ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Economic Review ▮ (▮▮▮▮) ▮▮▮–▮▮▮

EUROPEAN ECONOMIC REVIEW

www.elsevier.com/locate/eer

On the determinants of local government performance: A two-stage nonparametric approach

Maria Teresa Balaguer-Coll[a], Diego Prior[b], Emili Tortosa-Ausina[c,*]

[a] *Universitat Jaume I, Spain*

[b] *Universitat Autònoma de Barcelona, Spain*

[c] *Departament d'Economia, Universitat Jaume I and Instituto Valenciano de Investigaciones Económicas (IVIE), Campus del Riu Sec, 12071 Castelló de la Plana, Spain*

Received 16 October 2003; accepted 20 January 2006

Abstract

This article analyzes the efficiency of local governments in the Comunitat Valenciana (Spain) and their main explanatory variables. The analysis is performed in two stages. Firstly, efficiency is measured via (nonparametric) activity analysis techniques. Specifically, we consider both Data Envelopment Analysis (DEA) and Free Disposable Hull (FDH) techniques. The second stage identifies some critical determinants of efficiency, focusing on both political and fiscal policy variables. In contrast to previous two-stage research studies, our approach performs the latter attempt via nonparametric smoothing techniques, rather than econometric methods such as OLS or Tobit related techniques. Results show that efficiency scores, especially under the nonconvexity assumption (FDH), are higher for large municipalities. Thus, there is empirical evidence to suggest that resources may be better allocated by large municipalities. However, the inefficiency found is not entirely attributable to poor management, as second-stage analysis reveals both fiscal and political variables to be explicably related to municipality performance. Moreover, the explanatory variables' impact on efficiency is robust to the chosen technique—either convex DEA or nonconvex FDH.

JEL classification: D24; D60; H71; H72

Keywords: Efficiency; Kernel smoothing; Local government; Nonparametric regression

*Corresponding author. Tel.: +34 964387163; fax: +34 964728591.
E-mail address: tortosa@uji.es (E. Tortosa-Ausina).

0014-2921/$ - see front matter
doi:10.1016/j.euroecorev.2006.01.007

1. Introduction

All individuals, governments, and firms have an interest, one way or another, in improving efficiency and productivity in public sector activities. The study of public sector efficiency is of paramount importance if we consider that bureaucrats may have an incentive to waste resources, not to use inputs optimally and to produce too much, as claimed by Niskanen (1975). This ability shown by bureaucrats in pursuing objectives that are not in the interest of the citizen-voter may not be unlimited (Grosskopf and Hayes, 1993). For these and related reasons, many research studies have analyzed different aspects of efficiency and productivity in important public sector areas such as health, education, taxation or, in the case we are dealing with, different government departments such as municipalities.[1]

Research in the previous literature on the efficiency and productivity of municipalities all over the world varies widely in many aspects, ranging from their aims to their conclusions. Studies that are closer to our attempts and techniques include De Borger and Kerstens (1996), De Borger et al. (1994), Hayes and Chang (1990), or Deller (1992). For a survey of frontier efficiency measurement techniques in local government see Worthington and Dollery (2000). Other related studies include Davis and Hayes (1993) or Grosskopf and Hayes (1993) and, in general, the literature on local government efficiency and property values (see, for instance Brueckner, 1979; Grossman et al., 1999). It should also be noted that some studies focus on the efficiency of a single service rendered by local authorities, although this depends critically on the services and competencies of local authorities, which differ widely from country to country.

In this article we focus on the efficiency of local governments in the Valencian region (Comunitat Valenciana, Spain), along with its determinants. Most previous two-stage studies measure efficiency in the first stage via nonparametric techniques, and then regress it on environmental variables in the second stage in the search for exogenous factors that might affect municipalities' performance using parametric methods. In contrast, we propose a consistent (fully) nonparametric approach. Therefore, we have a twofold objective, i.e., to analyze the efficiency of Valencian local governments, and to use smoothing regression techniques, to date barely contemplated in the efficiency analysis setting.

Despite its scantness, the study of Spanish local government efficiency is relevant for several reasons. Amongst them, we find that the decentralization of public responsibilities to lower levels of government has grown rapidly in Spain over recent years,[2] fueling the debate on operational efficiency in local governments. During the last twenty-five years, both regional and local administrations in Spain have gained competencies at the expense of central administration, especially since the late seventies following the establishment of democracy.[3] Specifically, the Spanish Constitution granted higher independence to state and local governments, *Comunidades Autónomas*—states or regions, NUTS2 in European terminology—, *Provincias*—provinces, or NUTS3—, and *Municipios*—municipalities, or

[1]An introduction to these topics is provided by Fox (2001).

[2]Similar to the experience of the U.S., where one of the legacies of the Reagan Administrations was the decentralization of public functions to lower levels of government.

[3]However, the decentralization at the *first* level, i.e., regional, has largely surpassed that which occurred at the *second* level, i.e., local, which has actually taken place to a lesser extent—at least when juxtaposed.

NUTS5 (Prieto and Zofío, 2001). Proponents of decentralization of public functions to lower levels of government may argue that local responsibility contributes not only to a better match between public services and the needs or preferences of a diverse citizenry, but is also an effective way to control the overall growth of government (Deller, 1992; Marlow, 1988). Indeed, if effective policy is to be formulated, it becomes necessary to improve our understanding of managerial capacity in local governments, due to the increase in functions deriving from decentralization. However, the literature on the relative superiority of a city manager form of government is inconclusive (Hayes and Chang, 1990) as, on the other hand, some authors argue that the small size of operation for many local governments is inherently inefficient in economic terms, and hence costly. This is also related to the debate on whether economies/diseconomies of scale for municipalities exist, which considers whether it may be less costly for local governments to contract out some services to the private sector or, alternatively, to provide a diversified bundle of services (Grosskopf and Yaisawarng, 1990).

Other papers that analyze efficiency in Spanish municipalities are those by Prieto and Zofío (2001) and Giménez and Prior (2003); they differ from ours in several aspects, including the regions studied.[4] Other works concentrate on the evaluation of specific services: Garbage collection, urban public transport, water supply, local police, fire service, etc. As an example, Bosch et al. (2000) also deal with efficiency issues, but their study is confined to the measurement of efficiency in Spanish municipal garbage collection services. An alternative perspective is presented in this paper, with a focus on the evaluation of local organizations as a whole, such as decision-making units that organize the production process of multiple services. This perspective offers a link to the contributions by Vanden Eeckaut et al. (1993) and De Borger and Kerstens (1996), although we employ a set of techniques which do not exactly coincide with theirs.

As indicated above, we focus not only on efficiency but, more importantly, on its determinants. In this case, as suggested by De Borger et al. (1994), we find that many studies that set out to estimate inefficiencies in the public sector "simply do not attempt to explain the estimated inefficiencies in a systematic way". In other words, it may also be of interest to determine whether some factors affect the performance of municipalities. This is the spirit of the so-called two-stage, or two-step analyses.

On this point, previous literature has focused less thoroughly on the determinants of inefficiency. Our study differs substantially from previous contributions[5] in that we shift this focus by introducing techniques that have not previously been exploited in this field. The techniques we consider to measure efficiency are fairly standard. Specifically, we use nonparametric techniques, which are quite often employed in public sector studies. However, we also use a set of nonparametric techniques to analyze the determinants of efficiency, based on previous work by Deaton (1989), DiNardo et al. (1996), or Marron and Schmitz (1992) amongst others, in contrast to the more common set of parametric techniques—such as OLS, or Tobit censored regression model—used in the two-stage analyses. These techniques focus on graphical aspects of efficiency results, which provide a great deal of meaningful information, and which may be particularly relevant in our

[4]Other studies focusing on Spanish local government issues include those by Bosch et al. (2000), Gil Jiménez (2001), Solé-Ollé (1997), or Vela (1996).

[5]Such as, for instance, the application of De Borger and Kerstens (1996) for Belgian municipalities.

setting, given the peculiar distributions of nonparametric efficiency scores, which are bounded at unity.

The study is organized as follows. After this introductory section, Section 2 introduces the model for measuring efficiency. Section 3 presents the data, inputs and outputs definition, and results. Section 4 sets out both the different explanatory variables for efficiency and the techniques to be employed in the second-stage analysis, together with some results. Finally, Section 5 outlines the most relevant conclusions.

2. Nonparametric inefficiencies

As has been well established, when inefficiency is found, the analyst's decision on how to reach the frontier is conditional on the objectives of the units under evaluation and on the degree of control of the variables. Hence, when inputs are fixed and the market is growing it seems fairly reasonable to take an output orientation. In contrast, when output is exogenous, an input orientation seems more appropriate. It may also be the case that, for practical purposes, the most recurrent choice to avoid inefficiency is to mix the output and the input orientation or, using a more technical nomenclature, to take the directional distance functions measures (see Färe and Grosskopf, 2000). In our specific case study local governments take outputs as exogenous but they have the capacity to control inputs, especially in the long run. Thus, the input orientation seems the most appropriate choice of orientation when the analyst is interested in efficiency corresponding to municipalities.

To introduce some notation, let us assume that for S observations there are J inputs producing I outputs. Hence, each observation s uses an input vector $\mathbf{x}^s = (x_1^s, \ldots, x_j^s, \ldots, x_J^s) \in \mathbb{R}_+^J$ to produce an output vector $\mathbf{y}^s = (y_1^s, \ldots, y_i^s, \ldots, y_I^s) \in \mathbb{R}_+^I$. Production technology is defined by the set of feasible input and output vectors

$$F = \{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \text{ can produce } \mathbf{y}\}. \tag{1}$$

It is also useful to consider the input set associated with this technology. For a given output vector $\mathbf{y}^s$, the input set denotes all input vectors $\mathbf{x}$ capable of producing the output vector

$$L(\mathbf{y}^s) = \{\mathbf{x} | (\mathbf{x}, \mathbf{y}^s) \in F\}. \tag{2}$$

The next theoretical building block is the cost efficiency measurement.[6] To this end, we denote the observed total cost vector as $TC^s = \mathbf{p}^s\mathbf{x}^s$.

