Output complexity, environmental conditions, and the efficiency of municipalities

Maria Teresa Balaguer-Coll · Diego Prior · Emili Tortosa-Ausina

Abstract Over the last few years, many studies have analyzed the efficiency of local governments in different countries. An accurate definition of their output bundles—i.e., the services and facilities they provide to their constituencies—is essential to this research. However, several difficulties emerge in this task. First, since in most cases the law only establishes the minimum amount of services and facilities to provide, it may well be the case that some municipalities go beyond the legal minimum and, consequently, might have an uncertain effect on efficiency when compared to other municipalities which stick to the legal minimum. Second, municipalities face very different environmental conditions, which raises some doubts about the plausibility of an unconditional analysis. This study tackles these problems by proposing an analysis in which the efficiency of municipalities is evaluated after splitting them into clusters according to various criteria (output mix, environmental conditions, level of powers). We perform our estimations using order-\(m\) frontiers, given their robustness to outliers and immunity to the curse of dimensionality. We provide an application to Spanish municipalities, and results show that both output mix and, more especially, environmental conditions, should be controlled for, since efficiency differences between municipalities in different groups are notable.

Keywords Efficiency · Environmental conditions · Local government · Metafrontier · Order-\(m\)

JEL Classification D24 · D60 · H71 · H72

1 Introduction

Over the last few years, a wide range of studies has analyzed the efficiency of municipalities from multiple perspectives. The empirical evidence now available is increasing, and can be divided into two groups. On the one hand, some studies analyze efficiency in the provision of a specific service such as refuse collection (Brueckner 1981; Ruggiero 1995; Bosch et al. 2000). In some countries, however, this type of study presents certain disadvantages related to the difficulties in assigning the amount of input usage by each specific service. On the other hand, many studies have considered a global perspective, taking into account that local governments provide their constituencies with a wide variety of services and facilities from the municipal budget. The literature is fairly extensive, yet scattered in time—there has not been a continuous flow of research, rather studies have appeared sporadically over time. For more comprehensive reviews, see Tang (1997), or De Borger and Kerstens (2000).

See, for instance, Brueckner and Wingler (1984), Deller (1992), Taylor (1995), De Borger and Kerstens (1996), Grossman et al. (1999), Hughes and Edwards (2000) and, more recently, Sampaio de Sousa and Stošić (2005), Balaguer-Coll et al. (2007), or Barankay and Lockwood (2007).
In today’s complex scenario of economic and financial downturn and increasing public deficits in several euro-area countries, efficient management of resources at all levels of government (central, regional, and municipal) is essential. As indicated by De Witte and Geys (2011), a reasonable way to deal with increasing tasks and tightening budget requirements is to improve productive or technical efficiency, understood in terms of providing a maximum amount of output for a given level of inputs (Koopmans 1951; Fried et al. 2008). Yet in the particular case of municipalities we are dealing with, a relevant problem (shared by the second category of studies referred to in the first paragraph) is the difficulty to accurately define, and measure, what it is that local governments produce. In most cases the problems arise due to the impossibility of directly quantifying the supply of public services. The Spanish case, on which we focus, is not free from that criticism, although its magnitude is lower, largely due to the availability of data on most of the public services that municipalities are legally bound to provide.

However, the law only establishes the minimum services and facilities each municipality must provide, which depends upon population. Nothing prevents a particular municipality from going beyond this legal minimum and providing not only more of each compulsory service (such as discretionarily increasing the area of public parks), but also providing additional services and facilities whose input usage may be substantial. Although the implications for efficiency are not clear, it does not seem a priori fair to compare municipalities which stick to the legal minimum with those that go beyond it. These and other difficulties when modeling municipalities’ provision of services and facilities have led some authors to talk about output complexity (Haynes 2003).

The difficulties when measuring local governments’ service provision exacerbate not only when some municipalities go beyond the legal minimum, but also when some of them choose to provide services and facilities for which information is unavailable. Some studies that estimate the amount of these extra services and facilities for which there is no information acknowledge this complex reality, but results are generally based on surveys carried out on a limited number of municipalities (see Vilalta and Mas 2006). Previous studies dealing with similar issues in other countries found comparables results, i.e. that the additional costs incurred by many municipalities are quite large, although they also vary a great deal across observations (see, for instance Bennett and DiLorenzo 1982; Marlow and Joufafari 1989; Merrifield 1994). These studies constitute further evidence to support the hypothesis that some municipalities provide their constituencies with a larger amount of services and facilities, not only because different constituencies have different needs but also because of the varying environmental conditions each municipality faces.

There is a large body of literature on how different environmental conditions facing DMUs (Decision Making Units) affect their efficiency in different contexts—not only in local government. Since the pioneering contributions by Banker and Morey (1986a, b), many studies have analyzed the issue. Following De Witte and Kortelainen (2008), one may consider four main families of studies to incorporate the exogenous environment in nonparametric efficiency analysis. The cited studies by Banker and Morey (1986a, b) can be classified in the first category of studies controlling for the environment, which considers a one-stage approach; this approach has some drawbacks shared by some of the later variants (see, for instance Färe et al. 1989; Ferrier and Lovell 1990). The second category of studies considers a two-stage approach, examples of which would be Ray (1991) and Simon and Wilson (2007). There is a third category of studies based on the concept of metafrontier—or frontier separation. Although the concept has been popularized by Battese and Rao (2002) and Battese et al. (2004), among others, some of their ideas had been developed years before. According to metafrontier approaches, the efficiencies of DMUs that operate under a given production technology are not comparable with those of other DMUs operating under different technologies. In our setting, this would imply that it is not possible to compare the efficiencies of municipalities in different groups formed according to different criteria. Battese and Rao (2002) propose a solution based on the metafrontier using stochastic frontier analysis (SFA), which was later refined by Battese et al. (2004). In their paper, Battese and Rao (2002) assume there are two different data-generation mechanisms—one with respect to the stochastic frontier, the other with respect to the metafrontier model. In contrast, Battese et al. (2004) assume that the metafrontier function is an overarching function that encompasses the deterministic components of the stochastic frontier production functions for those DMUs operating under different technologies. Despite the advantages of the metafrontier, some problems that will be described below remain. Finally, a fourth approach is based on the conditional efficiency measures by Daraio and Simar (2005, 2007b), which provide some methods that developed ideas previously outlined by Cazals et al. (2002). Although the various advantages of these methods have contributed to their increase in popularity (see, for instance Bonaccorsi and Daraio 2008; De Witte and Dijkgraaf 2009; Daraio and Simar 2005, among other applications in different contexts), some questions such as the issue of bandwidth selection remain unsolved.

