<u>INSTITUT DE STATISTIQUE</u> <u>BIOSTATISTIQUE ET</u> <u>SCIENCES ACTUARIELLES</u> <u>(ISBA)</u>

UNIVERSITÉ CATHOLIQUE DE LOUVAIN



DISCUSSION PAPER

2013/01

Local government efficiency: The case of Moroccan municipalities

EL MEHDI, R. and C. HAFNER

Local government efficiency: The case of Moroccan municipalities

Rachida El Mehdi^{*} Christian Hafner[†]

January 4, 2013

Abstract

In times of tight public budgets and deficits, the topic of measuring local government efficiency is increasingly becoming important, both for policy makers in the central governments, as for researchers working on the methodology. In this work we focus on measuring the efficiency of Moroccan municipalities in terms of their financial autonomy. We use traditional nonparametric approaches, the Data Envelopment Analysis (DEA) and the Free Disposal Hull (FDH), combined with bias corrections using the bootstrap, which also allows us to construct confidence intervals for the estimated efficiencies, and test for the returns to scale. Our results indicate that very few municipalities are efficient or close to the frontier. Both DEA and FDH efficiency scores indicate that there is a negative relation between population size and efficiency scores, which is unlike previous studies for other countries.

JEL classification: C1; C44; C6

Keywords: Bootstrap; Efficiency; Data Envelopment Analysis; Returns To Scale; Financial Autonomy.

^{*}Mohammed First University of Oujda and Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain

[†]CORE and Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Voie du Roman Pays 20, 1348 Louvain-la-Neuve, Belgium; christian.hafner@uclouvain.be; corresponding author

1 Introduction

Analysis of local government efficiency has attracted considerable interest over recent years. In times of scarce public budgets and increasing public debt, the efficiency of local administration of decentralized budgets becomes important also for the central government. Examples of recent studies are Worthington (2000) for Australia, Afonso and Fernandes (2008) for Portugal and Balaguer-Coll et al (2007) for Spain. Afonso and Fernadez (2008) give an excellent account of the literature. In Appendix A, we review the concept of production efficiency introduced by Farrell (1957).

In most studies of local government efficiency, nonparametric approaches of frontier analysis are used, in particular data envelopment analysis (DEA) assuming convexity of the production set, and free disposal hull (FDH) without the convexity assumption. Parallel to applications, progress has been made in understanding the statistical properties of DEA and FDH efficiency estimates, which however has not yet been entirely taken into account in applied research. For example, the use of bias corrected estimators based e.g. on the bootstrap seems to be an important ingredient for reliable interpretation of estimated efficiency scores.

In this paper, we study the management of the financial resources of Moroccan municipalities. We use an aggregate output measure, the financial autonomy, which is economically meaningful and avoids the typical curse of dimensionality of nonparametric estimators in high dimensions. To the best of our knowledge, no results are available in the literature on local government efficiency in Morocco. We try to fill this gap, accounting for recent statistical research in DEA and FDH and comparing the results with those of other countries.

A general result of our study is that very few Moroccan municipalities are close to the frontier according to DEA. By construction, efficiency scores are higher for FDH, and more municipalities are close to the frontier. However, even for FDH, more than 90% of the municipalities are inefficient. Moreover, we find that there is a negative relation between population size and efficiency, both for DEA and FDH. This differs from the results of Balaguer-Coll et al (2007) for Spain.

The remainder of the paper is organized as follows. Section 2 presents the DEA framework, Section 3 describes the statistical approach of using the bootstrap to do inference on the efficiency score estimates, Section 4 presents FDH as an alternative to DEA with less restrictive assumptions, and Section 5 develops the application to our Moroccan data set. Section 6 concludes. We delegate to an appendix the definition of technical and allocative efficiency.

2 The DEA program

Data envelopment analysis (DEA) is a nonparametric programming approach used to determine the production frontier and to estimate technical efficiencies of decision making units (DMU) such as firms or countries. The concept of technical efficiency as introduced by Farrell (1957) has been widely used in empirical research on production efficiency, see the appendix for a short review.

We focus on input-oriented DEA, i.e. the question by how much input quantities can be proportionally reduced without changing the output quantities. This is the typical problem for public decision makers which have to ensure public services while trying to minimize the inputs, see e.g. Daraio and Simar (2007, p.30).

To describe the analytical structure of DEA, the input matrix is denoted by $X_{(n,m)}$ and the output matrix as $Y_{(n,r)}$, where m and r are respectively the number of inputs and outputs, and n is the number of DMU's under study. We suppose that each DMU produces the same r outputs in possibly different amounts using the same m inputs also in possibly different amounts. As in Sueyoshi (1999), we assume that all DMU's have linearly independent input and output vectors in their data domain. The DEA matrix formulation for a given point (X_0, Y_0) in the case of variable returns to scale is given by the following linear programming primal problem, which needs to be solved n times, once for each DMU, see e.g. Banker et al (2004),

$$\min \theta$$
s.t.
$$\begin{cases} \sum_{i=1}^{n} \lambda_i y_{ki} \ge y_{k0}, \quad k = 1, \cdots, r \\ \sum_{i=1}^{n} \lambda_i x_{ji} \le \theta x_{j0}, \quad j = 1, \cdots, m \\ \sum_{i=1}^{n} \lambda_i = 1, \\ \theta \ge 0, \ \lambda_i \ge 0 \quad \forall i = 1, \cdots, n \end{cases}$$
(2.1)

where θ corresponds to the level of technical efficiency, $Y_0 = (y_{10}, \dots, y_{k0}, \dots, y_{r0})$ and $X_0 = (x_{10}, \dots, x_{j0}, \dots, x_{m0})$ are levels of outputs and inputs of the considered DMU. The variables $\lambda_i, i = 1, \dots, n$ insure the convex hull of inputs and outputs in these data spaces. The restriction $\sum_{i=1}^{n} \lambda_i = 1$ corresponds to the assumption of variable returns to scale (VRS). It can be replaced by other assumptions on the returns to scale (RTS), namely $\sum_{i=1}^{n} \lambda_i > 1$ (increasing returns to scale) and $\sum_{i=1}^{n} \lambda_i < 1$ (decreasing returns

to scale). If the restriction on $\sum_{i=1}^{n} \lambda_i$ is dropped, one obtains constant returns to scale (CRS).