According to Färe et al. (1994), the minimal cost may be calculated as the solution to the mathematical programming problem:

$$\begin{aligned} OE^s = \min_{\theta,\lambda} \ & \theta \\ \text{s.t.} \quad & TC^s\theta - \boldsymbol{\lambda}\mathbf{c} \geqslant \mathbf{0}, \\ & -\mathbf{y}^s + \boldsymbol{\lambda}\mathbf{M} \geqslant \mathbf{0}, \\ & \vec{\mathbf{1}}\boldsymbol{\lambda} = 1, \end{aligned} \tag{3}$$

[6]We introduce cost vector notation since we have information relative to specific costs, although it is not possible to decompose it into inputs and input prices. Therefore, the mathematical programs defined in terms of cost—instead of inputs and input prices—are precisely the programmes solved in the empirical application. We used GAMS code to solve the mathematical programming problems considered here. They are available upon request.

where $\mathbf{c}$ is a vector containing the observed S total costs, $\mathbf{M}$ is a matrix containing the observed S output vectors and $\boldsymbol{\lambda}$ is the activity vector denoting the intensity levels at which the S observations are conducted.[7] The solution of mathematical programming problem (3) yields optimal values for OE^s (cost efficiency coefficient) and λ^* (activity vector) for each s municipality. The value of OE^s is smaller than unity for inefficient observations, and equals unity for efficient observations ($TC^* = OE^s \cdot TC^s$).

The activity vector λ allows actual production plans to be mixed in such a way that the implicit reference technology defined in the mathematical programming problem (3) is convex, since the frontier production plan is defined by combining some actual production plans. However, another reference technology is available if we do not postulate convexity, namely, the nonconvex reference technology defined by the free disposable hull (FDH) frontier (see Deprins et al., 1984).[8] As suggested by Tulkens (1993), it is straightforward to relate it to the convex VRS, since only one new restriction must be included in the mathematical programming problem (3), namely, $\lambda^s \in \{0, 1\}, s = 1, \ldots, S$.[9]

For the one output case, Fig. 1 represents the two cost frontiers. The nonconvex FDH frontier is represented by the inverted staircase line $ABCDEF$. This frontier is nested in the convex variable returns to scale frontier—determined by mathematical programming problem (3)—whose piecewise linear frontier is $ABDEF$. Apart from the convexity postulate, it is worth noting that unit C forms part of the FDH efficient frontier but is inefficient according to the convex reference technology. Another difference between convex and nonconvex cost frontiers is related to the output slacks. In Fig. 1 we observe how the convex cost frontier corresponding to unit G is located along the line connecting observations B and D, and with no slacks (say, maintaining the same production y_g). In contrast, the minimum cost reference in the nonconvex frontier is that corresponding to unit D. This has two implications: (i) for unit G, the FDH total cost is higher than the variable returns to scale DEA total costs; and (ii) while there is no output slack in the DEA frontier reference, there is important output slack in the FDH case.[10]

3. Computing efficiency measures for Spanish municipalities

3.1. Data, inputs, and outputs

The different types of data we deal with throughout the paper were provided by various institutions. Inputs came from the budget data of local authorities in the Valencian region, which presented information to the Valencian Audit Institution in the year of study (1995).

[7]The constraint $\overrightarrow{\mathbf{1}}\boldsymbol{\lambda} = 1$ accounts for the assumption of variable returns to scale (VRS). By removing such a constraint we would assume constant returns to scale (CRS) technology.

[8]As pointed out by one referee, FDH turns out to possess rather attractive statistical properties since it is a consistent estimator for any monotone boundary (by imposing only strong disposability). Asymptotic rationale such as imposing convexity is weak. When the true technology is convex the FDH estimator converges to the true estimator, albeit at a slow rate. By contrast, a convex model causes specification error when the true technology is nonconvex. See Park et al. (2000) or the literature review in Simar and Wilson (2000).

[9]See also Briec et al. (2004) for an up-to-date contribution. They consider a series of nonparametric, deterministic technologies and cost functions without maintaining convexity.

[10]For estimating slacks we followed Ali and Seiford (1993), who suggest a two-step procedure whereby first the efficiency measure is estimated and the slacks are determined in a second step, while maintaining fixed the optimal efficiency measure obtained in the first step. We are grateful to one of the referees who pointed out the relevance of including this explanation.

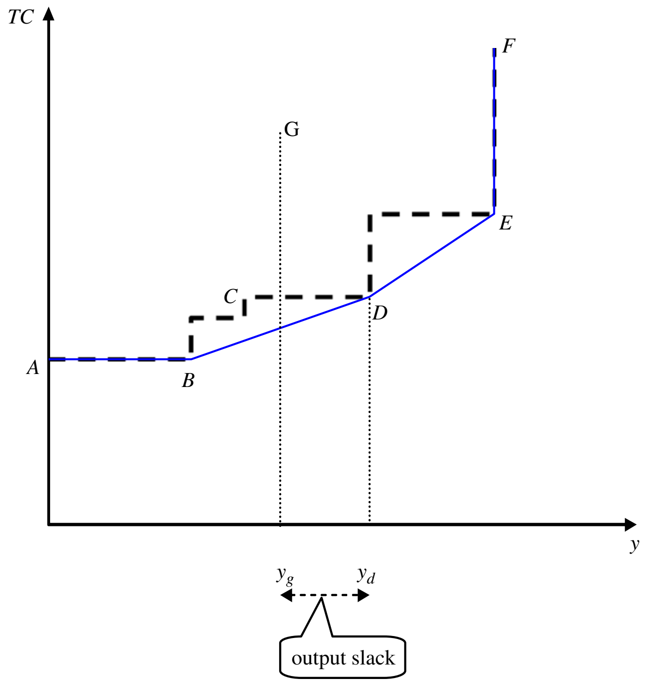

Fig. 1. Alternative shapes for the boundary of the cost frontier under convex (DEA) and nonconvex (FDH) reference technologies.

Outputs were obtained from information gathered in the survey on local infrastructures and facilities devised by the Spanish Ministry for Public Administration (*Encuesta de Infraestructuras y Equipamientos Locales*).

As stated by Fox (2001), it is difficult to measure the inputs and, more especially, the outputs of a hospital, a police force, or any government department. This was reflected long ago in, for instance, studies by Bradford et al. (1969), or Levitt and Joyce (1987). It also often constitutes a serious source of controversy. Fletcher and Snee (1985) ranked what hinders output measurement in the public sector, by firstly considering problems in setting the objectives, and secondly problems in measuring the outputs themselves. Given our aims, and considering we are dealing with municipalities, we are not free from criticism. However, in many cases this criticism is leveled as a result of database limitations, or simply because of variables that can hardly be justified as outputs. On this point, we must acknowledge the virtues of our database, which provides not only quantities for each output, but also includes an indicator of quality, which is crucial for assessing local government performance. This indicator is based on subjective perceptions; hence, it takes values of 1, 2 or 3 depending on whether the observed quality is poor, average, or good. Then, it is averaged for services on which quality data are available.

The selection of outputs is based on the minimum services provided by each municipality. Specifically, all local authorities must provide public street lighting, cemeteries, waste collection and street cleaning services, drinking water to households, access to population centers, surfacing of public roads, and regulation of food and drink. In some cases we have to select proxies for these services. For instance, as pointed out by

Table 1
Output indicators based on the minimum services provided

Minimum services provided	Output indicators
Public street lighting	Number of lighting points
Cemetery	Total population
Waste collection	Waste collected
Street cleaning	Street infrastructure surface area
Supply of drinking water to households	Population, street infrastructure surface area
Access to population centres	Street infrastructure surface area
Surfacing of public roads	Street infrastructure surface area
Regulation of food and drink	Total population

De Borger and Kerstens (1996), population is assumed to proxy for the various administrative tasks undertaken by municipalities, but it is clearly not a direct output of local production.[11] It may also constitute a proxy for measuring those services that a particular municipality could provide at its own expense, going beyond the legal minimum. Other important outputs, such as provision of primary and secondary education, do not fall within the responsibilities of Spanish municipalities.

Table 1 enumerates the service that attempts to measure, or to proxy, each output indicator. The list of outputs, along with summary statistics, is presented in Table 2. Our choice has several parallels with the studies that focus on the efficiency of other European local governments, since their competencies are similar on the whole. Most divergences are found among education competencies, which differ greatly from country to country, and largely explain the comparatively lower weight of Spanish local government budget on Spanish public expenditure. The publicly available information does not go much beyond from the outputs in Table 2; however, we have also included variable Y_5, i.e., the registered surface area of public parks, since it constitutes a mandatory local facility for municipalities with a population of over 5,000. In addition to our selection of outputs, we could only have included the registered surface area of public buildings, provision of social services, and *lonja*—a traditional building where merchants gathered to sell their goods. The inclusion of any of these is problematic, since their usages are not entirely clear.[12]

We also include an interesting variable designed to measure not only the quantity but also the *quality* of the services provided. This unusual type of data is particularly informative for municipality output. Often, local governments cannot directly affect, at least in the short-run, the quantity of outputs such as surface areas of street infrastructure or public parks. However, it may have a decisive impact on their quality. The information available on this variable is of a categorical nature—the quality of the services offered is arranged in three classifications: Good, average or bad. In accordance with this, we have computed a single quality indicator for each municipality. Since we have the specific quantity of each variable falling within each category (e.g., we have information on the number of "good", "average" and "bad" lighting points), we can construct a single quality

[11] See also Ladd (1994), who addresses the fiscal impact of population growth.

[12] The distribution of responsibilities among central, regional, and local administrations may be found at the URL http://www.igsap.map.es/cia/dispo/ce_ingles_index.htm.

Table 2
Summary statistics for inputs and outputs (1995)

Inputs[a]	Mean	Std.dev.
Wages and salaries (X_1)	117,335.95	198,594.00
Expenditure on goods and services (X_2)	94,683.56	153,212.62
Current transfers (X_3)	17,492.75	35,637.52
Capital transfers (X_4)	1,587.99	7,005.55
Capital expenditure (X_5)	74,603.31	115,218.63
Outputs		
Population (Y_1)	4,730.18	7,607.53
Number of lighting points (Y_2)	594.52	1,038.66
Tons of waste (Y_3)	13,093.19	75,751.95
Street infrastructure surface area[b] (Y_4)	126,010.25	196,552.32
Registered surface area of public parks (Y_5)	15,957.28	32,503.52
Quality (Y_6)	2.632	0.308

[a]In thousands of 1995 pesetas (1 euro = 166.386 pesetas).
[b]In square metres.

weighted indicator for each $s = 1, \ldots, S$ municipality as follows:

$$\text{Quality}^s = Y_6^s = \frac{\sum_{i=1}^{I-3} \frac{\sum_{k=1}^{3} q_{ki} c_k}{\sum_{k=1}^{3} q_{ki}}}{I-3}, \tag{4}$$

where the number of outputs is restricted to $I - 3$ (since the quality in itself constitutes an output, and we only have this information for variables Y_2, Y_4, and Y_5), q_{ki} is the quantity of i output of quality k, and c_k is the quality indicator, $k = 1, 2, 3$. Hence, our model includes both production and quality variables, enabling a joint assessment of both efficiency and quality to be made.[13]

Our selection of inputs is based on budgetary variables that reflect municipality costs. These are implemented rather than forecasted expenditures, given the usual discrepancies amongst these figures (forecasts often underestimate expenditure and overestimate revenues). In contrast to many other studies, we included capital measures—capital expenditures and capital transfers. In particular, our definition of inputs reflects the economic structure of Spanish local government expenditures, whose specifics are reported by Spanish legislation,[14] which considers three basic categories: Current—ordinary—expenditures, capital expenditures, and financial expenditures. Among them, current expenditures are further divided into four categories, which account for: (i) personnel expenditure; (ii) current goods and services expenditures; (iii) financial expenditures; (iv) current transfers. Capital expenditures are also broken down into either capital transfers (X_4), or real investments (X_5).