---

2 This is thoroughly described in De Witte and Kortelainen (2008).
3 The metafrontier function was first introduced by Hayami (1969) and Hayami and Ruttan (1970, 1971). As indicated by these authors (Hayami and Ruttan 1971, p. 82), “the metaproduction function can be regarded as the envelope of commonly conceived neoclassical production functions.”
We propose a contribution to this literature in the particular sub-field of the metafrontier approaches, in combination with cluster analysis. Specifically, we propose a method which combines both the second and third approaches referred to here, i.e. the metafrontier method and the two-stage method, since groups of observations are formed in a second stage. As indicated by De Witte and Kortelainen (2008), the metafrontier models have certain disadvantages, namely, they can only be applied to categorical environmental variables, in practice it is not possible to include several environmental factors, and comparison and statistical testing of efficiency differences between more than two classes can be at least challenging (De Witte and Kortelainen 2008). As we shall see below, our approach is more robust to these criticisms, since we consider a three-stage process in which efficiencies are estimated in the first instance, groups are built in the second stage (using cluster analysis), and finally we estimate conditional efficiency measures taking into account the information on cluster membership. This way to proceed was encouraged by (O’Donnell et al. 2008), who indicated the advantages of using multivariate statistical techniques when natural boundaries between groups are unavailable. However, as far as we know, few authors—if any—have combined the metafrontier with cluster methods.

In the particular case of metafrontier methods, most of the ensuing literature has been of an applied nature (see, for instance Bos and Kool 2006), although several theoretical refinements to the initial methodologies have also been proposed (see, for instance Ruggiero 2004). In the case of municipalities, the implications of metafrontier methods—as well as other methods for controlling for environmental conditions—would be related, among other issues, to the fact that different local governments face different constituencies in terms of economic and social conditions, to whose needs municipalities may be more or less responsive irrespective of the amount of services they are obligated to provide. Municipalities also have different characteristics in terms of geography (including, for instance, rugged terrain, or urban sprawl); different sectoral production characteristics (in terms of tourism, etc.); and other characteristics.

We consider municipalities may provide more services and facilities than the legally established minimum because of the different environmental conditions they face. Tourist municipalities may face budget strains due to their greatly increased personnel needs during the high season. Other municipalities may face higher costs because of urban sprawl. According to Horta-

Rico and Solé-Ollé (2010), the urban spatial structure of many Spanish cities not only has an environmental impact, but also a major impact on municipal finances. In other cases, reasons might be more involved, such as wealthy constituencies asking for additional services, or rising expenditures on security because of the rapid population increases experienced by some Spanish cities. As documented by Vilalta and Mas (2006), in a study applied to a sample from the province of Barcelona, more than 30% of municipal expenditures were discretionary. In these circumstances, one may reasonably expect that some municipalities will be mislabeled as inefficient (or, at least, more inefficient than other municipalities) simply because of this wide variety of scenarios. Therefore, it would be more appropriate to compare only local governments facing similar environmental conditions.

This study approaches these problems through a three-stage analysis that first analyzes the productive efficiency of municipalities, second it divides them into groups according to different classifications; and thirdly conditional efficiencies are measured. In the second stage we consider three criteria to classify municipalities into different groups. The first one considers clusters according to local governments’ output mixes, in order to compare the efficiency of municipalities with similar output bundles; in this way we can control for the fact that some of them might provide services and facilities beyond the legal minimum—hence, they will be more complex. The second criterion constructs groups of municipalities for which we include information on environmental conditions. The third criterion forms groups according to the level of powers, which is determined by the population living in each municipality (the levels of services each municipality must provide hinges on the number of inhabitants) since, as documented by Balaguer-Coll et al. (2010a, b), some links may exist between municipalities’ efficiency and decentralization issues. In our application to Spanish municipalities, results show that both output mix and the different environments that municipalities face are issues to control for, since the differences between municipalities affiliated to different groups turned out to be statistically significant.

Therefore, in the second stage municipalities are classified into groups according to different criteria given that, according to the hypotheses formulated, local governments in each group might be facing different production opportunities. They respond by making choices from different sets of feasible input-output combinations. These so-called technology sets differ because of variations in available stocks of physical, human and financial capital, economic infrastructure, resource endowments and any other characteristics of the physical, social and economic environment in which production, or service provision, takes place. As indicated by O’Donnell et al. (2008), such differences have led efficiency researchers to estimate separate frontiers for different groups of DMUs.4

In public sector applications such as the measurement of local government efficiency, most studies have used

4 For example, separate frontiers have been estimated for universities in Canada (McMillan and Chan 2006), Australia (Worthington and Lee 2008) and the United Kingdom (Glass et al. 1995), and for bank branches in South Africa (O’Donnell and Westhuizen 2002) and Spain (Lovell and Pastor 1997).
nonparametric techniques such as DEA (Data Envelopment Analysis), or its nonconvex version (Free Disposable Hull, FDH), for a variety of reasons (Fox 2001). O’Donnell et al. (2008) have recently filled this gap in the literature, by extending the metafrontier to DEA and alternative SFA approaches for estimating both metafrontiers and group frontiers. However, their solutions are not entirely satisfactory, mainly because of the curse of dimensionality that generally affects efficiency scores obtained using DEA. As indicated by Daraio and Simar (2007a), increasing the number of inputs or outputs, or decreasing the number of units being compared, leads to higher efficiencies, simply as a result of a statistical artifact. Multiple applications in disparate fields are affected by this issue (see, among many others Maudos et al. 2002), which is especially severe in fields such as mutual fund evaluation, where difficulties arise in defining the number of inputs and outputs (Joro and Na 2002). Our approach to deal with this problem is based in the order-

order-

order-

2.2 Measuring efficiency using robust techniques: order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-

order-
these problems some additional procedures are required in order to make FDH estimates more robust. Several approaches have already been proposed in the literature. Wilson (1993, 1995) introduced descriptive methods to detect influential observations in nonparametric efficiency estimations. More recent developments include the order-$m$ frontiers (Cazals et al. 2002; Simar 2003). The order-$m$ approach, based on the concept of expected maximal (minimal) output (input or cost) function, yields frontiers of varying degrees of robustness. Daraio and Simar (2007a, p.82), describe two ways to compute the order-$m$ efficiency coefficients. In what follows, we briefly describe the second one, based on a Monte Carlo algorithm.

