY Springer.

Conover, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). New York: John Wiley & Sons.

Cooper, W. W., Sieford, L. M., & Zhu, J. (2004). Data Envelopment Analysis. History, Models and Interpretations in *Handbook of Data Envelopment Analysis,* chap. 1. Kluwer Academic Publishers, London.

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data envelopment analysis. A comprehensive text with models, applications, references and dea-solver software* (2nd ed.). Springer.

Efron, B. (1979). Bootstrap methods: another look at the jacknife. *The Annals of Statistics*, 7(1), 1–26.

Färe, R., Grosskopf, S., & Lovell, C.A.K. (1994). *Production frontiers*, Cambridge University Press.

Färe, R., Grosskopf, S., & Russell, R. R. (1998). *Index numbers: Essays in honour of Sten Malmquist.* New York: Springer.

Forsund, F.R., Kittelsen, S.A.C., Lindseth, F., & Edvarsen, F. (2006), The tax man cometh - but is the efficient? *National Institute Economic Review,* 197,106-119.

Fuentes, R. (2008). Productivity at the SUMA tax offices. in VV.AA., *Proceedings of the 11th Toulon-Verona International Conference on quality in services* (pp. 323-333). Firenze University Press.

Fuentes, R. (2011). Efficiency of travel agencies: A case study of Alicante, Spain. *Tourism Management*, 32, 75-87.

Fuentes, R., & Álvarez-Suárez, A. (2011). Travel agencies in Alicante,Spain: a productivity analysis. *The Total Quality Management Journal*, 23(5), 560-577.

Gibbons, J. D., & Chakraborti, S. (1992). *Nonparametric statistical inference* (3rd ed.). New York: Marcel Dekker, INC.

González, X. M., & Miles, D. (2000), Eficiencia en la inspección de Hacienda, *Revista de Economía Aplicada*, 3(24), 203-219.

González-Páramo, J. M., & Onrubia, F. (2003). Información, evaluación y competencia al servicio de una gestión eficiente de los servicios públicos. *Papeles de Economía Española*, 95, 2-23.

Grifell-Tajté, E., & Lovell, C. A. K. (1996). Deregulation and productivity decline: The case of Spanish saving banks. *European Economic Review*, 40, 1281-1303.

Katharaki, M., & Tsakas, M. (2010). Assesing the efficiency and managing the performance of Greek tax offices. *Journal of Advances in Management Research*, 7(1), 58-75.

Köksal, C. D., & Aksu, A. A. (2007). Efficiency evaluation of A-group travel agencies with data envelopment analysis (DEA): a case study in the Antalya region, Turkey. *Tourism Management*, 28, 830–834.

Krüger, J. J. (2003). The global trends of total factor productivity: Evidence from the nonparametric Malmquist index approach. *Oxford Economic Papers*, 55, 265-286.

Leibenstein, H. (1966). Allocative efficiency versus X-efficiency. *American Economic Review*, 56(3), 392-415.

Löthgren, M., & Tambour, M. (1999). Bootstrapping the data envelopment analysis Mamquist productivity index. *Applied Economics*, 31, 417–425.

McLaughlin, C.P., & Coffey, S. (1990). Measuring productivity on services. *International Journal of Service Industry Management*, 1(1), 46-64.

Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de Estadística*, 4, 209-42.

Moesen, W., & Persoons, A. (2002). Measuring and explaining the productive efficiency of tax offices: a non-parametric best practice frontier approach. *Tijdschrift voor Economie en Management*, 17, 399-416.

Norman, M,, & Stoker, B. (1991), *Envelopment analysis: The assessment of performance*. Essex, UK: John Wiley & Sons.

Parsons, L.J. (1997). Productivity versus relative efficiency in marketing: Past and future. In G. Laurente, G.L. Lilien, & B. Prass (Eds.), *Research traditions in marketing* (pp. 169–198). Boston, MA: Kluwer.

Ruy, S., & Lee, S. (2013). An exploratory study of efficiency in tax jurisdictions. *Advanced Science and Technology Letters*, 34, 46-49.