[13]We also considered the proposal set out by Banker and Morey (1986), which involves breaking down the quality variable into two categorical variables, d_1 and d_2. Thus, for unit s, the values taken by d_1 and d_2 are $d_{s1} = d_{s2} = 0$, if the quality is bad, $d_{s1} = 1$ and $d_{s2} = 0$, if the quality is average, and $d_{s1} = d_{s2} = 1$, if the quality is good. Although this approach is more standard, results did not vary.

[14]See *Orden Ministerial*, September 20th, 1989.

3.2. Results

We estimate a common frontier, and results are presented in Table 3. It shows not only simple summary statistics such as mean and standard deviation but also additional statistics, which provide further insights into the distributions of efficiency scores, since their distributions (generally skewed) suggest it is of interest to use other summary statistics for a better understanding of what efficiency indices reveal.

In the convex reference set (DEA-VRS), only 32 (7.73%) out of 414 observations were found to be cost efficient—i.e., had an overall efficiency measure of 1. Table 3 also shows that the highest levels of efficiency are found amongst the most populated municipalities. In contrast, overall cost efficiency is quite low for the smallest group of municipalities (on average, 42.8%). In order to interpret the results, we should also bear in mind that municipalities with populations of under 5000 do not have to disclose as much accounting information as larger ones, which could lead to less effort being made to monitor expenditures.

The results as a whole reveal persistent growth in the overall cost efficiency with the size of the municipalities. Yet this fact is unrelated to the economies of scale since the reference technology assumes variable returns to scale (VRS). This suggests that for large municipalities there are decreasing returns to scale reflecting two realities: (i) public services offered in large cities are more complex than in small towns; and (ii) there are technical diseconomies of scale probably due to the existence of agglomeration diseconomies. Once decreasing returns to scale are controlled for, small towns are far and large cities are closer (and less dispersed) from their respective cost frontiers. Differences among sizes may be related to the quality of public management. In large cities, public managers and technical staff employ management tools similar to those used in private firms (financial and budgetary control, contracting out of some services, etc.), whereas in small municipalities the diffusion of innovation in public management is slower.

The results of the cost efficiency measure corresponding to the nonconvex FDH frontier are analogous: A growing average coefficient with size. Yet we must note that the percentage of efficient observations increases dramatically, suggesting that the situation represented by observation C in Fig. 1 describes very closely what the general case is for Valencian municipalities. Apart from large municipalities, the FDH evaluation also shows that approximately half the small municipalities and around 25% of the medium-sized municipalities have cost excesses.

In order to gain enhanced insights on the specifics of the distributions, we provide further summary statistics. As expected, skewness is found to be positive for DEA-VRS and negative for FDH, i.e., probability mass tends to concentrate rightwards when many observations reach the upper bound—which is the FDH case. Analysis by sizes corroborates this view, except for the DEA-VRS case for municipalities over 20,000 inhabitants, due to the large number of observations found to be efficient here. Measures for the ‘‘peakedness’’ of the distributions (kurtosis) are also heavily conditional on the number of efficient units to the point that, under the FDH evaluation, municipalities with over 20,000 inhabitants show an abnormally high value.

We also describe efficiency scores' distributions using Tukey's boxplots (see Fig. 2), which are particularly informative since both DEA-VRS and FDH efficiency scores are bounded by zero and unity, and in both cases a mass of observations reaches the upper bound, typically yielding skewed distributions. Furthermore, boxplots provide a more

Table 3
Summary statistics for overall cost efficiency by population sizes ($N = 414$)

Inefficiency	Size class	Mean	Std.dev.	Median	Skewness	Kurtosis	Minimum	Maximum	% of efficient observations
DEA-VRS	All sizes	0.531	0.246	0.492	0.373	−0.746	0.014	1.000	32/414(7.73%)
	$POP < 1000$	0.428	0.232	0.382	1.016	0.456	0.014	1.000	9/161(5.59%)
	$1000 \leqslant POP < 5000$	0.495	0.202	0.472	0.409	−0.091	0.080	1.000	4/144(2.78%)
	$5000 \leqslant POP < 20,000$	0.688	0.191	0.673	0.042	−0.790	0.260	1.000	8/82(9.76%)
	$POP \geqslant 20,000$	0.861	0.159	0.933	−0.942	−0.084	0.470	1.000	11/27(40.74%)
FDH	All sizes	0.902	0.193	1.000	−2.172	4.004	0.141	1.000	289/414(69.81%)
	$POP < 1000$	0.849	0.231	1.000	−1.567	1.533	0.141	1.000	91/161(56.52%)
	$1000 \leqslant POP < 5000$	0.919	0.187	1.000	−2.364	4.388	0.227	1.000	111/144(77.08%)
	$5000 \leqslant POP < 20,000$	0.947	0.107	1.000	−2.117	3.749	0.537	1.000	61/82(74.39%)
	$POP \geqslant 20,000$	0.994	0.029	1.000	−5.196	27.000	0.849	1.000	26/27(96.30%)

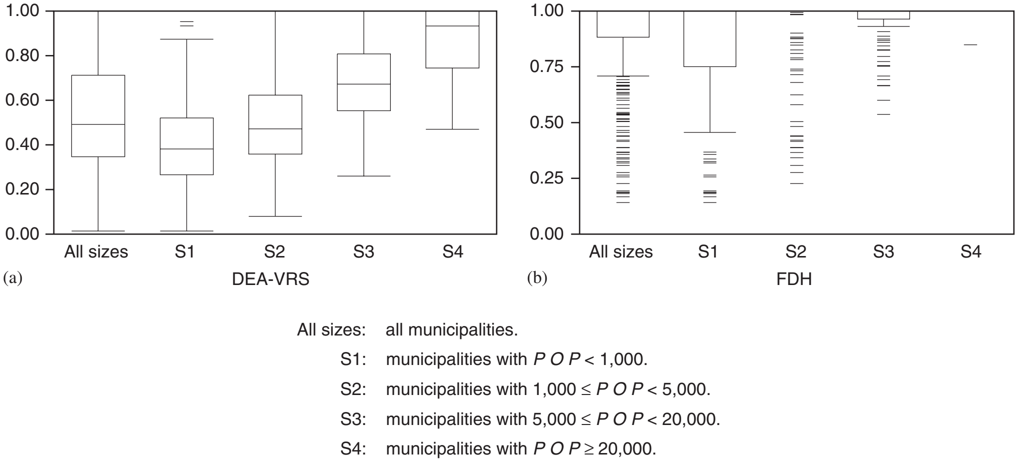

Fig. 2. Tukey's boxplots of overall cost efficiency scores by size classes.

scrupulous illustration of distributions, encoding further information on the features hidden by data such as the existence of outliers. Boxplots are shown in Fig. 2 for each size class and the interpretation is straightforward. For each size class, the box represents the 50% mid-range values of the efficiency scores. The length of each box represents the interquartile range (IQR). The line breaking down each box represents the median. The whiskers define the natural bounds of the distributions, i.e., the mean$\pm IQR$, while the short lines represent outliers lying outside the natural bounds.

It can be observed from the IQRs that the distribution of efficiency scores steadily shrinks with size. The most inefficient observations are found amongst the smallest size classes.[15] On the other hand, most of the largest municipalities are efficient, and the only

[15]Extreme minimum values (i.e., the very inefficient municipalities) reflect some problems regarding the non-audited information on particular small municipalities. When first verifying the results, we controlled for the 'super-efficient' (Andersen and Petersen, 1993) extreme observations, but we decided to keep extreme inefficient

FDH inefficient municipality is classified as an outlier. Therefore, although it seems that larger municipalities are closer to the frontier due to their greater resources (qualified staff, better information technology resources, etc.), we must bear in mind the scale effect, i.e., we are comparing equals.

4. Efficiency explanatory variables

The overall cost efficiency DEA-VRS and FDH measures (OE^s) yield only a first-stage measure of relative efficiency. What is not known is the reason for variations in such efficiency patterns, and it is clear that computing individual cost efficiency scores may not be enough for either consulting purposes or for policy analysis. Accordingly, a second-stage analysis is called for, as efficiency may be affected not only by inadequate management but also by exogenous factors beyond the control of each local government.

In our search for explanatory variables we considered the information provided by the Valencian Audit Institution on different financial and fiscal policy variables, such as tax revenue, grants, or financial liabilities. We also considered a political variable: The percentage of votes attained by the governing party in each municipality. Similarly to De Borger et al. (1994), the choice of independent variables was prompted both by the literature and by taking into account the institutional framework of local government financing in Spain. Specifically, the yardstick of selecting variables related to the income structure of each municipality was the main driver behind our set of determinants.

More specifically, the literature on Spanish local government deals primarily with fiscal policy issues. A major concern is that of how local governments should be financed. In the Spanish public sector, the most important tool, which has a decisive impact on both short- and long-run decision making, is the budget, and we take into account the financial perspective, by considering financial deficit due to cost excess relative to revenues as another explanatory variable for cost inefficiency.

However, our list of determinants is not exhaustive for several reasons. Other studies have included, for instance, personal income or the share of the adult population with higher education qualifications (De Borger et al., 1994), information we do not have access to. In addition, as mentioned previously, the competencies of Spanish local governments are different from those of their European peers, and therefore the information available differs. Summary statistics for each specific explanatory variable, whose likely impacts on efficiency are commented on below, are provided in Table 4.