Consider a positive fixed integer $m$. For a given level of input $(x_0)$ and output $(y_0)$, the estimation defines the expected value of maximum of $m$ random variables $(Y_1, \ldots, Y_m)$, drawn from the conditional distribution of the output matrix $Y$ observing the condition $Y_m \geq y_0$. Formally, the proposed algorithm (algorithm I) to compute the order-$m$ estimator has the following steps:

1. For a given level of $y_0$, draw a random sample of size $m$ with replacement among those $y_{sm}$, such that $y_{sm} \geq y_0$.
2. Compute Program (1) and estimate $\bar{z}_s$.
3. Repeat steps 1 and 2 $B$ times and obtain $B$ efficiency coefficients $\bar{z}_s^b (b = 1, 2, \ldots, B)$. The quality of the approximation can be tuned by increasing $B$, but in most applications $B = 200$ seems to be a reasonable choice.
4. Compute the empirical mean of $B$ samples as:

$$z^m_s = \frac{1}{B} \sum_{b=1}^{B} \bar{z}_s^b. \tag{2}$$

As $m$ increases, the number of observations considered in the estimation approaches the observed units that meet the condition $y_{sm} \geq y_0$ and the expected order-$m$ estimator in each one of the $b$ iterations ($\bar{z}_s^b$) tends toward the FDH ($z^m_{FDH}$). Thus, $m$ is an arbitrary positive integer value, but it is always convenient to observe the fluctuations of the $z^m_s$ coefficients depending on the level of $m$. For acceptable values of $m$, $z^m_s$ will normally present values smaller than unity (this indicates that these units are inefficient, as total costs can be reduced without modifying the production plan). When $z^m_s > 1$, the $s$ unit can be labeled as superefficient, as the order-$m$ frontier exhibits a higher total cost.

2.3 Adapting the order-$m$ estimators to different technologies

As mentioned above, the order-$m$ estimation is an excellent tool to mitigate the problems of dimensionality and the presence of extreme observations and outliers. However, this evaluation will be of little use if part of the inefficiency found hinges on the output complexity or the different environmental conditions that local governments face, which could lead to biased estimates of the frontier, and hence misleading policy implications.

Therefore, our objective is to define a process that can estimate the impact on efficiency of output complexity, environmental conditions and, in general, other options that consider classifying municipalities in different groups. This estimation is possible following the proposals of Battese et al. (2004) and O’Donnell et al. (2008) for estimating a metafrontier production function. This process (algorithm II) contains the following steps:

1. Use cluster analysis to classify the $S$ units in $S_1, S_2, \ldots, S_C$ groups.
2. Following the algorithm to estimate the order-$m$ efficiency coefficients, complete steps 1 to 4 of algorithm I to estimate the efficiency coefficients $(z^m_{s_1}, z^m_{s_2}, \ldots, z^m_{s_S})$ for the municipalities classified in each one of the clusters $S_1, S_2, \ldots, S_C$. In order to facilitate the cross-comparison of the results, irrespective of the number of units classified in each cluster, the same value for $m$ will be assigned in all the estimations. By doing this, the problems of dimensionality and the potential impact of the outliers will be neutralized.
3. After completing the conditional frontiers in step 2 of algorithm II, apply steps 1 to 4 of the order-$m$ estimation to the complete sample and estimate the efficiency coefficients with respect to the metafrontier $z^m_s$.
4. Estimate the technology gap ratio (TGR) separating the conditional and the metafrontier as the ratio $(z^m_s / z^m_{s_1}), (z^m_s / z^m_{s_2}), \ldots, (z^m_s / z^m_{s_S})$.

Figure 1 presents an illustration of a simple case with one output and the total cost. At a given output level, the technology gap ratio (TGR) is defined as the lowest possible cost within the metafrontier divided by the lowest total cost at the conditional specific cluster.\footnote{We have borrowed the concept of technology gap ratio from Battese and Rao (2002), Battese et al. (2004) and O’Donnell et al. (2008). They use the term technology because they consider DMUs operating under different technologies. Although we also talk about different technologies, in order to facilitate comparisons with their papers, we must acknowledge we do not explicitly test whether the municipalities in the different groups have different technologies. Therefore, when referring to technology we are only suggesting whether the effect of each hypothesis (output mix or environmental conditions) matters or not.}
We perform the analysis for a sample of Spanish municipalities with populations between 1,000 and 50,000 inhabitants for year 2000. Both input and output data are provided by the Spanish Ministry for Public Administration. Information on outputs is gathered through the survey on local infrastructures and facilities (Encuesta de Infraestructuras y Equipamientos Locales), which is performed with 5-year frequency and, consequently, constrains our sample period. Information on inputs basically consists of different types of costs, and is taken from local government budgetary data. This data is available for every year. The regions that meet our criteria (data for year 2000, and data for both inputs and outputs) were Andalusia, Aragon, Asturias, the Canary Islands, Cantabria, Castile-La Mancha, Extremadura, Murcia, La Rioja, and the Valencian Community. The final sample was made up of 1,198 municipalities. There was no information for the remaining regions for several reasons. At the time of the study, Madrid had not yet presented information on its outputs. Catalonia, the Basque Country and Navarra do not have to provide the Spanish Ministry for Public Administration with this information.

Measuring the production process at municipal level is usually more difficult than in other sectors/industries. We can distinguish three stages in this process of transforming inputs into outputs (Bradford et al. 1969). In the first stage primary inputs (labor, equipment and external services) are transformed into intermediate outputs (e.g., hours of traffic control or the extension of police services). In the second stage, intermediate outputs are transformed into direct outputs. This is what Bradford et al. (1969) call D-outputs, which are ready for consumption by the population. In the third stage, the direct outputs ultimately have welfare effects on consumers (e.g., increasing perceptions and feelings of safety and welfare). The third stage of the process can be directly captured by outcome indicators (labeled C-outputs by Bradford et al. 1969), which reflect the degree to which direct outputs translate into welfare improvements as perceived by consumers.

The efficiency of municipalities can be measured at each stage of this production process. However, under normal circumstances this will be difficult because data might be either unavailable or simply poor, making it difficult to distinguish between primary inputs, intermediate outputs, direct outputs, and final welfare effects. For this reason the analysis is usually confined to analyzing the first and second phases of this process, i.e., the links between primary inputs and direct outputs. We base our selection of outputs on the services and facilities provided by each municipality. 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 must select proxies for these services and facilities. As pointed out by 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. Other relevant outputs, such as provision of primary and secondary education, are not the responsibility of Spanish municipalities.