Let θ^* denote the optimal level of the efficiency score. In the case of one input and one output, θ_i is a radial measure of the distance between (x_i, y_i) and the corresponding frontier. The DMU is efficient when $\theta^* = 1$ and inefficient in case of $0 \le \theta^* < 1$.

The slack variables S_k^- and S_j^+ associated with the dual variables u_k and v_j respectively lead us to the following program:

$$\min\left(\theta + \sum_{k=1}^{r} S_{k}^{-} + \sum_{j=1}^{m} S_{m}^{+}\right)$$

$$s.t.\begin{cases}\sum_{i=1}^{n} \lambda_{i} y_{ki} - S_{k}^{-} = y_{k0}, & k = 1, \cdots, r\\\sum_{i=1}^{n} \lambda_{i} x_{ji} + S_{j}^{+} = \theta x_{j0}, & j = 1, \cdots, m\\\sum_{i=1}^{n} \lambda_{i} = 1, & \\\theta \ge 0, \ \lambda_{i}, \ S_{k}^{-}, \ S_{j}^{+} \ge 0 & \forall i = 1, \cdots, n, \ \forall k, \ \forall j\end{cases}$$
(2.2)

 DMU_{i_0} is Pareto-efficient or fully efficient if and only if $\theta_{i_0}^* = 1$ and all slacks S_k^{-*} and S_i^{+*} are zero, see e.g. Thanassoulis (2001).

Instead of solving the primal program, it is often easier to use the dual program. In our case, the dual program is given by

$$Max \sum_{k=1}^{r} u_{k}y_{k0} + u^{*}$$

$$s.t. \begin{cases} \sum_{j=1}^{m} v_{j}x_{j0} \leq 1, \\ \sum_{k=1}^{r} u_{k}y_{ki} - \sum_{j=1}^{m} v_{j}x_{ji} + u^{*} \leq 0, \quad i = 1, \cdots, n \\ u_{k}, v_{j} \geq 0 \quad \forall k, j \end{cases}$$
(2.3)

where $U^t = (u_1, \dots, u_k, \dots, u_r) \in \mathbb{R}^r_+$ and $V^t = (v_1, \dots, v_j, \dots, v_m) \in \mathbb{R}^m_+$ are row vectors of the dual variables related to the constraints of the primal problem. Furthermore, a constraint with a strict equality in the primal is replaced by a free (unrestricted) variable $u^* \in \mathbb{R}$ in the dual. Being free, this variable should be replaced by the difference between two positive variables t_1 and t_2 in a linear programming problem solved by the simplex method.

At a point (x_0, y_0) we have $\hat{\theta}(x_0, y_0) = \sum_{k=1}^r \hat{u}_k y_{k0} + \hat{u}^*$.

The sign of the optimal value of u^* is used to identify the type of RTS at a point (x_0^*, y_0^*) on the efficiency frontier. Being negative, zero, positive or free, this variable indicates NIRS, CRS, IRS or VRS respectively.

It is useful to indicate that input and output oriented models may give different results with respect to their returns to scale. Thus, increasing returns to scale may result from an input oriented model, while an application of an output oriented model may produce a decreasing returns to scale for the same data. Also, it is worthwhile to note that working in smaller dimensions tends to provide better estimates of the frontier.

Many software packages include algorithms to solve linear programming problems of the type discussed previously. The problem can be cast in a form treatable by the simplex method used by R.

3 Inference using the bootstrap

The bootstrap is a method which can be useful in many problems of statistical inference such as constructing confidence intervals and hypothesis tests. Its principle is to create a pseudo-replicate data set from the given data set, and then perform statistical inference using the replication set. The use of the bootstrap method in DEA goes back to Simar (1992).

The bootstrap method is based on the idea that the bootstrap distribution will mimic the original unknown sampling distribution of the estimators of interest (efficiencies) using a nonparametric estimate of their densities. Hence, a bootstrap procedure can simulate the data generating process (DGP) by using a Monte Carlo approximation and may provide a reasonable estimator of the true unknown DGP.

3.1 Data Generating Process

Consider a statistical model where a DGP P generates a random sample $\chi = \{(X_i, Y_i)_{i=1}^n\}$ of size n and suppose that we want to investigate the sampling distribution of the estimator $\hat{\theta}$ of an unknown parameter θ .

Using the nonparametric method described in (2.1) it is possible to estimate θ by $\hat{\theta}$ at a fixed point (x, y) for each DMU. As the DGP P is unknown, the bootstrap procedure is used to determine the DGP \hat{P} as an estimator of the true unknown DGP. Thus, since \hat{P} is known, we can generate a data set $\chi^* = \{(X_i^*, Y_i^*), i = 1, \dots, n\}$ from \hat{P} . This pseudo-sample defines the quantities $\hat{\theta}^*$ corresponding to the efficiencies $\hat{\theta}$ at the point (x, y).

Analytically, it may be difficult to compute the true distribution of $\hat{\theta}^*(x, y)$ resulting from a sample χ^* drawn from \hat{P} . Therefore, the Monte Carlo approximation can be employed to obtain the sampling distribution of $\hat{\theta}^*(x, y)$. Using \hat{P} to generate B pseudosamples χ_b^* for $b = 1, \dots, B$ and applying the model (2.1), we obtain a set of pseudo estimates $\left\{\hat{\theta}_b^*(x, y)\right\}_{b=1}^B$. These pseudo estimates give an approximation of the unknown sampling distribution of the efficiency scores $\hat{\theta}_b^*(x, y)$ conditional on \hat{P} .

3.2 Bootstrap correcting bias for DEA efficiency scores

The bootstrap algorithm allows us to obtain bias corrected estimators and to make inference on the DEA efficiency scores. Correcting for the bias introduces additional noise and thus increases the variance of the estimator. However, Daraio and Simar (2007) suggest that a bias correction should be considered in almost all practical situations. However, before defining the bias corrected estimator of DEA efficiency scores, we define the bias and the standard deviation of this estimator at a point (x, y).