Therefore, as part of the current revenues, the first variable to be considered is tax revenue per capita (*TAXES*). We may hypothesize a negative impact on efficiency, as it seems reasonable that a local government that is highly capable of generating current revenues would be less motivated to manage them efficiently. This idea also comes from the property rights and principal-agent literature, which outlines several reasons why politicians and public managers may lack proper incentives to effectively audit and control expenditures (De Borger and Kerstens, 1996). On the other hand, high taxes may increase voters' awareness of control of public expenditure (Davis and Hayes, 1993). Yet such an

(*footnote continued*)
values provided they do not shift the frontier. An alternative hypothesis justifies the existence of those extreme inefficient units (the provision of additional municipal services beyond the minimum requirement) although this phenomenon is mostly found among large municipalities.

Table 4
Summary statistics for explanatory variables (1995)

Variable name	Variable description	Mean	Std. dev.
TAXES	log(tax revenue/population)	3.569	0.495
GRANTS	log(grants received/population)	3.510	0.683
SELFG	log(own revenues/population)	3.663	0.503
SECLOAN[a]	log(financial liabilities/population)	1.993	2.543
DEFICIT	total expenditure/total revenue	0.985	0.136
VOTES	(votes received by governing party candidates, 1995)/population	0.433	0.453

[a]Logs were computed only for those municipalities with nonzero financial liabilities.

awareness might, instead, worsen efficiency, leading to a situation in which bureaucrats may prefer more visible rather than less tangible inputs since they are easier to justify in the appropriation process (Lindsay, 1976). For instance, police cars may be more visible than training (Grosskopf and Hayes, 1993).

The second fiscal policy variable to be considered is current transfers, or grants (*GRANTS*), representing transfers received from higher levels of government, that depend on variables such as number of inhabitants, number of schools, etc. More than 70% of these grants are unconditional grants for current spending. Following Silkman and Young (1982), this variable may be regarded as having a negative association with efficiency, as the cost of inefficient behavior is increasingly shared by a broader constituency.

Thirdly, self-generated revenues (*SELFG*) accounts for revenues other than *TAXES*, such as those from taxes on personal wealth and local government property transfers, which are all self-generated by each municipality. Some authors label this category as ‘‘patrimonial revenues’’ (Vela, 1996). Given its similar nature, the likely impact on efficiency is expected to resemble that of *TAXES*, yet in this case the intensity of monitoring by the citizens effect is partially lessened (Davis and Hayes, 1993). In addition, the relevance of this variable is minor compared to that of *TAXES* due to its lower magnitude.

Local governments with lower fiscal revenue capacity will probably be impelled to follow alternative paths to raise revenues, such as issuing securities or making loans. Both activities fall under the general category of financial liabilities. To proxy for these effects, we include the variable *SECLOAN*, which measures income generated by issuing debt and making loans. Its impact on efficiency is not a priori obvious. We may hypothesize that the municipalities that issue bonds and securities are those unable to raise revenues via taxes. In such a case, the taxpayers' incentives to effectively control expenditures may be low and, consequently, we might expect a negative relationship with efficiency. On the other hand, this variable is assumed to be inversely related to *TAXES* and *SELFG*, as the municipalities with higher levels of issued securities are possibly those unable to generate revenues in more traditional ways—i.e., taxes. In such a case, the impact on efficiency would be positive.

Another variable included is total expenses divided by total revenues (*DEFICIT*), which proxies the idea of deficit—although defined in a different way—and is expected to be negatively associated with efficiency. As deficit increases, we may face a higher social awareness to encourage its reduction; in such a case, local governments may adopt strategies to enhance efficiency. Municipalities carry out activities within the framework of a financial contingency that requires a balance between their receipts and expenditures, the rupture of which can put them in a situation of financial vulnerability. This may be due to inefficient management, or simply to structural insufficiency of resources. Thus, one may hypothesize an inverse relationship between efficiency and financial vulnerability. If we define financial vulnerability as the inability of a municipality to face its present and future financial commitments as they fall due, we may consider deficit a good proxy for it.

Apart from these budget-based fiscal policy variables, we include a political variable that represents the relative importance of votes held by the governing party (*VOTES*). If the governing party has an absolute majority, other parties or coalitions will possibly face greater difficulties in effectively controlling expenditures. Efficiency might not be the criterion used when awarding contracts for public works or building permits, hiring new workers, or setting local managers' wages, with the possible result that court orders are the only way to control this expenditure. In such a case the expected association between *VOTES* and efficiency would be negative. However, one might alternatively hypothesize a positive association since city managers would have an incentive to lower costs and increase efficiency in order to ensure their job security.[16]

4.1. Some problems with previous two-stage studies ideas

Previous studies performing two-stage analysis have considered estimation techniques such as OLS, or the Tobit censored regression model. An exhaustive file of previously published two-stage studies is provided by Simar and Wilson (2006). Unfortunately, these methods have some disadvantages related to the (likely) correlation between the efficiency scores and the explanatory variables (see De Borger and Kerstens, 1996). Other authors (see Charnes et al., 1988) have addressed the violation of normality assumption by employing the L_1-metric regression. Alternatively, Lovell et al. (1994) construct modified DEA efficiency scores (MDEA efficiency scores) which are only bounded by zero, following Andersen and Petersen (1993), whereas Ray (1991) considers a variant of corrected least squares in the second stage.

Simar and Wilson (2006) also point out that the efficiency scores generated by DEA, in the same way as other nonparametric techniques for efficiency measurement such as FDH, are serially correlated. Therefore, they present an alternative approach to succeed in two-stage settings where efficiency is estimated in the first stage and estimated efficiencies "are viewed as representing environmental variables". Specifically, they suggest a bootstrap methodology to address both the issues of describing the underlying data-generating process (DGP)—so as to understand what is being estimated in the two-stage approaches—and the existing serial correlation among DEA efficiency estimates. The basic underlying ideas, related to the dependency of the efficiency scores on each other in the statistical sense, had already been forcefully made by Harker and Xue (1999), who also suggested a bootstrap methodology to address the problem, in a similar way to Hirschberg and Lloyd (2002). However, Simar and Wilson (2006) pointed out that both studies are flawed due to their choice of a naive bootstrap rather than a *smoothed* bootstrap that would take into account the intricacies of efficiency scores' distributions.

[16]Obviously, the most desirable variable would be a dichotomous variable, taking values one or zero depending on whether an absolute majority governs the council. Unfortunately, that type of information is unavailable to date.

4.2. Kernel contributions: Nonparametric kernel regression and bivariate density functions

Here we consider an alternative set of techniques for disentangling the variables representing the factors likely to impact on efficiency performance of local governments. Our methods are not exactly coincidental with those employed by Simar and Wilson (2006) yet they share their most relevant underpinnings—namely, both have a clear nonparametric flavor, and both overcome most of the weaknesses found in previous two-stage studies. Specifically, we consider both nonparametric regression and nonparametric density estimation, which are less powerful in terms of prediction yet extremely informative for explanatory purposes. They are particularly useful in the case of local government studies where the number of observations is usually large. They build on previous work from authors such as Deaton (1989), DiNardo et al. (1996) or Marron and Schmitz (1992). Their main advantage is their nonparametric nature, conferring them an ability to provide easily comprehended graphical descriptions of the data that are directly informative about the problem in hand. In addition, this methodology is relatively robust to outlying observations.

4.2.1. Nonparametric kernel regression

In econometrics, the assumption of statistical adequacy or correct specification has been a constant concern for some time. This concern is present in our setting for the reasons set out above. In fact, as Haavelmo (1944) asks, "what is the use of testing, say, the significance of regression coefficients when maybe the whole assumption of the linear regression equation is wrong?" (Haavelmo, 1944, p. 66). The peculiar structure of our data, made explicit in Fig. 2, suggests the use of nonparametric techniques, which allow us not only to relax the assumptions of the underlying model, but also help in corroborating the goodness (or lack) of fit of parametric specifications.

The basic purpose of a regression analysis is to study how a variable Y, in our case efficiency estimates, responds to changes in another variable X. Then, the underlying trend would be a function such as

$$f(X) = E(Y|X). \tag{5}$$

A starting fit for (5) would be the simple linear regression model,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \ldots, n, \tag{6}$$

with errors $\boldsymbol{\varepsilon}$ (usually) taken to be independent and identically distributed, with zero mean and variance σ^2. If that model truly represented reality, then OLS estimates of $\boldsymbol{\beta}$ could be calculated, and inference and prediction followed (Simonoff, 1996). However, what if (6), or another more complex fit we might consider, does not truly represent reality? Or, what if assumptions, as stated above, are not met?

The more general nonparametric regression model avoids both issues. Specifically, it considers some unspecified m "smooth" function that needs to be estimated from (X, Y),

$$y_i = m(x_i) + \varepsilon_i, \tag{7}$$

where the regression curve $m(x)$ is the conditional expectation $m(x) = E(Y|X = x)$, with $E(\varepsilon|X = x) = 0$, and $V(\varepsilon|X = x) = \sigma^2(x)$ is not necessarily constant—i.e., we avoid both the misspecification and assumption violation problems.

So to prevent convoluting notation, these two variables Y and X will be labeled EFF and z, respectively, where EFF may refer to either overall cost estimated through DEA-VRS or FDH. If we have data on these variables, then smoothing of a data set $\{(z_s, EFF_s)\}_{s=1}^{S}$ involves the approximation of the mean response curve m in the regression relationship:

$$EFF_s = m(z_s) + \varepsilon_s, \quad s = 1, \ldots, S, \tag{8}$$

where ε_s represents the error term. The less we know about the nature of m, the more desirable a nonparametric estimation approach is. These methods impose a minimum of structure on the regression function. Therefore, the main advantage of this technique is that the data are allowed to choose the shape of the function, and there is nothing that forces the points to lie along a straight line, or along a low-order polynomial (Deaton, 1989). Nonparametric methods are also known as *smoothing* methods, i.e., methods aimed at sanding away the rough edges from a set of data or, in other words, to remove data variability that has no assignable cause and to thereby make systematic features of the data more apparent (Hart, 1997).

Nonparametric regression may be accomplished by several alternatives. The most popular ones are kernel estimators—among which may be considered either the Nadaraya–Watson estimator, or the Gasser–Müller estimator (Fan and Gijbels, 1996). In addition to this, local polynomial estimation may be considered, which basically consists of a generalization of the Nadaraya–Watson estimator; specifically, the Nadaraya–Watson estimator corresponds to a local constant least squares fit. In contrast, local polynomial regression generalizes it by fitting a polynomial in a neighborhood of x; the basic difference is that the local linear fit reacts more sensitively on the boundaries of the fit.