Spanish law requires municipalities to provide minimum services depending on their size. Some of the minimum services and facilities must be provided by all municipalities, while others are only binding for larger municipalities (with populations of over 5,000, 20,000, and 50,000, the boundaries that define the different categories). The second column in Table 1 reports information on the minimum services that each category of municipalities must provide. The third column indicates the selected output indicators to measure the different services and facilities. Our output choice was driven by the minimum services and facilities. The list of outputs for year 2000, along with summary statistics, are reported in Table 2. The choice was also driven by previous studies on local government efficiency in other European countries, since for the most part they are endowed with the same competencies. 6

Therefore, we will be considering eight outputs which measure the minimum amount of services and facilities that municipalities must provide to their constituencies. However, some of these outputs are not proper outputs but proxies for them. This is the case of population size \( Y_1 \), which is used as a proxy for several services—those indicated in Table 2. We admit this assumption is crude, but the amount of available information, although relatively

---

6 Differences are basically confined to the area of education which, in Spain, corresponds to regional and central governments.
high, has to measure a very complex reality. In this line, it would presumably be a better indicator of “needs” or “demand” rather than municipal “production”. This assumption has an wide acceptance in the literature, as shown by studies such as De Borger and Kerstens (1996) or Vanden Eeckaut et al. (1993), basically due to a generalized lack of information on municipal services. Other services and facilities can be measured more directly. This is the case of public street lighting (measured by the number of lighting points, \(Y_2\)), waste collection (measured by the tons of waste collected, \(Y_3\)), provision of market (measured by the market surface area, \(Y_6\)), or public parks (measured by the registered surface area of public parks, \(Y_7\)). Other outputs measure several services and facilities. For instance, street infrastructure surface area (\(Y_4\)) measures not only the paving of public roads and street cleaning but it is also a proxy for access to population centers, fire prevention and extinction, or urban passenger transport service, and public buildings surface area (\(Y_5\)) measures the provision of public libraries, the provision of social services, or the provision of public sports facilities. Finally, some services and facilities such as the provision of social services are measured by several outputs (total population, \(Y_1\), public building surface area, \(Y_5\), and the surface area of assistance centers, \(Y_8\)).

## Table 1
**Output indicators based on minimum services provided**

| Minimum services provided | Output indicators |
|---------------------------|-------------------|
| All local governments     |                   |
| Public street lighting    | Number of lighting points |
| Cemetery                  | Total population  |
| Waste collection          | Waste collected   |
| Street cleaning           | Street infrastructure surface area |
| Supply of drinking water  | Population, street infrastructure surface area |
| to households             |                   |
| Access to population      | Street infrastructure surface area |
| centres                   |                   |
| Paving of public roads    | Street infrastructure surface area |
| Regulation of food and    | Total population  |
| drink                     |                   |
| In local governments      |                   |
| with populations of over  |                   |
| 5,000, in addition        |                   |
| Public parks              | Surface of public parks |
| Public library            | Total population, public buildings surface area |
| Market                    | Market surface area |
| Treatment of collected    | Waste collected   |
| waste                     |                   |
| In local governments      |                   |
| with population of over    |                   |
| 20,000, in addition       |                   |
| Civil protection          | Total population  |
| Provision of social       | Total population, public buildings surface area, assistance centers surface area |
| services                  |                   |
| Fire prevention and       | Street infrastructure surface area |
| extinction                |                   |
| Public sports facilities  | Total population, public buildings surface area |
| Abattoir                  | Total population  |
| In local governments      |                   |
| with populations of over   |                   |
| 50,000, in addition       |                   |
| Urban passenger transport  | Total population, street infrastructure surface area |
| service                   |                   |
| Protection of the         | Total surface area |
| environment               |                   |

## Table 2
**Summary statistics for inputs and outputs, year 2000**

|                | Mean     | SD       |
|----------------|----------|----------|
| **Inputs**     |          |          |
| Wages and salaries (\(X_1\)) | 6,304.18 | 9,044.30 |
| Expenditure on goods and services (\(X_2\)) | 4,953.61 | 7,862.17 |
| Current transfers (\(X_3\)) | 1,072.51 | 1,748.62 |
| Capital expenditure (\(X_4\)) | 5,905.78 | 8,513.16 |
| Capital transfers (\(X_5\)) | 276.52  | 1,031.29 |
| **Total cost (\(TC\))** | 18,512.60 | 28,199.54 |
| **Outputs**    |          |          |
| Population (\(Y_1\)) | 6,118.54 | 6,991.48 |
| Number of lighting points (\(Y_2\)) | 1,100.07 | 1,022.89 |
| Tons of waste collected (\(Y_3\)) | 2,984.87 | 4,787.84 |
| Street infrastructure surface area (\(Y_4\)) | 256,097.13 | 261,653.68 |
| Public buildings surface area(b) (\(Y_5\)) | 599.84  | 1,198.27 |
| Market surface area(b) (\(Y_6\)) | 748.21  | 1,867.12 |
| Registered surface area of public parks(b) (\(Y_7\)) | 202.44  | 1,193.32 |
| Assistance centers surface area (\(Y_8\)) | 1,858.22 | 4,015.22 |
| **# of observations** | 1,198    |          |

|                | *a*      | *b*      |
|----------------|----------|----------|
| *a* In thousands of euros, converted from 1995 pesetas (1 euro = 166.386 pesetas) |
| *b* In square meters |

\[7\] The correlation coefficient between the total surface area and the different outputs measured using the surface they occupy is low. This is to be expected because many municipalities are quite large in terms of total surface, which does not necessarily mean that the space they devote to public buildings or parks will also be large.
These choices are in line with the previous literature, especially that focusing on Europe, such as De Borger and Kerstens (1996) or, more recently, Geys et al. (2010), although to a lesser extent than those focusing on the US case (see, e.g. Hughes and Edwards 2000). With respect to these previous contributions we have, in general, a more complete database which a priori would lead to more accurate results. However, some intricate issues still remain. For instance, in their recent paper, De Witte and Geys (2011) propose a methodology to analyze efficient public good provision in stage one only, as compared to the two stages we are focusing on. These authors consider that if the observed value of a given output such as, for instance, educational outcomes (although not in the case of Spain, where these powers are in the hands of regional governments) is demand-driven, then it would lie partially beyond the control of the public service provider. According to this line of reasoning, some particular services such as waste collection would fall in this category of outputs, since residents partially choose the amount of waste they produce—by recycling, etc. Although local governments still have a say here—they can, for instance, either facilitate or hamper recycling—we must admit it might not be as appropriate as some of the other variables chosen to measure municipal output.8

Therefore, our view of public good provision partly coincides with several previous contributions, but differs from other recent ones such as De Witte and Geys (2011),9 who adapt Hammond’s (2002) approach, which accounts for the fact that municipal outcomes might not be solely determined by the public service provider, and that a distinction between service potential and final outputs should be made (two stages). The benefits of this approach are that we avoid the bias induced by using final outputs that are influenced by citizens’ coproduction. Although undoubtedly interesting, this approach has some limitations such as: (1) the need to assume that public service providers provide services fitting to local preferences; and (2) a distinction between the first and second stage is not always clear, and public service providers could be thought to have some influence over the second-stage of the production process (De Witte and Geys 2011, p.322).