Sheskin, D. (2000). *Handbook of parametric and nonparametric statistical procedures* (2nd ed.). Chapman & Hall/CRC. (Test 12).

Silverman, B. W. (1986). *Density estimation for statistics and data an*alysis. London: Chapman and Hall.

Simar, L., & Wilson, P.W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44, 49–61.

Simar, L., & Wilson, P.W. (1999). Estimating bootstrapping Malmquist indices. *European Journal of Operational Research*, 115, 459–471.

Simar, L., & Wilson, P.W. (2000). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis*, 13, 49–78.

Simar, L., & Wilson, P.W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136, 31–64.

Sufian, F. (2007). Malmquist indices of productivity change in Malaysian islamic banking industry: Foreign versus domestic banks. *Journal of Economic Cooperation*, 28(1), 115-150.

Tortosa-Ausina, E., Grifell-Tatje´, E., Armero, C., & Conesa, D. (2008). Sensitivity analysis of efficiency and Malmquist productivity indices: an application to Spanish savings Banks. *European Journal of Operational Research*, 184, 1062–1084.

Tsakas, M., & Katharaki, M. (2014). Impact of environmental factors on the efficiency of tax organizations. *Serbian Journal of Management*, 9(1), 31-43.

Wöber, K. W., & Fesenmaier, D. R. (2004). A multi-criteria approach to destination benchmarking: a case study of state tourism advertising programs in the United States. *Journal of Travel & Tourism Marketing*, 16(2/3), 1–18.

Table 1. Characteristics of inputs and outputs.

	Year		2004	2005	2006
OUTPUTS	Number of tax returns	average	93491.73	104355.23	115423.67
		square deviation	52876.44	56072.87	62180.53
	Number of taxpayers to whom services are provided	average	11095.67	12661.20	13695.80
		square deviation	10743.53	10557.31	9612.69
INPUTS	Area	average	192.85	192.85	192.85
		square deviation	90.54	90.54	90.54
	Number of employees	average	6.73	6.93	7.19
		square deviation	3.97	4.02	4.20

Source: Authors

Table 2. Results of M, E and TC for 04-05.

	04-05 period		
DMU	**E**	**TC**	**M**
ORIHUELA	**1.1351**	**1.0679**	**1.2122**
CALLOSA SEGURA	0.9546	**1.1183**	1.0675
ALMORADÍ	**0.9710**	**1.1170**	**1.0847**
TORREVIEJA-MATA	1.0000	1.0000	1.0000
PILAR HORADADA	**1.2822**	**0.9874**	**1.2660**
SANTA POLA	0.9051	**1.1600**	1.0499
GUARDAMAR	1.0000	**1.0525**	1.0525
CREVILLENTE	1.0638	**1.0375**	**1.1037**
ASPE	1.0824	**1.1166**	**1.2086**
ELDA	0.9019	1.0933	0.9860
PETRER	0.8924	**1.1215**	1.0008
NOVELDA	0.9261	**1.1242**	1.0412
VILLENA	**1.1009**	**1.1016**	**1.2128**
ALCOY	0.8410	**1.1201**	**0.9420**
IBI	0.9588	0.9510	**0.9118**
ONIL	**1.3909**	**1.1112**	**1.5456**
EL CAMPELLO	1.0000	1.0354	1.0354
SAN VICENTE	0.9366	**1.0594**	0.9922
SAN JUAN	**1.0210**	**1.0254**	1.0470
VILLAJOYOSA	1.0267	1.0721	**1.1007**
ALFAZ DEL PÍ	**1.2343**	0.9822	**1.2123**
LA NUCIA	**0.9340**	**1.1157**	**1.0420**
ALTEA	**1.6110**	**1.1614**	**1.8710**
BENIDORM	1.0417	0.7056	**0.7350**
DENIA	**1.0358**	0.6225	**0.6448**
PEDREGUER	**0.8517**	1.1941	1.0171
PEGO	**0.9787**	**1.1102**	**1.0865**
CALPE	1.0000	1.0000	1.0000
BENISA	**0.8255**	**1.1899**	**0.9823**

TEULADA	0.9197	1.1162	**1.0266**
Mean	1.0160	1.0471	1.0639

Source: Authors.