The basic idea of smoothing lies in a *local averaging procedure*, which is constructed in such a way that it is defined only from observations in a small neighborhood around z, since EFF-observations from points far away from z will have, in general, very different mean values.[17]

More formally, the procedure can be defined as:

$$\widehat{m}(z) = S^{-1} \sum_{s=1}^{S} W_{S,s}(z) EFF_s, \tag{9}$$

where $\{W_{S,s}(z)\}_{s=1}^{S}$ denotes a sequence of weights which may depend on the whole vector $\{z_s\}_{s=1}^{S}$.

An alternative way of looking at the local averaging formula (9) is to suppose that the weights $\{W_{S,s}\}$ are positive and sum to one for all z, i.e., $S^{-1}\sum_{s=1}^{S} W_{S,s}(z) = 1$. Then $\widehat{m}(z)$ is a *least squares estimate* at point z since we can write $\widehat{m}(z)$ as a solution to the following minimization problem:

$$\min_{\theta} S^{-1} \sum_{s=1}^{S} W_{S,s}(z)(EFF_s - \theta)^2 = S^{-1} \sum_{s=1}^{S} W_{S,s}(z)(EFF_s - \widehat{m}(z))^2. \tag{10}$$

According to formula (10), we realize that the basic idea of local averaging is equivalent to the procedure of finding a local weighted least squares estimate.

[17] We have left out most details on nonparametric regression. See, for instance Härdle (1990), Härdle and Linton (1994), or Pagan and Ullah (1999), for good expositions.

The Nadaraya–Watson estimator of $m(z)$ as a local average of EFF_s is written as:

$$\widehat{m}(z) = \frac{\sum_{s=1}^{S} EFF_s K_h(z - z_s)}{\sum_{s=1}^{S} K_h(z - z_s)} \tag{11}$$

which may be decomposed as

$$\widehat{m}(z) = \sum_{s=1}^{S} EFF_s W_s(z) \tag{12}$$

and

$$W_s(z) = \frac{K_h(z - z_s)}{\sum_{s=1}^{S} K_h(z - z_s)}, \tag{13}$$

where $K_h(\bullet)$ is a kernel function satisfying different properties.

The more general (local) polynomial fit differs by considering a Taylor expansion of the unknown conditional expectation function $m(\bullet)$ for t in a neighborhood of the point z, i.e.:

$$m(t) \approx m(z) + m'(z)(t - z) + \cdots + m^p(z)(t - z)^p \frac{1}{p!}. \tag{14}$$

Thus, Eq. (10) gains some complexity, since scalar θ turns into a vector $\beta = (\beta_0, \beta_1, \ldots, \beta_p)'$:

$$\min_{\beta} \sum_{s=1}^{S} W_{S,s}(z)(EFF_s - \beta_0 - \beta_1(z_s - z) - \cdots - \beta_p(z_s - z)^p)^2, \tag{15}$$

The estimator $\widehat{m}_p(z)$ is the intercept term $\widehat{\beta}_0$, since expressions (14) and (15) result in $m(z) \approx \beta_0(z)$. As may be inferred from the above formulæ, both local polynomial regression and the Nadaraya–Watson estimator coincide when $p = 0$. Results may differ slightly for $p > 0$, although only if problems at boundaries are present. Since results did not show marked differences for both approaches, only those relative to the Nadaraya–Watson estimator are provided.

The choice of kernel, $K_h(\bullet)$, consists of several alternatives. For reasons of simplicity, we used a Gaussian kernel, whose formula is based on the Gaussian density function

$$K_h(z - z_s) = \frac{1}{(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\left(\frac{z - z_s}{h}\right)^2\right], \tag{16}$$

where the quantity h is the bandwidth or smoothing parameter and controls the smoothness of $\widehat{m}$—likewise the window width controls the smoothness of a moving average. Indeed, our choice of kernel is related to the right amount of smoothing, i.e., the choice of bandwidth. This is the most crucial decision in nonparametric regression. Every smoothing method must be tuned by a smoothing parameter which balances the degree of fidelity to data against the smoothness of the estimated curve. Although more popular techniques are available, such as cross-validation, we focus on the more up-to-date approaches. Specifically, we follow the contribution by Ruppert et al. (1995), who suggest plug-in rules. As stated by DiNardo et al. (1996), plug-in methods do not exhibit the discretization problems associated with cross-validation. Obviously, if we want smoothing to be merely a descriptive device, the "by eye" technique may be satisfactory. However, this approach is inadequate in many circumstances. Thus, the practical implementation of any scatterplot, or nonparametric regression smoother is greatly enhanced by the availability of a reliable rule for automatic selection of the smoothing parameter.

4.3. Nonparametric bivariate density estimation

Unfortunately, the information provided by nonparametric regression is somewhat skeletal. To fill out this information, we provide nonparametric estimation of the joint density functions of efficiency and each of the explanatory variables. Its basic ideas closely resemble those underlying scatterplot smoothing (or nonparametric regression). The main difference is that now observations are counted and weighted not in an interval band around each point, but in a two-dimensional elliptical band. For clearer interpretation, we also provide contour maps, where points linked by a contour have the same density and the contours are equally spaced.

To estimate joint density functions, in which the OX axis represents the explanatory variable and the OY axis represents the dependent variable, i.e., efficiency, we also chose nonparametric techniques. Again, one of the most popular alternatives is kernel smoothing. The basics are quite similar to nonparametric regression. While the approach to smoothing data is exactly the same (kernel smoothing), the other two decisions are not strictly coincidental. Regarding the choice of kernel, and following Wand and Jones (1995) and, in his applied work, Deaton (1989), here we consider the Epanechnikov kernel. We proceed with this for the sake of efficiency. Its expression is as follows:

$$\widehat{f}(\mathbf{x};\mathbf{H}) = S^{-1}\sum_{s=1}^{S} K_{\mathbf{H}}(\mathbf{x}-\mathbf{v}_s), \tag{17}$$

where $\mathbf{v}_s = (z_s, EFF_s)$, $\mathbf{x} = (x_1, x_2)$ is the point of evaluation, $\mathbf{H}$ is a $d \times d$ (in our case, 2×2) bandwidth matrix, and $K_{\mathbf{H}}$ is a kernel function:

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2}K(\mathbf{H}^{-1/2}\mathbf{x}). \tag{18}$$

We consider $\mathbf{H} \in \mathscr{D}$, where $\mathscr{D} \subseteq \mathscr{F}$ defines the subclass of diagonal positive definite matrices 2×2 ($\mathscr{D} = \{\text{diag}(h_1^2, h_2^2) : h_1, h_2 > 0\}$), and $\mathscr{F}$ defines the class of positive definite and symmetric matrices. Hence, for $\mathbf{H} \in \mathscr{D}$, we have $\mathbf{H} = \text{diag}(h_1^2, h_2^2)$.

According to the above statements, provided $\mathbf{h} = (h_1, h_2)$, the *bivariate* density function to be estimated is

$$\widehat{f}(\mathbf{x};\mathbf{h}) = (Sh_1h_2)^{-1}\sum_{s=1}^{S} K\left(\frac{x_1 - z_s}{h_1}, \frac{x_2 - EFF_s}{h_2}\right). \tag{19}$$

Finally, we chose the Epanechnikov kernel, whose expression for the bivariate case is as follows:

$$K_e(\mathbf{x}) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-\mathbf{x}^{\mathrm{T}}\mathbf{x}) & \text{if } \mathbf{x}^{\mathrm{T}}\mathbf{x} < 1, \\ 0 & \text{otherwise,} \end{cases} \tag{20}$$

where c_d is the volume of the d-dimensional unit sphere: $c_1 = 2$, $c_2 = \pi$, $c_3 = 4\pi/3$, etc.

Unfortunately, we have to deal with a problem that has still not been satisfactorily addressed in the bivariate nonparametric density estimation literature—namely, the choice of bandwidth. In this case, the state of the art is in a much more preliminary stage than in the univariate or the nonparametric regression cases. One of the most up-to-date contributions is the work by Wand and Jones (1994), also based on plug-in methods, which is precisely our choice. Although previous research considers by-eye criteria (Deaton, 1989), we do not completely agree with this view, as it partly undermines the impartiality of the analysis. When possible, data-driven methods are preferable.

This complementary approach is extremely informative, as it provides clear awareness of what exactly underlies each regression line. Hence, we know exactly the probability mass or, put another way, the empirical evidence supporting each claimed sign for the differing associations. The precise interpretation would suggest that when a negatively sloped regression line is observed, supporting an inverse association between a certain variable and efficiency, its corresponding bivariate density function should exhibit probability mass also concentrated along a ‘‘hypothetical’’ negative slope in each contour map. In other words, we would observe a collection of (possibly) consecutive bumps, possibly with high peaks, with probability vanishing at other locations of the figure. In the likely event that these bumps were firmly concentrated along the negative slope, it would indicate empirical evidence to support our prediction. This outcome would constitute a parallelism with the case of (parametric) linear regression when a high value for the t-statistic is achieved. On the other hand, if probability mass in the bivariate figure does not follow the path determined by the nonparametric regression line, the scenario would mirror the case of linear regression when a poorer value for the t-statistic is obtained.