First, denote the estimator at point (x, y) of DEA efficiency score $\theta(x, y)$ by $\hat{\theta}(x, y)$ and its bootstrap estimator by $\hat{\theta}^*(x, y)$.

These bias and standard deviation of $\hat{\theta}(x, y)$ cannot be computed because its sampling distribution is unavailable and its asymptotic approximation is too complicated to handle. Nevertheless, a bootstrap approximation is available and given by

$$\widehat{bias}^{*}\left(\widehat{\theta}\left(x,y\right)\right) \approx \frac{1}{B} \sum_{b=1}^{B} \widehat{\theta}_{b}^{*}\left(x,y\right) - \widehat{\theta}\left(x,y\right);$$
(3.1)

$$\widehat{std^{2}}^{*}(\hat{\theta}(x,y)) \approx \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_{b}^{*2}(x,y) - \left(\frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_{b}^{*}(x,y)\right)^{2};$$
(3.2)

Then, the bias corrected estimator is

$$\tilde{\theta}(x,y) = \hat{\theta}(x,y) - \widehat{bias}^* \left(\hat{\theta}(x,y) \right) = 2\hat{\theta}(x,y) - \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^*(x,y); \quad (3.3)$$

In (3.3), the correction is done by the mean. If the distribution of $\hat{\theta}^*(x, y)$ is asymmetric, the correction by the median can be used and may be more appropriate. In that case, we define the bias corrected estimator by

$$\tilde{\theta}(x,y) = 2\hat{\theta}(x,y) - \mathrm{median}\left(\hat{\theta}_b^*(x,y), \ b = 1, \cdots, B\right);$$
(3.4)

3.3 Confidence intervals for DEA efficiency scores

To do statistical inference and, in particular, to construct confidence intervals we need the distribution function of the variable of interest for computing or estimating the quantiles. Since in DEA the sampling distribution of $W = \hat{\theta}(x, y) - \theta(x, y)$ is unknown, the bootstrap method will provide an appropriate approximation, see e.g. Daraio and Simar (2007).

By definition, the efficiency's confidence interval at level $1 - \alpha$, for all $\alpha \in [0, 1]$, is

$$P\left(\hat{\theta}\left(x,y\right) - a_{1-\frac{\alpha}{2}} \le \theta\left(x,y\right) \le \hat{\theta}\left(x,y\right) - a_{\frac{\alpha}{2}}\right) = 1 - \alpha; \tag{3.5}$$

The method adopted to build the bootstrap confidence interval for efficiency is the basic bootstrap method that adjusts automatically for the bias of the DEA estimator. The bootstrap approximation of the confidence interval for $\theta(x, y)$ is given by

$$P\left(\hat{\theta}\left(x,y\right) - \hat{a}_{1-\frac{\alpha}{2}} \le \theta\left(x,y\right) \le \hat{\theta}\left(x,y\right) - \hat{a}_{\frac{\alpha}{2}}\right) \approx 1 - \alpha; \tag{3.6}$$

where $\hat{a}_{\alpha'} = \hat{c}_{\alpha'} - \hat{\theta}(x, y)$ and $\hat{c}_{\alpha'}$ is the α' -quantile of the empirical distribution of the estimators $\left\{\hat{\theta}_b^*(x, y)\right\}_{b=1}^B$.

As usual, the precision will be higher when the DEA frontier above (x, y) is determined by many sample points (X_i, Y_i) , as the length of the confidence interval will be smaller, and vice versa. Simar and Wilson (2000) have shown that the naive bootstrap described above is inconsistent, but a smoothed version of it can be shown to be consistent. The smoothed bootstrap of FEAR (the Frontier Efficiency Analysis with R) can be used to generate bootstrap replications, and this method is statistically consistent.

3.4 Testing returns to scale

The bootstrap can also be used for hypothesis tests, e.g. testing the returns to scale. The least restrictive model for returns to scale is the varying returns to scale (VRS) situation where the returns to scale are allowed to be locally increasing, then constant and finally non-increasing. Testing returns to scale (RTS) is carried out according to the following procedure where we test CRS against VRS, see e.g. Simar and Wilson (2002) and Daraio and Simar (2007). Let Ψ be the production set, defined by

$$\Psi = \left\{ (x, y) \in \mathbb{R}^{m+r}_+ | x \in \mathbb{R}^m_+, y \in \mathbb{R}^r_+, (x, y) \text{ feasible} \right\}$$

To test the null hypothesis $H_0: \Psi^{\partial} is CRS$ against the alternative $H_1: \Psi^{\partial} is VRS$, we first estimate the efficiency scores at all points (X_i, Y_i) for the two cases CRS and VRS,

denoted respectively $\hat{\theta}_{CRS}(X_i, Y_i)$ and $\hat{\theta}_{VRS}(X_i, Y_i)$. Then we define the test statistic

$$T\left(\chi_{n}\right) = \frac{1}{n} \sum_{i=1}^{n} \frac{\hat{\theta}_{CRS,n}\left(X_{i}, Y_{i}\right)}{\hat{\theta}_{VRS,n}\left(X_{i}, Y_{i}\right)}.$$
(3.7)

Under H_0 , $T(\chi_n)$ will be close to one since $\hat{\theta}_{CRS,n}(X_i, Y_i)$ and $\hat{\theta}_{VRS,n}(X_i, Y_i)$ are close to each other. By construction, $\hat{\theta}_{CRS,n}(X_i, Y_i) \leq \hat{\theta}_{VRS,n}(X_i, Y_i)$, and hence, under the alternative, $T(\chi_n)$ will be close to zero. Therefore, we reject H_0 for small values of $T(\chi_n)$, or formally, at level $\alpha \in (0, 1)$ if $p - value < \alpha$, where

$$p - value = P\left(T\left(\chi_n\right) < T_{obs} \mid H_0 \text{ is } True\right)$$

$$(3.8)$$

and T_{obs} is the value of $T(\chi_n)$ computed with the original observed sample χ_n . This probability cannot be computed analytically but we can approximate it by using the bootstrap by

$$p - value = \frac{1}{B} \sum_{i=1}^{B} I\left(T^{*b} \le T_{obs}\right),$$
 (3.9)

where $T^{*b} = T(\chi_n^{*b})$ is the value of T computed for each bootstrap sample, B is the number of pseudo samples χ_n^{*b} , and $I(\cdot)$ is the indicator function.