Our definition of inputs is based on budgetary variables reflecting the economic structure of Spanish local government expenditures (costs), details of which are determined by Spanish legislation,10 that considers three basic categories: current or ordinary expenditures, capital expenditures, and financial expenditures. Within these, current expenditures are further divided into four chapters, or categories, which account for: (1) personnel expenditure; (2) current goods and services expenditures; (3) financial expenditures; (4) current transfers. Capital expenditures are also broken down into either real investments, or capital transfers. The former is what Table 2 refers to as capital expenditures (X3), i.e., all expenditures local governments implement: (1) to produce or acquire capital goods; (2) to acquire necessary goods to provide local services in the right conditions; or (3) financial expenditures that are suitable for amortization. On the other hand, capital transfers (X5) refer to the payments to institutions to finance certain investments. Descriptive statistics for year 2000 are provided in Table 2. Since our analysis is entirely confined to overall cost efficiency, the fact that some local government departments may be actually sharing some costs does not raise any particular issue.11

Another issue concerning the data is the consideration of the aggregate total costs (in monetary units), or overall cost, as opposed to considering the amount of the inputs (in physical units) and their correspondent input prices, that can be non competitive. In such circumstances, following Camanho and Dyson (2008) what we are estimating is a coefficient of economic efficiency that is dependent on two basic sources of inefficiency: market efficiency—dependent on the input prices—and Farrell efficiency, dependent on the physical units consumed. Unfortunately, the lack of data on physical units prevents us from decomposing the economic efficiency, but the meaning of the economic inefficiency is maintained as the cost excess due to the two aforementioned basic factors.

We must also select those variables to include when carrying out the cluster analysis for classifying municipalities into groups according to the different hypotheses. This task is not easy given that we face certain relevant constraints. First, there is no well-established theory as to which variables constitute the “environmental conditions” that might impact on each municipality’s cost structure. Second, the available information is also limited. In a number of contexts the relevance of considering groups is unquestionable. As indicated by O’Donnell et al. (2008), in most practical settings DMUs can be grouped a priori on the basis of geographical, economic and/or political

---

8 We are grateful to one of the referees for this comment.
9 Other approaches have been reviewed by, for instance, Cordero-Ferrera et al. (2008).
10 See Ministerial Order (Orden Ministerial), September 20th, 1989.
11 There is a certain amount of subcontracting going on in some Spanish municipalities. However, this information is not publicly available. The only way to account for it would be to assume that an abnormally high value for X3 (expenditures on goods and services) and a low value of X1 (wages and salaries) would signal the existence of subcontracting, but we consider such an assumption to be too heroic. This and related issues have been considered recently by, for instance, Zafra Gómez et al. (2011).
boundaries, to name a few. However, it is not entirely clear how groups must be formed. In the absence of “natural” boundaries for the different groups, multivariate statistical techniques such as cluster analysis are available for determining both the number of groups and group membership. In our particular setting, the a priori classification or “natural” boundaries would be those based on size, since the limits of each group are determined by the extent of their powers. In contrast, we need the cluster analysis technique to classify municipalities according to either output mix or environmental conditions.

Regarding the classification based on output complexity, the variables selected to construct the groups are similar to those chosen as outputs, divided by population. This helps to control for the fact that, as pointed out in the introduction, some municipalities might go beyond the legal minimum and provide an amount of services which does not correspond to their size. Therefore, we only compare municipalities with similar output mixes, i.e., only those in the same group. We have selected as many variables to construct the clusters as outputs. The details are reported in Table 3.

Regarding the classification based on environmental conditions, we have chosen some of the variables provided by the Anuario Estadístico de La Caixa. Although they can be partly judged as ad hoc, and although some factors influence the amount, allocation and distribution of local public spending, we consider these provide a rough idea of the environmental conditions facing each municipality that might have an impact on municipal finances. Summary statistics are reported in Table 4, and the particular definition of each variable considered for performing the cluster analysis follows:

**Total surface area:** (divided by population). This indicator is similar to density, which is usually defined as urbanized land per person. Although some authors (Hortas-Rico and Solé-Ollé 2010) consider it is more appropriate to use urbanized land per person, this variable was not available for all municipalities in our sample. We consider this is an appropriate proxy for urban sprawl, although other indicators can be considered (for instance, street surface area divided by total surface). (ENV1)

**Tourist index:** this index may contribute to rising municipal expenditures, at least during some months of the year. Comparing these municipalities only with their peers might be more appropriate. This index is constructed by taking into account information on the economic activity tax (“Impuesto de Actividades Económicas”, IAE), based on the type of touristic accommodation (number of rooms and annual occupancy rates). The value of the index indicates each municipality’s share (of 100,000) of total national revenues due to this type of activity. (ENV2)

**Economic status:** municipalities with wealthier populations might be facing higher requirements. These constituencies may be willing to pay more taxes but, in return, they will demand more services and facilities. Some additional reasons are explained in the next section. This is an indicator of gross disposable income per capita, estimated for each municipality, which is divided in ten classes, from the poorest (level 1) to the richest (level 10). We define this indicator as the sum of total family income in the analyzed period, which is estimated using several covariates. (ENV3)

**Industrial activities:** (divided by population). Municipalities in which industrial activity is high may be affected by a different, probably higher, cost structure, since this type of activity requires higher investments in infrastructures, security, anti-pollution policies, etc., which may—although not necessarily—be offset by higher tax revenues. This variable is defined as the number of industrial activities subject to taxes. This number is very similar to the number of industrial establishments per municipality. (ENV4)

**Total number of cars:** (divided by population). This might be related to ENV3. Some wealthy constituencies also have higher levels of education (in relative terms) and might have a preference for non-polluting means of transport. (ENV5)

**Unemployment:** higher unemployment might entail higher crime and, therefore, increased demands for security. This variable is lower than what the official statistics indicate because it is defined as total unemployment divided by total population instead of working population. We cannot use the unemployment figures issued by the Spanish Bureau of Statistics (INE, Instituto Nacional de Estadística) because they are constructed at the provincial, regional and national levels, but not at municipal level. (ENV6)

**Total population growth, 1991–1998 (percentage):** municipalities facing higher population growth might have had to increase the services and facilities at their

---

12 As also indicated by O’Donnell et al. (2008), if the analysis were conducted in an SFA framework, it would also be possible to conduct statistical tests concerning the number of groups. El-Gamal and Inanoglu (2005) propose other methods to circumvent the use of multivariate analysis techniques.

13 An annual report provided by the largest Spanish savings bank, La Caixa.

14 More specific details on the definitions of the variables and the methodologies used to construct them are available from La Caixa Foundation upon request.