Thus, the main advantage of our approach over the aforementioned approaches (both parametric two-stage studies and also nonparametric methods such as those using resampling techniques) is that, in our case, the term nonparametric refers to the flexible functional form of the regression curve, and neither the error distribution nor the functional form of the mean function are prespecified. While bootstrap approaches may be more informative in terms of predictive power, ours is more graphical and informative for explanatory purposes. However, if this were required, semiparametric methods such as those reviewed by Ruppert et al. (2003) combine the simple additive structure of the parametric regression model with the flexibility of the nonparametric approach (Härdle et al., 2004); as a result, they provide a reasonable balance between predictive and explanatory aims. We do not enter into the specifics of semiparametric regression here since it goes beyond the scope of our study; however, they are extremely closely related to its underpinnings, since when estimating semiparametric models we must use nonparametric techniques.[18]

In any case, either our approach or others based on resampling techniques (primarily Simar and Wilson, 2006) perform much better than any parametric model in general and in particular, the Tobit model—or related limited-dependent models. As suggested by Härdle et al. (2004), limited-dependent models, like other parametric models, are based on rather strong functional forms and distributional assumptions which, in addition, are often difficult to justify from an economic theory approach. Obviously, should parametric estimates be unaffected by model violations, no problem arises. Unfortunately, as shown

[18]If we considered, for instance, a function describing log wages (Y) as dependent from schooling ($SCHOOL$) and labor market experience (EXP) we would be interested in estimating $E(Y|SCHOOL, EXP) = \alpha + g_1(SCHOOL) + g_2(EXP)$, $g_1(\bullet)$ and $g_2(\bullet)$ being two unknown, smooth functions and α an unknown parameter. In these circumstances, semiparametric techniques would make use of nonparametric regression estimators to estimate the unknown functions $g_1(\bullet)$ and $g_2(\bullet)$ (Härdle et al., 2004).

by Härdle et al. (2004) for the Logit model (easily extendable to the Tobit setting), asymptotic statistical theory shows that violating the assumptions of the model leads parameter estimates to be inconsistent, i.e., the maximum-likelihood estimator (MLE) will not converge to the *true* parameter value in probability. In their application, Härdle et al. (2004) rely on simulations to collect evidence of small-sample performance in the presence of misspecification, and conclude that the MLE underestimated the true value by 11%.[19]

4.4. Results

Figs. 3a–f are nonparametric regressions of the efficiency scores on each of the explanatory variables estimated by kernel smoothing, using the Nadaraya–Watson estimator.[20] The solid line represents DEA-VRS, whereas the dashed line denotes FDH. They generally support the hypothesized signs of the regressions, although to varying extents. Regarding fiscal policy variables, we observe that efficiency decreases with tax revenues, (*TAXES*), self-generated revenues (*SELFG*), and deficit (*DEFICIT*), as expected. Grants (*GRANTS*) offer a clearly decreasing pattern, which is robust for both DEA-VRS and FDH. In the case of loans and issued securities (*SECLOAN*), where the sign was hypothesized to be positive, the pattern is less clear, although the visual snapshot partly suggests a negative relationship.

If we examine the particular details, we observe that the explanatory power of the grants variable (*GRANTS*) offers a steadily negative relationship which, in addition, does not exhibit remarkable ups and downs. However, this behavior shows some peculiarities, since disparities between DEA-VRS and FDH are more ostensible in the upper and middle ranges; more affinities are more apparent among inefficient firms. Yet robustness is guaranteed, since both regression curves parallel each other greatly. Our findings coincide with those of Silkman and Young (1982), or De Borger and Kerstens (1996), who suggest that grants may not only encourage local service provision, but also stimulate inefficiency, as the cost of inefficient behavior is increasingly shared by national, or regional, taxpayers. In the case of self-generated revenues (*SELFG*), the association is also negative, although it exhibits some particularities, since linearities are more present in the middle range, whereas the situation is less clear at boundaries. A plausible explanation, examined below, could come from the existence of extreme observations. This variable is highly related to tax revenues (*TAXES*). In fact, they only differ because *TAXES* accounts solely for revenues entirely generated via taxes.

In contrast, our variable with the strongest political content, i.e., the governing party share of votes (*VOTES*) does exhibit a fuzzy pattern (Fig. 3f), and no conclusion may be drawn—at least at first sight. The behavior is not entirely coincidental for DEA-VRS and FDH, the latter offering a much smoother pattern—although this finding is clearly driven by the difficulties in finding an appropriate smoothing parameter for this particular variable. Accordingly, we cannot state that there is sufficient empirical evidence to corroborate the hypothesis that municipalities managed by governments with a higher

[19] Simar and Wilson (2006) also ran Monte Carlo trials, coming to a similar conclusion: The model estimated was misspecified, finally concluding that "this approach results in catastrophe".

[20] The *OX*-axis in each sub-figure contains information depending on the explanatory variable considered. Table 4 states precisely what each explanatory variable stands for.

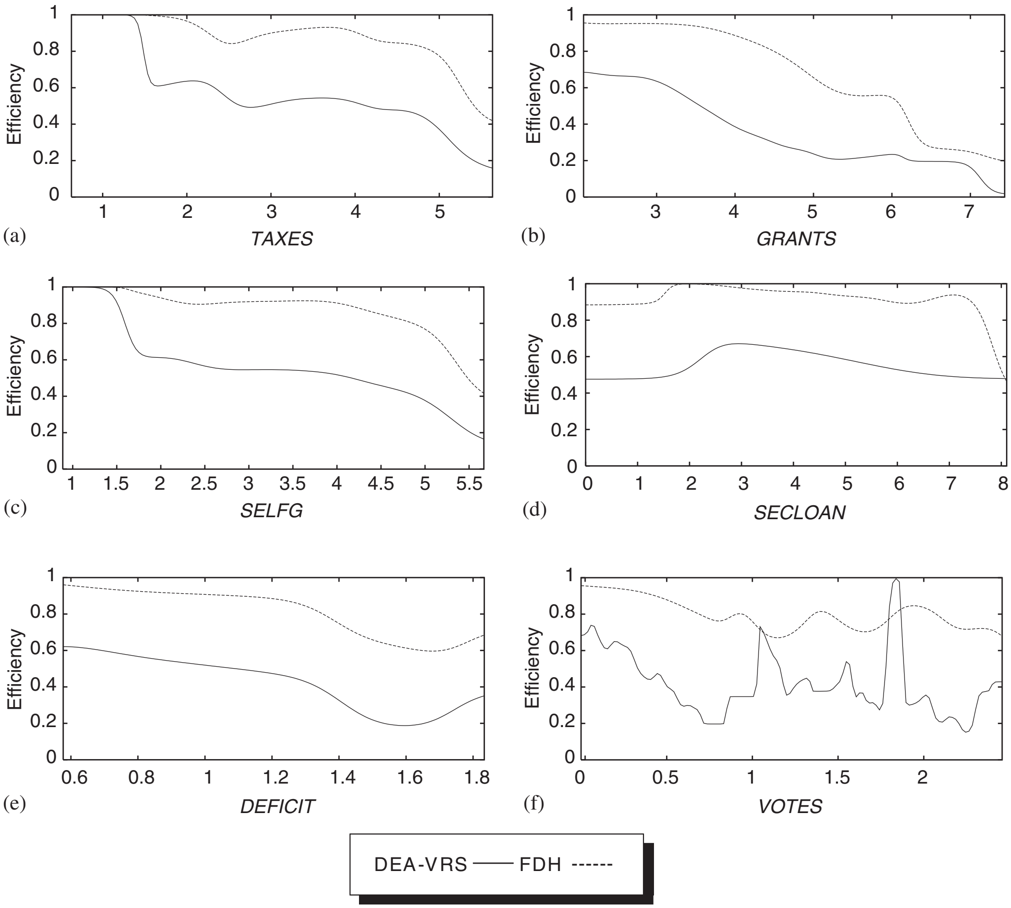

Fig. 3. Efficiency determinants, Nadaraya–Watson nonparametric regression (overall cost efficiency).

percentage of votes, more unlikely to face monitoring by other parties, have fewer incentives to manage their resources efficiently.

Joint density functions for DEA-VRS overall cost efficiency and each of the explanatory variables are shown in Fig. 4.[21] Their respective contour maps are given in Fig. 5. A paradigm is constituted by Fig. 4e. Its nonparametric regression counterpart, i.e., Fig. 3e, shows an unstable pattern, especially on the right-hand side, as the relationship between efficiency and deficit is negative except for those municipalities with higher deficit, which shift the regression line upwards. But Fig. 4e, together with its contour plot in Fig. 5e, reveal that the units driving the regression line to shift upwards are very few, as probability mass in the 3-d plot is virtually nonexistent on its right-hand side.

Similarly, the nonparametric regression for securities and loans issued by municipalities exhibits no clear pattern, especially for DEA-VRS (see Fig. 3d). Yet its 3-d plots

[21]We have restricted the 3-d analysis to DEA-VRS, to avoid over-straining the limits of space. Results, similarly to the case of nonparametric regression, proved to be robust.

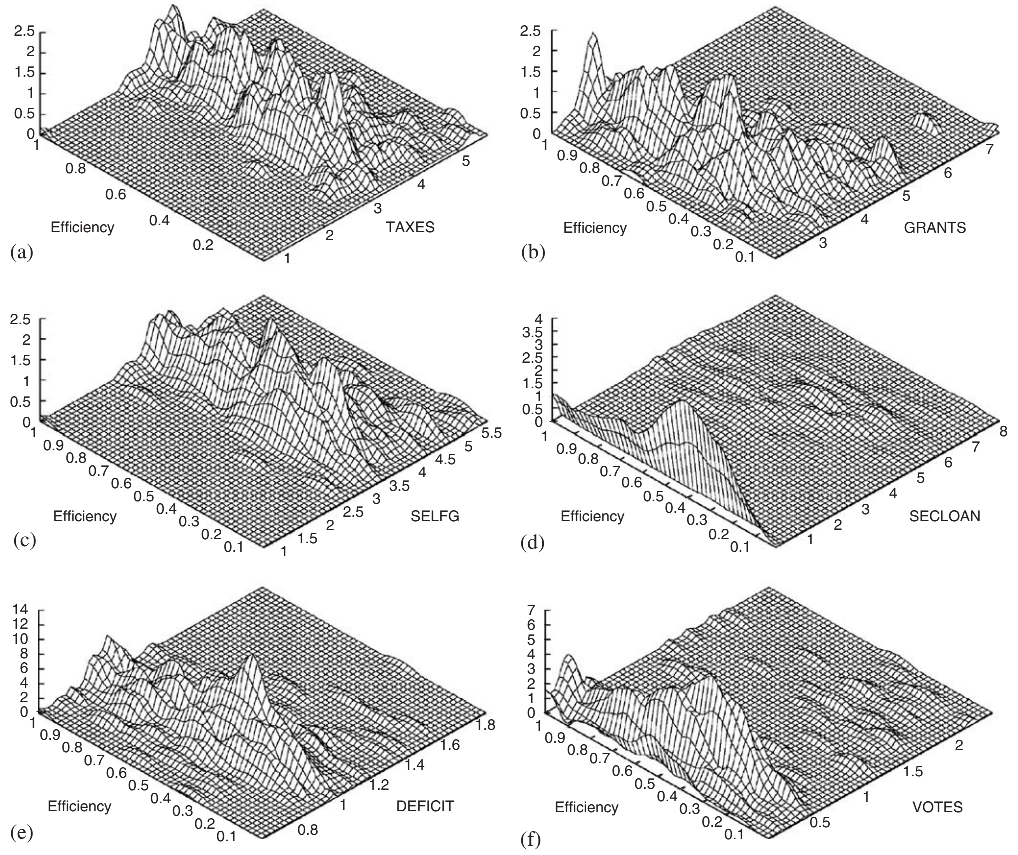

Fig. 4. Efficiency determinants, joint densities, overall cost efficiency.

counterpart (Figs. 4d and 5d) unmask what this is determined by: We cannot infer any link for most municipalities, as they simply do not issue securities. For others, a negative association is observed. But the probability mass supporting the latter assertion is quite scarce, as revealed by a low number of contours (see Fig. 5d).