4 The free disposal hull approach

The DEA approach is based on a restrictive convexity assumption on the structure of the production set Ψ . Deprins, Simar and Tulkens (1984) have proposed an estimator supposing that the frontier of the production set is simply the boundary of the free disposal hull (FDH) of the data set. The FDH approach to estimate the frontier only requires strong disposability of inputs and outputs and variable returns to scale. Hence, the convexity assumption is not required in the FDH approach.

A DMU is declared inefficient if it is dominated by at least another DMU, meaning that it is possible to produce more outputs with less or the same inputs. Consequently, if a DMU is not dominated by any other DMU, it is declared FDH efficient, see e.g. De Sousa and Schwengber (2005).

In the input oriented case, the FDH efficiencies at a fixed point (x_0, y_0) , denoted $\hat{\theta}_{FDH}(x_0, y_0)$, can be estimated by solving the following linear program that has $\lambda_i \in (0, 1)$

instead of $\lambda_i \geq 0$ in comparison with the DEA linear program:

$$\min \theta$$

$$s.t. \begin{cases} \sum_{i=1}^{n} \lambda_i y_{ki} \ge y_0, \quad k = 1, \cdots, r \\ \sum_{i=1}^{n} \lambda_i x_{ji} \le \theta x_0, \quad j = 1, \cdots, m \\ \sum_{i=1}^{n} \lambda_i = 1, \\ \lambda_i \in \{0, 1\} \quad \forall i = 1, \cdots, n \end{cases}$$

$$(4.1)$$

In applications, these scores can be calculated as follows. For the sample $\chi = \{(X_i, Y_i), i = 1, \dots, n\}$ where $X_i \in \mathbb{R}^m_+$ and $Y_i \in \mathbb{R}^r_+$, let D_0 be the set of observations which dominate (x_0, y_0) ,

$$D_0 = \{i \mid (X_i, Y_i) \in \chi, X_i \le x_0, Y_i \ge y_0\}.$$

Then,

$$\hat{\theta}_{FDH}\left(x_{0}, y_{0}\right) = \min_{i \in D_{0}} \left\{ \max_{j=1, \cdots, m} \left(\frac{X_{i}^{j}}{x_{0}^{j}} \right) \right\},\tag{4.2}$$

where X_i^j is the j^{th} component of $X_i \in \mathbb{R}^m_+$ and x_0^j is the j^{th} component of $x_0 \in \mathbb{R}^m_+$.

First, the maximum part of the algorithm identifies the dominant DMU's relative to which a given DMU is evaluated. Then, the estimators of the FDH efficiency scores are calculated from the minimum part of the algorithm. For each DMU declared inefficient by the FDH approach, it is possible to find at least one DMU in the set D_0 that presents a superior performance relative to the first DMU.

Simar and Wilson (2000) have established the statistical properties of the FDH estimator in a multivariate context, in order to do inference either by using asymptotic distributions or by means of bootstrap. The FDH estimator, as other nonparametric estimators such as DEA, suffers from the curse of dimensionality due its slow convergence rate in high dimensions. In our application we will reduce the dimension of input and output space to avoid this problem.

5 Application to local government efficiency in Morocco

To reduce the monopoly of the central administration in decision making, the kingdom of Morocco opted, since the first years of independence, for a system of decentralization. This system allows to involve the citizens with the management of local business and to give a sense of responsibility to the local leaders. Since the 1960's, the country tried progressively to transfer certain responsibilities and certain authorities of the central government towards well defined local authorities. This transfer of responsibilities was accompanied by a transfer of financial resources to confront expenses.

Since 1997, the Moroccan local authorities include 16 regions, 68 prefectures and provinces and 1546 districts, of which 248 are urban and 1298 rural. The different local entities are managed by a council elected for a period of six years. Their financial transactions are established according to rules defined by the legislator and put back in a document called the budgetary document which describes the budget of the entity.

The budget is an act by which is planned and authorized all the loads and the resources of the local authority or their grouping. It is prepared, approved and executed according to the current laws, regulations and instructions. Nevertheless, local authorities have the possibility of establishing secondary budgets for specific operations.

The main budget contains two parts. The first one describes the operating budget, and the second is the budget of equipment or investment. Each of them contains two parts, one describes receipts and the other one the expenses. In this framework, to facilitate the statistical analysis of the budgets of local authorities, the various budgetary columns were numbered according to a well defined nomenclature. The budget is then divided in its two parts of recipes and expenses in Sections, Chapters, Articles and Paragraphs.

5.1 The data

We estimate the efficiency of the Moroccan rural districts by giving a particular attention to those of the oriental region. This region contains 91 rural districts. The inputs are constituted by ten variables which represent the categories of the financial resources of the local authority during the budgetary year 1998/1999. The ten variables describing the inputs are: The urban tax, the tax on the collection of the waste, the tax of the licence, the product of the forest domain, the taxes and assimilated taxes, the product of services, the product and the income of the goods, the concessions, the subsidies and competition and finally the order receipts.

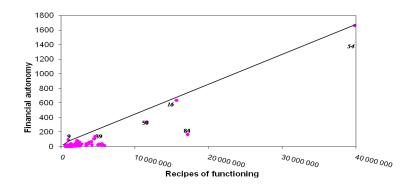
In order to reduce the dimension of the input space and thus to avoid the curse of dimensionality of nonparametric estimation, we decide to aggregate these ten input variables into a single one. See e.g. Daraio and Simar (2007, p.148) for a general justification of aggregate input and/or output measures. In our case, all input variables have the same scale and their unweighted sum has an economic meaning, which we call operat-

ing receipts. The operating receipts less the subsidies are called the own receipts of the municipality.

With respect to output, we define a variable which measures the financial autonomy of the municipalities, defined as the ratio of the own receipts of the municipality and its operating expenses. If this ratio is one or larger, then the municipality is financially autonomous, but not necessarily efficient. If the ratio is smaller than one, than it is not financially autonomous. Thus, we consider DEA and FDH efficiency estimates with a single input variable and a single output variable.