15 Please note that we only have information for municipalities with population between 1,000 and 50,000.
own expense, because of the speed at which the population’s demands have increased. If central and regional governments have not reacted promptly, they might face a sudden imbalance between their revenues and their costs. (ENV7)

**Construction:** (divided by population). Municipalities where construction was higher have also raised high tax revenues, which might have led to inefficient management of these increased revenues (see, for instance Silkman and Young 1982). In addition, the population levels in these municipalities might have increased sharply, which may have driven some municipalities to increase their social expenditures, as well as other expenses such as civil protection, security, etc., some of which are not included in the list of minimum services municipalities must provide. This variable is constructed using analogous criteria as those used for the industrial activities. (ENV8)

**Agricultural vehicles/Total number of vehicles** (percentage): this information would indicate whether it is a rural municipality whose needs might differ from others with different sectoral specializations. It might also be a proxy for urban sprawl. (ENV9)

**Number of bank branches:** (divided by total population). This would indicate another type of specialization and, in addition, it might proxy for the economic level of the municipality (ENV10).

We admit the selection of these variables is somewhat ad hoc. However, we consider they represent a realistic summary of the different socioeconomic conditions affecting each municipality.

Some decisions involved in performing cluster analysis are the measure of similarity as well as the clustering method. Regarding the former, one of the most popular choices is the Euclidean square distance. Regarding the latter, although there are several alternatives, the Ward hierarchical clustering method has the advantage of maximizing intra-group homogeneity and inter-group heterogeneity. In addition, the technique is robust to outliers and groups are not too dissimilar in size. However, the criterion that must ultimately guide the decision on both the methodology and the optimal number of groups is whether the final groups are sensitive in any way, and whether statistical differences exist among group centroids. Battese et al. (2004) point out the importance of analyzing whether all municipalities share the same technology. If all municipality-level data were generated from a single production function and the same underlying technology, there would be no good reason for estimating the efficiency levels of municipalities relative to a metafrontier. We assume that if statistically significant differences existed among the different groups, it would constitute evidence in favor of comparing municipalities with those in their group only.

### 4 Results

Table 5 provides summary statistics on *unconditioned* efficiency for order-$m$ efficiency scores. Results are reported for all municipalities, and also for the different size categories, given the differences in their powers. Results have been obtained by specifying a common frontier—the metafrontier—for all 1,198 observations.
Thus, although results are split into different municipality size categories, they correspond to the same common frontier. The results corresponding to all municipalities, regardless of their powers, are displayed in the first row. Average efficiency is 91.18%, which is a high value given that municipalities would become fully efficient if they were able to decrease their total costs by 8.82% only. However, this is an average effect which varies across municipalities. The values at both tails of the distribution suggest that a remarkable variety of behaviors exist, since the minimum is 23.44%, whereas the maximum is 142.07%. In the former case, cost inefficiency is high, whereas the latter refers to cases of super-efficiency—units which lie beyond the frontier and can be regarded as outliers. This finding is important, since it constitutes a clear advantage of the order-m frontiers over DEA or FDH, which are strongly affected by the existence of outliers. In the case of order-m frontiers these extreme observations are labeled as super-efficient and do not affect the efficiencies found for other observations.

Table 5 also reports order-m efficiencies for the different categories of municipalities split by population and, consequently, levels of powers. The smallest municipalities in the sample (those with populations between 1,000 and 5,000) are, on average, the most inefficient. Mean efficiency is 89.04%, close to the global mean value of 91.18%. These results are partly similar because this is, by and large, the category with the most observations. In contrast, medium sized municipalities (with populations between 5,000 and 20,000) and large municipalities (with populations over 20,000) show higher efficiencies. Not only is average efficiency higher (94.17% and 97.87%, respectively), but also the number of municipalities lying on the frontier (i.e., fully efficient municipalities) is much higher (59.79 and 90.14%, respectively).\footnote{17}

However, the most interesting results are those obtained for the different clusters, constructed either using output mix or environmental variables. Descriptive statistics (medians) on the variables used to build the clusters are provided in Table 3 (clusters based on output mix) and 4 (clusters based on environmental variables). In order to facilitate interpretations, we provide a lower panel below each table that reports a summary of the variables included to form the clusters.

Regarding the clusters based on output mixes, as reported in Table 3, differences between the municipalities in each group (in terms of the selected variables) are noteworthy, even though the size of some of these clusters is remarkable. For instance, group 1 is made up of 300 municipalities, roughly 1/4 of the total sample. Ideally, it would be desirable to have clusters containing fewer observations to facilitate comparisons. However, some clusters were difficult to split into further groups, despite considering the Ward method to cluster observations (which tends to form equally-sized clusters). This is an interesting finding, which would corroborate the fact that many municipalities indeed do different things, making comparisons misleading.

In some cases, the medians for some clusters and variables differ substantially; this is the case for cluster 2 in OUTMIX5, cluster 3 in OUTMIX6, cluster 4 in OUTMIX3, cluster 5 in OUTMIX4, cluster 6 in OUTMIX2 and OUTMIX5, cluster 7 in OUTMIX6, cluster 8 in OUTMIX2, and

\footnote{17}{For explanations for the different levels of inefficiency found see, for instance, Balaguер-Coll et al. \cite{2007}.}
cluster 9 in OUTMIX. Therefore, the clusters excel in some particular variables, even taking into account that some of them contain many observations—compared with the total sample size. In addition, although there is a wide consensus that the multivariate technique of cluster analysis is flawed, especially because of the multiple decisions it involves, the MANOVA analysis indicated that the differences between the identified groups and variables were indeed significant.18

As indicated in Table 4, the differences found among the clusters based on environmental variables (with respect to the variables included in the analysis) are also noteworthy, and significant at the 1 % level. In this case, differences are more difficult to distinguish, because of the narrow range of variation for some of these variables (for instance, ENV2 or ENV3). However, certain groups excel in some variables. See, for instance, cluster 3 in ENV1, cluster 8 in ENV4, cluster 2 and 6 in ENV7, cluster 8 in ENV8, cluster 7 in ENV9 or cluster 6 in ENV10.

4.1 Describing the clusters

As indicated in the introduction, Battese et al. (2004) propose a method for comparing the efficiencies of DMUs in different groups in the context of Stochastic Frontier Analysis, which has been extended to the DEA context by O’Donnell et al. (2008). Our methodology, based on order-m indicators, allows us to control for group membership in the context of efficiency measurement via nonparametric techniques and at the same time, takes into account the severity of the curse of dimensionality. We report the results obtained following our approach in Table 6. On average—and this result holds for both categories of clusters—municipalities are much closer to their frontiers, as documented by average efficiencies closer to unity. This result holds for both clusters based on output mix variables (99.10 %) and clusters based on environmental variables (98.92 %). In the case of clusters based on size, or level of powers, as one might a priori expect, results are quite similar to those of the unconditioned case (Table 5).