The remaining fiscal policy variables—*TAXES* (Figures 4a and 5a), *SELFG* (Figs. 4c and 5c), and *DEFICIT* (Figs. 4e and 5e)—widely confirm what was revealed by regression, as probability mass has a tendency to concentrate along the hypothetical negative slope diagonal. However, in all cases we notice a remarkable amount of multi-modality. Related to this, we face the issue of bandwidth choice, which is critical in nonparametric kernel density estimation, and has not yet been fully addressed in the bivariate case.[22]

Although not displayed, there is a greater difficulty in fitting a regression line of Y (or EFF, i.e., efficiency) on X (or Z, i.e., explanatory variables) for the FDH reference technology, since all cases show substantial probability mass concentrations for efficient municipalities. Yet when analyzing DEA-VRS (Figs. 4 and 5) the observed patterns are

[22]The best performers—in terms of balance between bias and variance—are those using plug-in methods, such as the ones put forward by Wand and Jones (1994).

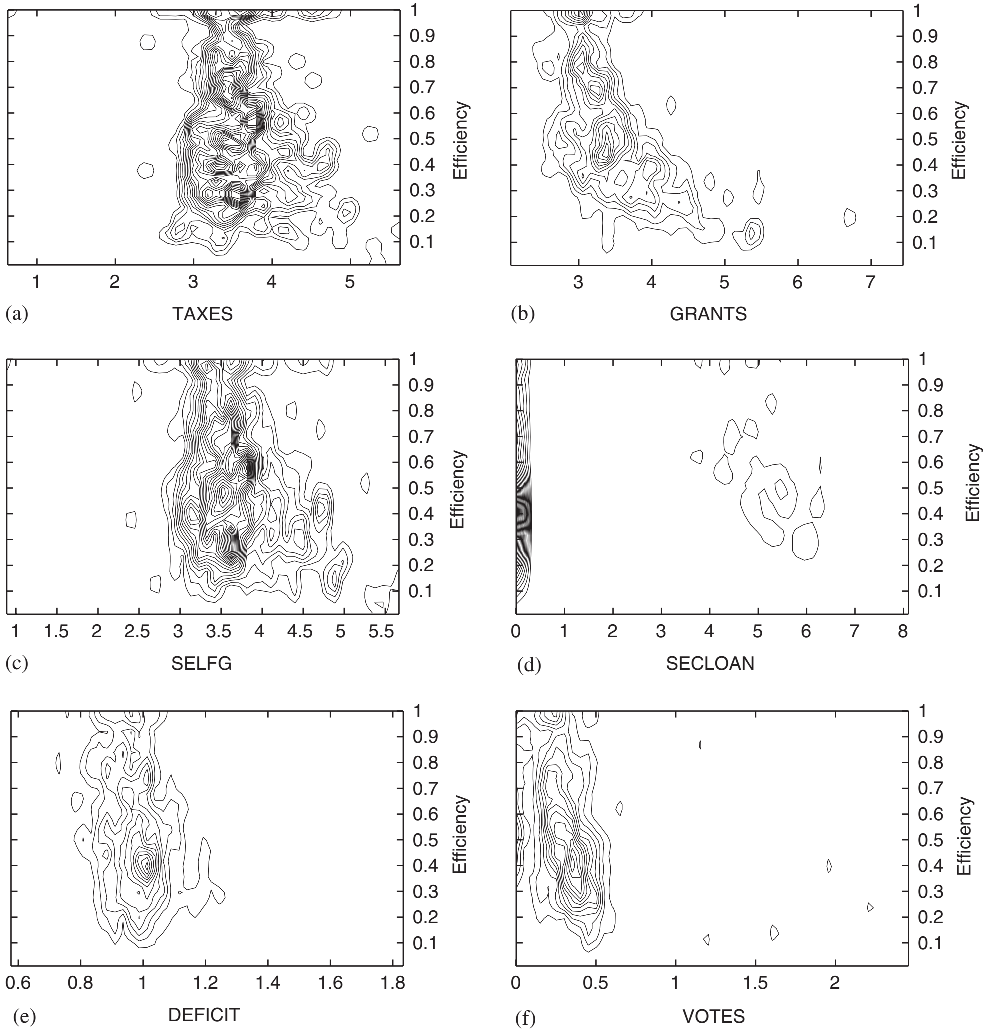

Fig. 5. Efficiency determinants, joint densities, overall cost efficiency (contour plots).

not exactly coincidental. In these cases the number of efficient municipalities is ostensibly lower, preventing probability mass concentrations at the top—either left or right—of each sub-figure. This feature makes the association between variables clearer in some instances. This is the case of grants and cost efficiency (Figs. 4b and 5b), for which the negative relationship is ostensibly apparent. For other variables the relationship is less clear, yet the choice of both OX and OY ranges might have quite an effect. Specifically, although both taxes and self-generated revenues exhibit a negative association for some municipalities, this does not hold for all of them (Figs. 4a, c, 5a and c).

Of special note is the case of *VOTES*, whose fuzzy nonparametric regression line (Fig. 3f) does not allow us to infer any clear association between the votes of the governing party divided by population and efficiency. Yet the nonparametric bivariate density estimation reveals that the number of observations with $VOTES > 1$ is very low. If we disregard these municipalities, and consider only those with $VOTES < 1$, which are those with the overwhelming majority, the apparent relationship is negative. Thus, it seems that our "preferred" hypothesis is accepted, as governing parties with a lower percentage of votes might be facing strong monitoring from other parties and coalitions, which contributes to enhancing efficiency.

5. Concluding remarks

This article considers an alternative methodology to find out which factors affect cost efficiency in Spanish municipalities using nonparametric methods both in the first stage, when the cost efficiency evaluation is determined, and in the second stage, when the factors causing inefficiency are sought. From a methodological point of view, our nonparametric application guarantees that the same hypotheses hold in both stages and, for this reason, results are obtained using consistent methods.

The organizations analyzed (Valencian local governments) are particularly complex for several reasons, namely: (i) local governments are public sector organizations, and nobody can ex ante guarantee whether they fall in line with the interests of their constituency; (ii) objectives are multiple and different targets can be pursued; (iii) the strategy on contracting out, or outsourcing services can range from totally centralized to essentially decentralized local councils; (iv) the level of services to be provided grows with the size, and the level of quality is very variable, and (v) the specific characteristics of the technology of production are not entirely clear. Indeed, the existing complexity demands a method to evaluate efficiency without assuming further strong behavioral hypotheses, or defining any ungranted technological reference. On this point, it is worth noting that Spanish laws on municipality management provide us with a useful guide to design the appropriate evaluation: The services provided by local councils must be produced without generating financial deficit. Thus, regardless of the services provided, their quality, and the size of the municipality, the generally accepted principle is that the diversity of services must be offered by minimizing total expenditures (in this paper we define expenditure as notion of cost, despite the fact that in some specific contexts these variables may differ).

Therefore, the requirement on expenditure control enables us to orient efficiency evaluation towards the total cost frontier minimization. This objective is reinforced when considering that outputs are basically exogenous and very stable, especially in the short run. Accordingly, we have the behavioral criteria for the evaluation, but the above-mentioned complexity requires the use of very flexible methods to compute inefficiency. The required flexibility is provided by nonparametric methods. In order to determine the cost efficiency corresponding to each municipality, we define two different technological references: Convex and nonconvex nonparametric frontiers. Convex technologies assume that combinations amongst efficient observations can be made, whereas nonconvex only accepts cost comparisons between two individual observations; thus, the only technological assumption made in the second case is that two different observations can always be compared.

Maintaining this reasonable objective (cost minimization), together with the convex and the nonconvex reference technologies, a second stage analysis was carried out to determine the factors affecting the cost inefficiency found. Taking into account the extant literature on local government efficiency in particular, and efficiency studies in general, we found that most two-stage analyses quite often become dangerously inconsistent since efficiency scores are determined using nonparametric methods, whereas the factors affecting inefficiency are detected by using parametric methods. This is, precisely, the main objective of the empirical part of the study: To show how it is possible to carry out the second stage without having to abandon the nonparametric field.

The application to Valencian municipalities yields a general picture of persistent growth in overall cost efficiency with the size of the municipalities. This idiosyncratic result suggests that small towns are far from their respective frontier, whereas large cities are closer—and less dispersed—from their respective cost frontier. Our mathematical programming techniques also allowed subsequent analysis of the output slacks detected, i.e., whether output underprovisions (or shortfalls) existed, and what their magnitude was. Quite a few disparities are found, suggesting that both the number of lighting points and street infrastructure surface area, essentially related to capital expenditure, suffer large underprovisions. Indeed, street infrastructure surface area is more important than what one might a priori expect, since it also proxies for access to population centers, street cleaning, or supply of drinking water to households. It is also important to bear in mind that our output indicators include a quality variable, of special relevance when measuring the efficiency of municipalities. Results also vary according to local government size, as large municipalities generally perform better. Yet this result is not entirely attributable to a more proficient management.

Although our findings indicate that there is a wide margin available to public managers for optimization in the use of public resources, some of these inefficiencies might be caused by variables disregarded when measuring efficiency; some partly lie within local governments' control while others remain outside it. Specifically, we considered a set of both fiscal policy and political variables out of which self-generated revenues, grants, deficit, and governing party votes over total population were found to have a negative impact on efficiency, although they varied with the type of efficiency considered. More precisely, the empirical evidence was quite relevant in the case of overall cost efficiency for unconditional grants received from higher layers of government, and votes for the governing party over total population. The latter results were obtained via nonparametric smoothing techniques, as opposed to either OLS or Tobit models. This is important due to the special distributional features of DEA and FDH efficiency scores which, by construction, are bounded by zero and unity. As we use mathematical programming techniques, a mass of efficiency scores reaches the upper bound. Although some authors have attempted to address this issue, they did not consider the statistical dependency of efficiency scores, which leads to the violation of important hypotheses. The nonparametric flavor of the techniques considered here lessens such problems. Although some bootstrap techniques designed to address the issue have been introduced in recent times, we consider that the set of techniques presented here fits the local government setting properly, since they handle particularly well with the large data sets of both observations and variables that usually confront the analyst in this type of study.