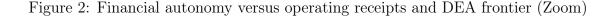
The data consists of pairs (X_i, Y_i) where X_i represents the input expressed by the operating receipts of the DMU_i used to produce the output Y_i expressed by the financial autonomy for the same DMU_i . The relationship between the output and the input reveals a possible positive trend, as with higher operating receipts the financial autonomy increases, as shown in Figures 1 and 2. These figures also suggest a possible existence of outliers in our data set.

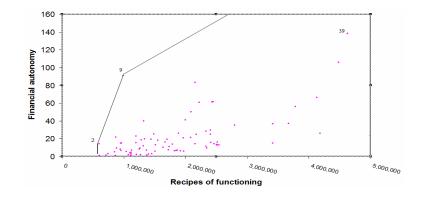
Figure 1: Financial autonomy versus operating receipts and DEA frontier



We see from Figures 1 and 2 that the districts RAS ASFOUR (2), LABKHATA (9) and AIN LEHJER (54) are estimated as efficient. However, LAAOUINATE (16) is almost efficient because it is very close to the frontier. The AIN LEHJER (54) district is isolated from the others, so it may be possible that it is an outlier. If it dominates several districts, removing it with other outliers may declare some districts efficient which were previously inefficient such as TIOULI (39). This finding confirms that frontier analysis is sensitive to outliers.

Since the presence of outliers may influence the efficiency scores, we used the procedure





of Wilson (1993) of detecting outliers in deterministic nonparametric frontier models. If outliers are identified, they will be deleted from the data set and efficiency scores will be re-estimated. The number of outliers being arbitrary in Wilson (1993), we first set it equal to ten and the procedure indicates that there are four possible outliers which are the districts of LAAOUINATE (16), IKSANE (50), AIN LEHJER (54) and SELOUANE (84). Results show that deleting these observations from the data set does not influence substantially the efficiency scores of the other DMU's. Thus, we keep these districts in the data set.

5.2 Interpretation of the results

In the following, we present the estimation results and their interpretation, first for DEA and then for FDH.

5.2.1 DEA results

Before applying DEA to our data, we performed a hypothesis test about the returns to scale (RTS) in order to decide which DEA linear program we shall adopt. The statistical test described in Section 3 reveals the existence of variable returns to scale (VRS). Therefore, the linear program used to estimate the scores of efficiencies of DEA is that represented in the envelopment model expressed in (2.1) or the Multiplier model expressed in (2.3). To avoid the inconsistency of the naïve bootstrap, we used the smoothed bootstrap of FEAR (the Frontier Efficiency Analysis with R).

From Table 1 we can see that the initial DEA efficiency estimators of all districts are well included in the unit interval. Furthermore, only three districts are efficient: RAS ASFOUR (2), LABKHATA (9) and AIN LEHJER (54), and one is close to the frontier with a score equal to 0.9899.

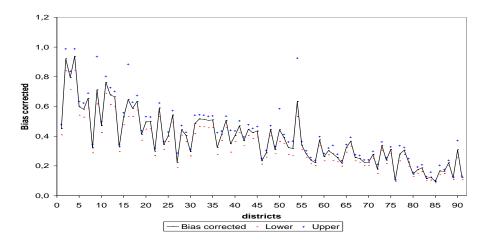
The districts in Table 1 are ranked with ascending population size. This suggests that rural districts having a large population size tend to have weak efficiency score and are consequently far from the efficiency frontier as shown in Figure 3. For instance, the estimator of the efficiency score has reached 0.1024 for the district of AREKMANE (ordered 85th among 91 according to population size). The negative relation between population size and efficiency is confirmed by a Kendall test which strongly rejects that the correlation between the population size and the DEA efficiency estimates is equal to zero with a p - value < 2.2e - 16 and indicates an inverse relation between them with a Kendall's tau estimate given by $\tau = -0.6395$. The truncated regression estimation on the Shephard efficiency scores as an environmental variable of the two-stage procedure described in Simar and Wilson (2007) confirm indirectly this relationship.

We used the bias correction given by expression (3.3). Denoted by $\tilde{\theta}$, the bias corrected estimators indicate that there are no efficient rural districts and that only 6% of the districts are close to the efficiency frontier with a score estimated above 0.70. In addition, the district of AIN LEHJER (54), which was declared efficient by the initial estimator of the DEA efficiency score, is inefficient with an estimator $\tilde{\theta}$ equal to 0.6336, as the bias was important. This indicates that even with a financial autonomy of 1656%, this district clearly fails to reach the efficiency frontier. The district of AREKMANE (85) recorded the lowest score of efficiency estimated at 0.0944.

In order to test if the bias can be disregarded, the ratio of the estimated bias to the standard error of the bootstrap estimates defined by $|\widehat{bias}^*(\hat{\theta}(x,y))|/\hat{\sigma}^*$ has been analyzed. Following a suggestion by Efron (1982), the bias is significant since the ratio exceeds 0.25. It is possible to conclude in our case that the bias correction should be used and the bootstrap bias correction provides more accurate results. On the other hand, we note that the bias corrected estimator for each district is obviously in the corresponding confidence interval, but the range of the interval is more important for some districts such as RAS ASFOUR (2), LABKHATA (9), LAAOUINATE (16) and AIN LEHJER (54), which have relatively high bias corrected estimators. This can also be viewed in Figure 3.

It should be noted also that the initial DEA estimators of the efficiency scores are often outside the corresponding confidence interval. They are also close to the upper bound of





the interval, since they are positively biased.

5.2.2 FDH results

As pointed out in Table 1, eight districts are declared efficient using the FDH approach. This represents only 8.79% of the total of the population under study. Nevertheless, this percentage is almost three times higher than that given by the DEA approach. Note also that districts which are declared efficient by DEA approach are also efficient by the FDH approach, which follows by construction because the FDH approach is less restrictive than that of DEA. On the other hand, the most populated districts have generally a weak FDH efficiency reflecting the same finding as for the DEA approach. The correlation between estimated DEA and FDH efficiency scores is 0.845, which shows that both approaches tend to give similar results where mainly the distance from the frontier differs.