However, Table 6 also reveals that classifying municipalities into different groups does not per se explain away the remaining efficiencies. The maximum values for the three clustering criteria are well above the unity, suggesting that a non-negligible number of outliers exists.19 These are what Andersen and Petersen (1993) call super-efficient units.20

More specific results are reported in Tables 7, 8 and 9. They report basic summary statistics of the technology gap ratio, the efficiencies obtained from the group frontiers (CEg), and the metafrontier (CE*). In the case of clusters based on output mix, the widest gap between group efficiencies and metafrontier efficiencies corresponds to municipalities in cluster 2, for which average CEg = 1.0508 and average CE* = 0.9071. As a result, the technology gap ratio is the lowest (TGR2 = 0.8776). In contrast, for cluster 8 the technology gap ratio is the highest (TGR8 = 1.0009), which should be interpreted inversely, i.e., the efficiencies for municipalities in this group are very similar for the group frontier (CEg = 1.0006) and metafrontier (CE* = 1.0016). In general, although some groups show remarkable discrepancies between group efficiencies and metafrontier efficiencies (clusters 1 and 2), for many others the gap is narrower (well above 0.90). Although this result might constitute evidence against our initial hypothesis, we should bear in mind that clusters 1 and 2 are indeed the largest ones, with 300 and 407 observations, respectively. Therefore, for more than half of the municipalities in our sample, it is more reasonable to compare them only with the municipalities in their output mix group (Table 7).

Table 8 reports analogous information to that in Table 7 for clusters constructed using environmental variables. In this case, supporting evidence for our initial hypothesis is stronger as, on average, the technology gap ratios are much lower than in Table 7. The widest gap is found for cluster 8, whose average TGR8 = 0.8718, whereas the lowest gap is found for cluster 3 (TGR3 = 0.9421). However, in this case the clusters are, on average, much closer to their respective group frontiers than in the case of output mix clusters—on average, most of them show CEg values in the vicinity of 1. Therefore, regardless of the cluster considered, the environmental conditions faced by the different

---

18 Although, technically, MANOVA compares the means of the groups, not the medians, as reported in Tables 3 and 4.

19 The consideration of outliers was based on some outputs for which there were remarkable disparities among municipalities.

20 The existence of outliers also partly underlies the remarkably high average values for the efficiency scores.
municipalities play a remarkable role, leading us to mislabel them as inefficient. Note also that the differences found between \( CE^g \) and \( CE^*/C_3^* \) are irrespective of the number of observations in each cluster.

The cluster with the highest (average) discrepancies between the group efficiency and the metafrontier is cluster 8 \( (CE^g_8 = 0.8476) \), although there are also discrepancies within the group \( (CE^g = 0.9655) \). Multiple reasons related to the characteristics of the cluster could explain this relatively low level of efficiency. This group is characterized by high levels of economic status \( (ENV_3) \), industrial activities \( (ENV_4) \), number of cars \( (ENV_5) \), population growth \( (ENV_7) \) and, especially, construction activities \( (ENV_8) \), or a high number of bank branches \( (ENV_10) \), i.e. municipalities populated by rich constituencies and whose levels of economic activity is quite high. In some particular

### Table 6
Order-\( m \) efficiency scores and metatechnology ratios, all municipalities, summary statistics, year 2000

| Group                             | Index                        | Mean    | Median  | Max.    | Min.    | SD    | % Eff. obs. |
|-----------------------------------|------------------------------|---------|---------|---------|---------|-------|-------------|
| Clusters based on output mix variables | Cluster-specific efficiencies | 0.9910  | 1.0000  | 1.9583  | 0.2425  | 0.1832 |             |
|                                   | Technology gap               | 0.9249  | 0.9996  | 1.1949  | 0.3659  | 0.1343 |             |
| Clusters based on environmental variables | Cluster-specific efficiencies | 0.9892  | 1.0000  | 1.5218  | 0.2493  | 0.1447 |             |
|                                   | Technology gap               | 0.9203  | 0.9895  | 1.1682  | 0.2766  | 0.1264 |             |
| Clusters based on size (or levels of powers) | Cluster-specific efficiencies | 0.9264  | 1.0000  | 1.6983  | 0.2428  | 0.1807 |             |
|                                   | Technology gap               | 0.9849  | 1.0000  | 1.2117  | 0.4559  | 0.0598 |             |