Policy implications of the results presented here would aim to split short-run from long-run inefficiencies. Short-run inefficiencies can be controlled for by modifying the variables under control. In contrast, long-run inefficiencies are harder to remove, since they require a strategic perspective, and time to modify structural and exogenous circumstances related to the budget composition.

To conclude, this paper presents a diagnosis that focuses on the fundamental causes of inefficiency in the Spanish public sector on a local scale. However, the need to continue and improve the study, to incorporate certain factors that are uncontrollable at this stage, is undeniable. The authors consider the most important extensions to be related to: (i) considering specific prices so as to complement the decomposition by introducing technical and allocative components; (ii) the dynamic analysis of cost evolution; and (iii) the inclusion in the analysis of municipalities with different financial regulations and territorial structures.

Acknowledgements

We are grateful for the comments by George Battese, Tim Coelli, Kevin Fox, Quentin Grafton, Joseph Hirschberg and other participants at the Economic Measurement Group Workshop '02 held at the School of Economics, University of New South Wales, Sydney. We are also grateful for the comments by Rolf Färe and two anonymous referees, which contributed greatly to a general overhaul of the article. Maria Teresa Balaguer-Coll and Diego Prior acknowledge the financial support of the Instituto Valenciano de Investigaciones Económicas (IVIE), Fundació Caixa Castelló (02I199.01/1) and the Ministerio de Educación y Ciencia (SEC2003-04770). Emili Tortosa-Ausina acknowledges the financial support of Fundació Caixa Castelló (P1-1B2005-03) and the Ministerio de Educación y Ciencia (SEJ2005-01163/ECON).

References

Ali, A.I., Seiford, L.M., 1993. The mathematical programming approach to efficiency analysis. In: Fried, H., Lovell, C.A.K., Schmidt, S. (Eds.), The Measurement of Productive Efficiency: Techniques and Applications. Oxford University Press, Oxford, pp. 120–159.

Andersen, P., Petersen, N.C., 1993. A procedure for ranking efficient units in data envelopment analysis. Management Science 39 (10), 1261–1264.

Banker, R.D., Morey, R.C., 1986. The use of categorical variables in data envelopment analysis. Management Science 32, 1613–1627.

Bosch, N., Pedraja, F., Suárez-Pandiello, J., 2000. Measuring the efficiency of Spanish municipal refuse collection services. Local Government Studies 26 (3), 71–90.

Bradford, D., Malt, R., Oates, W., 1969. The rising cost of local public services: Some evidence and reflections. National Tax Journal 22, 185–202.

Briec, W., Kerstens, K., Vanden Eeckaut, 2004. Non-convex technologies and cost functions: Definitions, duality and nonparametric tests of convexity. Journal of Economics 81 (2), 155–192.

Brueckner, J.K., 1979. Property values, local public expenditure, and economic efficiency. Journal of Public Economics 11, 223–245.

Charnes, A., Cooper, W.W., Sueyoshi, T., 1988. A goal programming/constrained review of bell system breakup. Management Science 34, 1–26.

Davis, M.L., Hayes, K., 1993. The demand for good government. Review of Economics and Statistics 75, 148–152.

De Borger, B., Kerstens, K., 1996. Cost efficiency of Belgian local governments: A comparative analysis of FDH, DEA, and econometric approaches. Regional Science and Urban Economics 26, 145–170.

De Borger, B., Kerstens, K., Moesen, W., Vanneste, J., 1994. Explaining differences in productive efficiency: An application to Belgian municipalities. Public Choice 80, 339–358.

Deaton, A., 1989. Rice prices and income distribution in Thailand: A nonparametric analysis. The Economic Journal 99, 1–37.

Deller, S.C., 1992. Production efficiency in local government: A parametric approach. Public Finance/Finances Publiques 47, 32–44.

Deprins, D., Simar, L., Tulkens, H., 1984. Measuring labor-efficiency in post offices. In: Marchand, M., Pestieau, P., Tulkens, H. (Eds.), The Performance of Public Enterprises: Concepts and Measurement. North-Holland, Amsterdam, pp. 243–267 (Chapter 10).

DiNardo, J., Fortin, N.M., Lemieux, T., 1996. Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. Econometrica 64 (5), 1001–1044.

Fan, J., Gijbels, I., 1996. Local Polynomial Modelling and Its Applications. Monographs on Statistics and Applied Probability, vol. 66. Chapman & Hall, London.

Färe, R., Grosskopf, S., 2000. Theory and application of directional distance functions. Journal of Productivity Analysis 13, 93–103.

Färe, R., Grosskopf, S., Lovell, C.A.K., 1994. Production Frontiers. Cambridge University Press, Cambridge.

Fletcher, J., Snee, H., 1985. The need for output measurements in the service industries: A comment. The Service Industries Journal 5 (1), 73–78.

Fox, K.J., 2001. Efficiency in the Public Sector. Studies in Productivity and Efficiency, vol. 1. Kluwer Academic Publishers, Boston.

Gil Jiménez, F., 2001. Une analyse des desequilibres financiers des collectivites locales en Espagne. Revue d'Economie Regionale et Urbaine 0 (3), 369–392.

Giménez, V.M., Prior, D., 2003. Evaluación frontera de la eficiencia en costes. Aplicación a los ayuntamientos de Cataluña. Papeles de Economía Española 95, 113–124.

Grosskopf, S., Hayes, K.J., 1993. Local public sector bureaucrats and their input choices. Journal of Urban Economics 33, 151–166.

Grosskopf, S., Yaisawarng, S., 1990. Economies of scope in the provision of local public services. National Tax Journal 43 (1), 61–74.

Grossman, P.J., Mavros, P., Wassmer, R.W., 1999. Public sector technical inefficiency in large U.S. cities. Journal of Urban Economics 46, 278–299.

Haavelmo, T., 1944. The probability approach in econometrics. Econometrica, Monograph, vol. 12.

Härdle, W., 1990. Applied Nonparametric Regression. Cambridge University Press, Cambridge.

Härdle, W., Linton, O., 1994. Applied nonparametric methods. In: Engle, R., McFadden, D. (Eds.), Handbook of Econometrics, vol. 4. North-Holland, Amsterdam.

Härdle, W., Müller, M., Sperlich, S., Werwatz, A., 2004. Nonparametric and Semiparametric Models. Springer Series in Statistics. Springer, New York.

Harker, P.T., Xue, M., 1999. Overcoming the inherent dependency of DEA efficiency scores: A bootstrap approach. Working paper, Wharton Financial Institutions Center.

Hart, J.D., 1997. Nonparametric Smoothing and Lack-of-Fit Tests. Springer Series in Statistics. Springer, New York.

Hayes, K.J., Chang, S., 1990. The relative efficiency of city manager and mayor-council forms of government. Southern Economic Journal 57, 167–177.

Hirschberg, J.G., Lloyd, P.J., 2002. Does the technology of foreign-invested enterprises spill over to other enterprises in China? An application of post-DEA bootstrap regression analysis. In: Lloyd, P.J., Zang, X.G. (Eds.), Modelling the Chinese Economy. Edward Elgar Press, London.

Ladd, H.F., 1994. Fiscal impacts of population growth: A conceptual and empirical analysis. Regional Science and Urban Economics 24, 661–686.

Levitt, M.S., Joyce, M.A.S., 1987. The Growth and Efficiency of Public Spending. Cambridge University Press, Cambridge.

Lindsay, C., 1976. A theory of government enterprise. Journal of Political Economy 84, 1061–1077.

Lovell, C.A.K., Walters, L.C., Wood, L.L., 1994. Stratified models of education production using modified DEA and regression analysis. In: Charnes, A., Cooper, W., Lewin, A.Y. (Eds.), Data Envelopment Analysis. Theory, Methodology and Applications. Kluwer Academic Publishers, Boston, pp. 329–352 (Chapter 17).

Marlow, M.L., 1988. Fiscal decentralization and government size. Public Choice 56, 259–270.

Marron, J.S., Schmitz, H.P., 1992. Simultaneous density estimation of several income distributions. Econometric Theory 8, 476–488.

Niskanen, W., 1975. Bureaucrats and politicians. Journal of Law and Economics 18, 617–643.

Pagan, A., Ullah, A., 1999. Nonparametric Econometrics. Themes in Modern Econometrics. Cambridge University Press, Cambridge.

Park, B.U., Simar, L., Weiner, C., 2000. The FDH estimator for productivity efficiency scores. Econometric Theory 16 (6), 855–877.

Prieto, Á.M., Zofío, J.L., 2001. Evaluating effectiveness in public provision of infrastructure and equipment: The case of Spanish municipalities. Journal of Productivity Analysis 15 (1), 41–58.

Ray, S.C., 1991. Resource-use efficiency in public schools: A study of Connecticut data. Management Science 37 (12), 1620–1629.

Ruppert, D., Sheather, S.J., Wand, M.P., 1995. An effective bandwidth selector for local least squares regression. Journal of the American Statistical Association 90 (432), 1257–1270.

Ruppert, D., Wand, M.P., Carroll, R.J., 2003. Semiparametric Regression. Cambridge University Press, Cambridge.

Silkman, R., Young, D.R., 1982. *X*-efficiency and state formula grants. National Tax Journal 35, 383–397.

Simar, L., Wilson, P.W., 2000. Statistical inference in nonparametric frontier models: The state of the art. Journal of Productivity Analysis 13 (1), 49–78.

Simar, L., Wilson, P.W., 2006. Estimation and inference in two-stage, semi-parametric models of productive processes. Journal of Econometrics, forthcoming.

Simonoff, J.S., 1996. Smoothing Methods in Statistics. Springer, New York.

Solé-Ollé, A., 1997. Tax exporting and redistributive politics: An empirical investigation on the determinants of Spanish local tax structure. Public Finance 52 (1), 102–124.

Tulkens, H., 1993. On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts, and urban transit. Journal of Productivity Analysis 4, 183–210.

Vanden Eeckaut, P., Tulkens, H., Jamar, M.-A., 1993. Cost efficiency in Belgian municipalities. In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency. Oxford University Press, Oxford, pp. 300–334 (Chapter 12).

Vela, J.M., 1996. Latest developments in local government accounting in Spain. Financial Accountability & Management 12 (3), 205–224.

Wand, M.P., Jones, M.C., 1994. Multivariate plug-in bandwidth selection. Computational Statistics 9, 97–116.

Wand, M.P., Jones, M.C., 1995. Kernel Smoothing. Chapman & Hall, London.

Worthington, A.C., Dollery, B.E., 2000. An empirical survey of frontier efficiency measurement techniques in local government. Local Government Studies 26, 23–52.