Column 4 of Table 1 reports the number of districts which dominate a given district (including the own district). Each FDH inefficient district is dominated by at least another district. For instance, the inefficient district AREKMANE (85) is dominated by 46 districts. This means that with the same quantity of resources, ratios of the financial autonomy of the 46 districts exceed that of AREKMANE. Furthermore, it can reach efficiency if it reduces its resources by 90%, meaning that it can be efficient with only 10% of its resources.

6 Conclusion

The technical efficiency determination in the input orientation of the Data Envelopment Analysis requires testing returns to scale in order to define the primal linear programming problem. A procedure for the determination of the dual from the primal model was developed for the case of constant returns to scale. Since the efficiency scores are often over-estimated, a bootstrap procedure is used to correct the bias by the mean or by the median. The bootstrap efficiency scores allow us to make statistical inference on the DEA efficiency by using them to build confidence intervals.

In this study, we estimate efficiency scores of the financial autonomy of the Moroccan rural districts in the oriental region for the budgetary year 1998/1999. The inputs are expressed as the operating receipts for the DMU_i to produce the output expressed as the financial autonomy for the same DMU_i . Statistical tests suggested variable returns to scale (VRS) for the data. Bias corrected results indicated that they are well in the unit interval and in the corresponding confidence interval. They indicated also that there are no efficient rural districts and that only 6% of the districts are close to the efficiency frontier with a score estimated above 0.70. In addition, the most efficient district is AIN LEHJER, but even with a financial autonomy of 1656% it fails to reach the efficiency frontier.

Being less restrictive than the DEA approach, the FDH analysis delivered efficiency scores generally larger than those of DEA but the DMU ranking is very similar for both approaches.

Finally, we found that generally rural districts having a large population size have a weak efficiency score. If data become available, a detailed analysis using the two-stage procedure described in Simar and Wilson (2007) could be done on the socio-economic and demographic factors such as the geographical distance from the center and the training level of the local council members, which may explain these inefficiencies.

Acknowledgements

We would like to thank Léopold Simar for his great help, constructive comments and suggestions. Thanks are also addressed to the Belgian Technical Cooperation (CTB) for the funds allocated for this research and to the Mohammed First University of Oujda for the free time assigned to the realization of this work. Table 1: Efficiency scores of FDH, DEA with VRS and their confidence intervals

11 0.9878Upper 0.83750.68970.34590.9360.49680.72650.70140.34630.88420.67380.5313Bound 0.47770.98770.63370.62340.80310.55770.62910.43610.52920.31950.62350.36450.42870.2865U 0.57240.47210.4268Lower Bound 0.83950.71560.84360.54190.52850.61680.42450.68630.6130.59950.29620.47650.53360.53330.57570.3730.44760.45270.26970.53330.3107 U 10 0.58910.28750.36350.48920.19040.40230.36480.4085 Std DEAс О 0.0260.05830.04580.05410.03450.03780.02470.04570.01880.03050.74330.03680.02370.03440.01970.03390.03670.40770.02710.04380.03810.0370.02870.02010.02560.03110.10960.0260.02330.0783 Bias DEA ∞ 0.04420.05180.03820.04820.03620.02920.28830.02620.04280.05660.03860.02940.04870.04230.02530.03060.02450.03620.02380.03150.08050.03090.02010.04210.03250.02270.28710.0287Eff 0.9217 Bias COLL. DEA 1-0.79550.93810.58180.32090.71170.47190.76250.67790.66490.32750.52970.64650.58710.63490.41250.49580.50060.29830.34260.45110.59830.65510.58970.4010.54260.22340.44360.4053Mean Boot Eff 9 1.09340.89071.04970.68080.68740.3861.76150.52840.74820.37090.59311.8950.69360.72630.46710.58880.35220.39420.4954DEA0.51320.73310.80260.56680.47210.61060.51040.45430.85470.6678Effic 0.69130.53790.98990.63650.630.73450.70350.63580.67720.43780.53120.3228DEA ഹ 0.83970.80530.34760.93360.62590.36640.43350.30390.47980.35010.49810.55910.57410.47450.428Domin FDH 4 102 2 4 ŝ 9 က 21 2 2 26132 ŝ $\frac{18}{18}$ 2 6 1 9 179 10 x 26Effic 0.69130.70350.34760.4378FDH က 0.47980.83970.98990.63650.90790.52420.49810.80530.55910.91610.67720.82430.53120.4430.62590.36640.47450.4280.60440.5741OULAD SIDI ABDELHAKEM OULAD DAOUD ZKHANINE SIDI MOUSSA LEMHAYA OULAD M'HAMMED SIDI BOULENOUAR AL BARKANYENE OULAD GHZIYEL ABBOU LAKHAL AIN-CHOUATER SIDI BOUBKER TAFOUGHALT LAAOUINATE SIDI BOUHRIA RAS ASFOUR BNI MATHAR MESTEGMER GUENFOUDA MESTFERKI **FAZAGHINE** LEBKHATA LABSARA M'HAJER RISLANE EL ATEF AIN SFA GTETER GAFAIT AFSOU MRIJA District 2 Pop ഹ Ord 4 9 1 ∞ 6 2 က 10 1213 1415 $\begin{array}{c} 16\\ 17\\ 18\\ 18\end{array}$ 19 2022 23 24 25 26 11 2127 28 29 Table 1: Efficiency scores of FDH, DEA with VRS and their confidence intervals