Note: The results in bold type represent values that are significantly different from one (95%).

Table 3. Results of M, E and TC for 05-06.

		05-06 period	
DMU	**E**	**TC**	**M**
ORIHUELA	0.9307	**1.0799**	1.0051
CALLOSA SEGURA	1.0475	1.0417	1.0912
ALMORADÍ	**1.1389**	**0.9791**	**1.1151**
TORREVIEJA-MATA	**0.9451**	**0.9116**	**0.8615**
PILAR HORADADA	0.9014	1.0374	0.9352
SANTA POLA	**1.1415**	**1.0273**	**1.1726**
GUARDAMAR	1.0000	**1.0149**	1.0149
CREVILLENTE	1.1641	**0.9881**	**1.1502**
ASPE	1.0808	0.9929	1.0731
ELDA	**1.1006**	**1.0593**	**1.1658**
PETRER	**1.5170**	0.9683	**1.4689**
NOVELDA	0.9450	1.0199	0.9637
VILLENA	**0.9274**	**0.9895**	**0.9177**
ALCOY	0.9838	1.0312	1.0145
IBI	**1.1088**	**1.0376**	**1.1505**
ONIL	1.0312	1.0073	1.0387
EL CAMPELLO	1.0000	1.0000	1.0000
SAN VICENTE	**1.0755**	**0.9712**	1.0446
SAN JUAN	0.9100	1.0258	**0.9334**
VILLAJOYOSA	1.0000	1.0000	1.0000
ALFAZ DEL PÍ	0.8448	1.0437	**0.8817**
LA NUCIA	**0.8670**	**1.1872**	1.0293
ALTEA	**1.1619**	**1.0277**	**1.1941**
BENIDORM	1.0919	**1.0606**	**1.1582**
DENIA	**0.8322**	**1.0776**	**0.8968**
PEDREGUER	1.1077	**1.0518**	1.1651
PEGO	1.1296	**0.9386**	1.0603
CALPE	1.0000	1.0000	1.0000
BENISA	**1.2173**	0.9339	**1.1368**
TEULADA	**1.1175**	0.9678	1.0815
Mean	1.0359	1.0144	1.0509

Source: Authors.

Note: The results in bold type represent values that are significantly different from one (95%).

Table 4. Aggregated development of M, E and TC for 04-06.

PERIOD	E	TC	M
04-05	1.0160	1.0471	1.0639
05-06	1.0359	1.0144	1.0509
Mean	1.0259	1.0306	1.0573

Source: Authors

Table 5. Mean values of E and TC from statistically significant results.

Period	Mean TC	Mean E	Mean M
04-05	1.0987	1.0865	1.1162
05-06	1.0222	1.0758	1.0860

Source: Authors

Table 6. Average E and TC values for the significant most productive offices in each period.

Period	Mean TC	Mean E
04-05	1.1100	1.2936
05-06	1.0379	1.1279

Source: Authors

Table 7. Results of the Mann Whitney U Test based on the significant results of the bootstrap test for each period.

Null hypothesis	Test	Asymptotic significance (bilateral test) 04.05/05.06	Decision
The population distribution between 04.05 and 05.06 is the same for both efficient and inefficient offices	Mann Whitney U Test	0.127/0.463	The null hypothesis is confirmed (population had no influence on productivity in the periods 04.05/05.06)
The distribution of the number of municipalities per office during the period between 04.05 and 05.06 is the same for both the efficient and inefficient offices	Mann Whitney U Test	1.0000/0.835	The null hypothesis is confirmed (the number of municipalities had no influence on productivity in the periods 04.05/05.06)

The table shows asymptotic significances. The level of significance is 0.05

Source: Authors

Hightligths

-We study productivity of tax offices using DEA-based Malmquist productivity index
-In addition a smoothed bootstrap technique is used to provide confidence intervals
-A Mann Whitney U test is utilized to study the influence of specific context variables
-No statistical evidence that population or number of municipalities affect productivity
-Results reveal an increase in productivity due to technology and resource management