Π 0.54140.5419Upper 0.43460.5350.43960.39530.45290.5849Bound 0.31190.54380.5340.53630.42410.43550.50250.47720.46510.25010.30270.47030.3310.4120.3610.36650.92580.36330.30180.25570.23785 Lower Bound 0.41830.46470.46310.45630.37120.45730.29280.36270.4260.33530.40480.38490.39340.21210.25880.40180.28130.36390.26790.21770.2018U 10 0.26670.45810.27740.35210.27720.53210.3080.2547 Std DEAс О 0.06630.02940.15980.02380.0290.16740.02340.02820.02570.01470.01640.02570.01950.35330.04580.06361.70520.01380.01690.02950.02910.02940.030.02960.02810.02240.0217 0.01860.01430.0659 Bias DEA ∞ 0.02930.11590.02280.02950.11750.03830.02980.03240.03550.017 0.02490.02630.17040.04530.05630.0175 0.01790.01760.02940.02810.03570.01910.0220.02750.02330.02840.03610.36640.48560.5160Eff Bias COLL. DEA 1-0.29540.51410.50700.50940.41280.50710.34860.40400.46970.36990.44650.42360.43480.23380.28670.44660.30880.44370.39120.32280.31770.63360.33980.23980.22190.32440.28179 Mean Boot 0.64410.57940.57720.56830.57000.71510.46200.57060.73700.48340.55340.43510.52530.49410.51250.27550.32350.50010.36641.08780.43850.43230.46222.98900.2778Eff DEA 0.333330.40000.33270.2609Effic 0.5515DEA 0.3130.54540.54340.44030.5080.4560.47030.25290.47150.2573Ŋ 0.53540.53750.43560.53660.43970.39970.48260.30370.61410.41320.3050.23980.46610.33510.36810.3740.3673Domin FDH 4 342 11 1215 $\frac{18}{18}$ 2 33 11 ŝ 4 6 4 6 S 24 36 24 16- 28 2 \mathbf{S} $-\infty$ 6 2934Effic 0.56190.4560.30370.4715FDH က 0.3130.74590.54540.54340.53540.53750.45910.43560.53660.66350.72590.67840.65050.36140.33510.41320.49440.51220.41850.25730.23980.4731BNI OUKIL OULAD MHAND MECHRAA HAMMADI MELG EL OUIDANE OULAD BOUBKER BNI-SIDEL-LOUTA OULAD AMGHAR BNI MARGHNINE HASSI-BERKANE BOUCHAOUENE **3NI-SIDEL-JBEL** BOUMERIEME BNI KHALED SIDI LAHSEN *TIZTOUTINE* TANCHERFI MAATARKA AIN LEHJER BOUANANE OUARDANA FEZOUANE AMEJJAOU AIT-MAIT BNI GUIL TALILIT IKSANE TIOULI TSAFTAZLAF District ISLY 2 Pop 4835 36 37 $339 \\ 339$ 40424344 454647 49 50 $52 \\ 53 \\ 54$ 56Ord 30 313233 3441515557 58 Table 1: Efficiency scores of FDH, DEA with VRS and their confidence intervals

11 0.2836Upper 0.39590.320.3380.27880.23420.34360.23990.24030.29870.21090.36160.25430.1103Bound 0.39240.27510.27010.32810.33550.32370.24860.1560.19260.20690.1580.20365 0.1271 0.1011 0.1753 Bound Lower 0.23390.19880.29250.3330.23380.22490.20340.20420.24870.14960.30650.21670.28050.09360.23360.27690.19730.15080.162U 10 0.33850.27370.23510.22740.1297 0.09950.10440.0860.14020.1486 Std DEAс О 0.0217 0.01740.07930.01360.01920.01620.01860.01420.02060.04360.02130.0140.01780.00650.08060.01760.02520.02110.01390.02150.02310.02270.0111 0.02270.06530.006 0.05430.0103 0.0141 Bias DEA ∞ 0.02230.02530.01860.06290.02620.01760.02350.02950.02180.02220.01820.0190.02460.03590.02750.01660.01850.00840.06420.01870.02650.02140.0230.01420.04280.0080.04390.01360.01310.3749Eff 0.26230.2813 Bias COLL. DEA 0.30270.25650.21860.32220.36620.25670.25050.22430.22420.2770.17870.3380.2390.31070.10320.27910.30620.22580.14480.17370.18660.11460.12461-0.09440.16610.16369 Mean Boot Eff 0.31930.34280.45390.31610.25700.37320.43050.30450.29980.26420.26590.33160.27440.39820.2750.35060.12150.45680.34670.28760.17410.22430.2410.1480.2688DEA 0.42300.1119 0.29350.1933Effic 0.28760.1116DEA 0.32130.34420.30160.21460.32920.34330.3249Ŋ 0.28270.23620.34570.39570.27850.24250.24320.36550.25560.25230.15790.20960.12880.17720.39720.27270.19510.16740.10240.21Domin FDH 4 21 1-30 4 1-361813 2211 27 32 ∞ 9 11 33 30 34 \mathfrak{S} 296 17106 15က 47 က 31 Effic 0.32120.2556FDH က 0.39720.42410.40650.41070.23620.34570.39570.27850.41150.34660.24320.4550.26140.52240.32920.15810.4030.32490.35340.23650.26950.28950.17790.67510.25530.10240.2394SIDI ALI BELQUASSEM DAR-EL-KEBDANI **JULAD SETTOUT** BNI-BOUIFROUR AHL OUAD ZA **JERMAOUAS HADDADENE** TEMSAMANE IAAZZANENE BNI-TADJITE AHL ANGAD TROUGOUT AREKMANE TENDRARA AIN ZOHRA LAATAMNA CHOUIHIYA FARKHANA SELOUANE RAS-EL-MA BOUDINAR **BEN TAIEB** TAFERSIT MADAGH TALSINT AGHBAL BOUARG MIDAR IFERNI District 2 Pop 596263 6566 $67 \\ 68$ 69 71 72 72 73 75 75 77 77 79 77 79 80 82 84 85 Ord 0061648183 86

	Upper Bound CI	11	0.2368 0.1277 0.3703 0.131
	Lower Bound CI	9 10	$\begin{array}{c} 0.2012 \\ 0.1076 \\ 0.2575 \\ 0.1064 \end{array}$
Table 1: Efficiency scores of FDH, DEA with VRS and their confidence inter- vals	Std DEA	6	$\begin{array}{c} 0.0134 \\ 0.0083 \\ 0.0869 \\ 0.0869 \\ 0.0113 \end{array}$
	Bias DEA	8	0.017 0.0101 0.0689 0.0127
	Bias corr. Eff DEA	6 7	0.2384 0.2583 0.2214 0.1293 0.1415 0.1192 0 0.377 0.4972 0.3081 0 0.1329 0.1491 0.1202 0
VRS and	Mean Boot Eff DEA	9	$\begin{array}{c} 0.2583 \\ 0.1415 \\ 0.4972 \\ 0.1491 \end{array}$
EA with	Effic DEA	5	37 0.2384 24 0.1293 3 0.377 16 0.1329
of FDH, D	Domin FDH	4	37 24 3 16
1: Efficiency scores c	Effic FDH	3	$\begin{array}{c} 0.2384 \\ 0.1981 \\ 0.4453 \\ 0.1916 \end{array}$
Table vals	District	2	BOUGHRIBA BNI-CHIKER ZEGZEL DRIOUCH
	Pop Ord	1	88 89 90 91

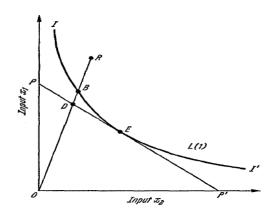
Column 1: Number of district, ordered with respect to increasing population size; 2: Name of district; 3: FDH efficiency score; 4: Number efficiency scores; 8: estimated bias of DEA efficiency scores using the bootstrap; 9: standard deviation of bootstrap DEA efficiency scores; 10: of dominating districts (including the own one); 5: DEA efficiency score; 6: Mean of bootstrapped DEA efficiency scores; 7: bias-corrected DEA lower bound of 95% bootstrap confidence interval; 11: upper bound of 95% bootstrap confidence interval

Appendix

Since its introduction in 1957, Farrell's efficiency measure has been widely used in empirical research on production efficiency in order to measure the efficiency of firms, countries, or other decision making units, see e.g. Färe (1984) for a detailed survey.

In Figure 4, for the case of a firm which uses two inputs to produce one output under constant returns to scale, consider the isoquant II' and the isocost line PP' that minimizes total cost of producing one unit of output. Let R be a vector of input quantities to produce a unit output belonging to the input correspondence image set L(1).

Figure 4: Technical and allocative efficiencies



Then, for a given input vector, Farrell defines the degree of technical efficiency (TE) as the ratio $\frac{OB}{OR}$, the allocative efficiency (AE) as $\frac{OD}{OB}$ and finally the overall productive efficiency (OPE) or the total economic efficiency as $\frac{OD}{OR}$. The Farrell technical efficiency is denoted θ . Note that the product of technical and allocative efficiencies provides the overall efficiency and all three measures of efficiency are between zero and one.

Furthermore, the distance DB represents the reduction in production cost that would occur if production were to occur at the allocatively (and technically) efficient point E. The distance BR can also be interpreted in terms of a cost reduction. The line PP'represents the input price ratio.

By definition, the technical efficiency reflects the ability of the firm to obtain maximal output from a given set of inputs; the allocative efficiency reflects the ability of the firm to use the inputs in optimal proportions, given their prices. These two efficiency components are combined to obtain the total economic efficiency.

References

- Afonso A. and Fernandes S., 2008. Assessing and Explaining the Relative Efficiency of Local Government. Journal of Socio-Economics, Vol 37, N^o 5, pp 1946-1979.
- [2] Balaguer-Coll M.T., Prior D. and Tortosa-Ausina, E., 2007. On the determinants of local government performance: A two-stage nonparametric approach. European Economic Review, Vol 51, 425-451.
- [3] Banker R. D., Cooper W. W., Seiford Lawrence M., Thrall Robert M. and Zhu Joe, 2004. Returns to scale in different DEA models. European Journal of Operational Research 154 (2004) 345-362.
- [4] Daraio C. and Simar L., 2007. Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications. Springer.
- [5] De Sousa M. and Schwengber S., 2005. Efficiency estimates for judicial services in Brazil: Nonparametric FDH (Free Disposal Hull) and the expected order-M efficiency scores for Rio Grande Do Sul courts. Encontro da anpec, 33th.
- [6] Deprins D., Simar L., and Tulkens H., 1984. Measuring labor-efficiency in post offices. In Marchand, M., Pestieau, P. and Tulkens, H. (Eds.). The performance of public enterprises: concepts and measurement. Amsterdam: North-Holland, pp. 243-267.
- [7] Efron B., 1982. The Jackknife, the Bootstrap, and Other Resampling Plans. CBMS-NSF Regional Conference Series in Applied Mathematics. #38. Philadelphia: SIAM.
- [8] Färe R., 1984. The Dual Measurement of efficiency. Journal of Economics, Vol. 44, N^o 3, pp. 283-288.
- [9] Farrell M.J., 1957. The measurement of productive efficiency. Journal of the Royal Statistical Society, Vol. 120, 253-281.
- [10] R Development Core Team, 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.
- [11] Simar L. and Wilson P. W., 2007. Estimation and inference in two-stage, semiparametric models of production processes. Journal of Econometrics 136, pp. 31-64.

- [12] Simar L. and Wilson P. W., 2006. Statistical Inference in Nonparametric Frontier Models: Recent Developments and perspectives. Oxford University Press(2008)
- [13] Simar L. and Wilson P. W., 2002. Nonparametric tests of returns to scale. European Journal of Operational Research, 139, pp. 115-132.
- [14] Simar L. and Wilson P. W., 2000. Statistical Inference in Nonparametric Frontier Models: The State of the Art. Journal of Productivity Analysis, 13, 49-78.
- [15] Simar L., 1992. Estimating Efficiencies from Frontier Models with Panel Data: A Comparison of Parametric, Non-Parametric and Semi-Parametric Methods with Bootstrapping. The Journal of Productivity Analysis, 3, 171-203.
- [16] Sueyoshi T., 1999. DEA Duality on Returns to Scale (RTS) in Production and Cost Analyses: An Occurrence of Multiple Solutions and Differences Between Production-Based and Cost-Based RTS Estimates. Management Science, Vol. 45, No. 11 (Nov., 1999), pp. 1593-1608.
- [17] Thanassoulis E., 2001. Introduction to the theory and application of data envelopment analysis: A foundation text with integrated software. Springer, 281 p.
- [18] Wilson P. W., 1993. Detecting Outliers in Deterministic Nonparametric Frontier Models With Multiple Outputs. Journal of Business & Economic Statistics, Vol. 11, No. 3.
- [19] Worthington A., 2000. Cost efficiency in Australian local government: A comparative analysis of mathematical programming and econometric approaches. Financial Accountability and Management, Vol. 16, No. 3, pp. 201-224.