### Table 7
Order-\( m \) efficiencies and metatechnology ratios, clusters based on output mix variables

| Group          | Mean    | Median  | Minimum | Maximum | SD    | % Eff. obs. |
|----------------|---------|---------|---------|---------|-------|-------------|
| Cluster 1      | 0.9651  | 1.0000  | 0.3222  | 1.6823  | 16.69 | 19.00       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 0.9083  | 0.9678  | 0.4094  | 1.0466  | 0.1266 |           |
| Metafrontier, \( CE^*/C_3^* \) | 0.8809  | 1.0000  | 0.2672  | 1.4207  | 20.47  |           |
| Cluster 2      | 1.0508  | 1.0000  | 0.4205  | 1.9583  | 22.78 | 38.08       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 0.8776  | 0.9529  | 0.4089  | 1.0000  | 0.1489 |           |
| Metafrontier, \( CE^*/C_3^* \) | 0.9071  | 1.0000  | 0.3772  | 1.2286  | 17.31  |           |
| Cluster 3      | 0.9300  | 1.0000  | 0.2944  | 1.1426  | 15.82 | 30.37       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 0.9751  | 1.0000  | 0.4835  | 1.0384  | 0.0870 |           |
| Metafrontier, \( CE^*/C_3^* \) | 0.9090  | 1.0000  | 0.3049  | 1.1412  | 17.95  |           |
| Cluster 4      | 0.9796  | 1.0000  | 0.3445  | 1.2002  | 10.43 | 16.38       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 0.9785  | 1.0005  | 0.3659  | 1.1949  | 0.1272 |           |
| Metafrontier, \( CE^*/C_3^* \) | 0.9611  | 1.0040  | 0.2710  | 1.3073  | 16.99  |           |
| Cluster 5      | 0.9225  | 1.0000  | 0.2425  | 1.0874  | 16.23 | 23.00       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 0.9922  | 1.0006  | 0.4667  | 1.1652  | 0.0908 |           |
| Metafrontier, \( CE^*/C_3^* \) | 0.9165  | 1.0000  | 0.2536  | 1.1682  | 18.54  |           |
| Cluster 6      | 0.9972  | 1.0000  | 0.6334  | 1.1546  | 7.49  | 48.61       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 0.9418  | 1.0000  | 0.3846  | 1.0730  | 0.1252 |           |
| Metafrontier, \( CE^*/C_3^* \) | 0.9412  | 1.0000  | 0.3923  | 1.2117  | 15.27  |           |
| Cluster 7      | 0.9942  | 1.0000  | 0.9207  | 1.0191  | 2.29  | 60.00       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 0.9883  | 1.0000  | 0.7597  | 1.0145  | 0.0494 |           |
| Metafrontier, \( CE^*/C_3^* \) | 0.9832  | 1.0000  | 0.7076  | 1.0339  | 6.19   |           |
| Cluster 8      | 1.0006  | 1.0000  | 1.0000  | 1.0088  | 0.23  | 57.14       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 1.0009  | 1.0000  | 0.8551  | 1.1033  | 0.0502 |           |
| Metafrontier, \( CE^*/C_3^* \) | 1.0016  | 1.0000  | 0.8551  | 1.1129  | 5.17   |           |
| Cluster 9      | 0.9686  | 1.0000  | 0.3371  | 1.0187  | 14.87 | 65.00       |
| Group efficiency, \( CE^g \) |  |         |  |  |       |           |
| Technology gap ratio | 0.9625  | 1.0000  | 0.6663  | 1.1369  | 0.1215 |           |
| Metafrontier, \( CE^*/C_3^* \) | 0.9411  | 1.0000  | 0.2344  | 1.1581  | 19.63  |           |
cases (ENV3, ENV4, ENV5 and ENV8) the values are the highest among the different groups. It is especially remarkable the case of population growth (ENV7), with a 16.15 % increase between 1991 and 1998, which might partly be due to both foreign-born and Spanish national migration,\(^{23}\) attracted by the relatively high industrial and construction activities of the towns in this cluster. Although this is an hypothesis that would require specific testing, some recent papers such as those by Hierro and Maza (2009) or Peri and Requena-Silvente (2010), although dealing with different issues, indicate it is a plausible conjecture. Such a remarkable population increase in a relatively short period of time might have caused some municipalities to increase the amount of services and facilities provided to their inhabitants proportionally, without this being reflected in the list of outputs (civil protection, etc.) which, as stated in the introduction, would \textit{a priori} lead to a deterioration in their efficiency levels. The low efficiency levels could also be due to the increased revenues of municipalities resulting from the increased construction activities which could be managed inefficiently. This view is supported in the literature by several studies such as Spann (1977) or Silberman and Young (1982), who indicate that, “at the local level, higher incomes increase the fiscal capacity of municipalities and may foster featherbedding of politicians and public managers, thereby increasing the scope for inefficient operation” (De Borger and Kerstens 1996).

Another cluster whose mean efficiency is relatively low is cluster 6 \((CE_6 = 0.8805)\). It is also made up of relatively rich municipalities since the median for the economic status variable (ENV3) is among the highest and the unemployment rate (ENV6) is the lowest \((2.4000)\). The variables most strongly linked to economic activity (industry and construction, i.e. ENV4 and ENV8) are also among the highest—second only to group 8. Of special note is the number of bank branches per inhabitant \((ENV10)\), which presents the highest median \((2.1816)\). The low efficiencies found for the municipalities in this group may have a myriad of causes, among which we might mention not only those pointed out in the previous paragraph but also other rationale by De Borger and Kerstens (1996), who point out that the incomes and wealth of citizens affect the incentives of both politicians and taxpayers to monitor expenditures, or the fact that rich constituencies might ask for more services and facilities. These patterns are similar for cluster 4, whose environmental conditions are similar to those found for cluster 6, although less marked (for instance, the economic status, ENV3, is slightly worse). Consequently, in this case the average group efficiency (compared to the metafrontier) is higher \((CE_4 = 0.9176)\).

In contrast to clusters 6 and 8, cluster 2 has a higher average efficiency (compared to the metafrontier, \(CE_2 = 0.9144\)) and, in addition, disparities within the cluster are lower. The trends found for population (ENV8) are similar to those found for cluster 8, since they are much higher than the rest of the clusters—for cluster 2, its median value is even higher \((25 \%)\). These similarities are also shared by other variables such as the tourist index (ENV2) or the economic status (ENV5). The median number of cars (ENV5) is also similar, which was foreseeable because it is related to economic status. However, this variable is also strongly related to urban sprawl which, as indicated by Horta-Rico and Solé-Ollé (2010), may have a relevant impact in terms of costs (and efficiency) for the municipality. Compared to the metafrontier, this cluster’s mean efficiency is higher than that of cluster 8 \((CE_2 = 0.9144 > 0.8476 = CE_8)\), as well as the number of efficient municipalities in it \((46.30 \text{ vs. } 31.82 \%)\). This superior relative efficiency might also be related to the lower levels found for the variables reflecting economic activity, either industry (ENV4) or construction (ENV8), which require the provision of some services and facilities, despite raising some additional revenues (in terms of higher taxes) to the municipality.

Cluster 3 has the highest average cost efficiency (metafrontier), \(CE_3 = 0.9326\). Some of its main characteristics in terms of environmental conditions differ substantially from those corresponding to the clusters described above. Unemployment is higher, and economic activities (industry, construction and agriculture) are relatively low. As a consequence, the number of bank branches divided by population \((ENV10)\) is also much lower. The relatively lower level of economic activities might have translated into lower municipal tax revenues. Although the average density of this cluster is the lowest \((ENV1 \text{ is the inverse of population density})\) and, as indicated by the literature (De Borger and Kerstens 1996), cost inefficiency rises with lower population density, in this case the lower revenues that require more efficient management might have offset the effect of lower density.

The rest of the clusters based on environmental variables (clusters 1, 5 and 7) also have some characteristics which differentiate the municipalities included in them from those in other clusters. Nonetheless, these characteristics are not as clear as those for the groups described above. However, the cluster analysis performed indicates that the differences among them are indeed significant.

In contrast to the clusters described in the previous paragraphs, clusters 1 and 5 are those with the lowest economic status \((ENV3)\), which is also proxied by the number of cars \((ENV5)\). This variable is also the lowest of all the clusters \((0.4057 \text{ and } 0.4087 \text{ for clusters 1 and 5, respectively})\). The

\(^{23}\) The Spanish population increased from less than 40 million in 1996 to more than 46 in 2008 (and more than 47 in 2010), mostly due to migration, but such an increase was unevenly distributed across regions.
Table 8 Order-$m$ efficiencies and metatechnology ratios, clusters based on environmental variables

| Group       | Mean | Median | Minimum | Maximum | SD  | % Eff. obs. |
|-------------|------|--------|---------|---------|-----|-------------|
| Cluster 1   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 1.0008 | 1.0000 | 0.3251  | 1.5098  | 16.03 | 22.97       |
| Technology gap ratio | 0.9155 | 0.9573 | 0.4106  | 1.0131  | 0.1019 |             |
| Metafrontier, $CE^*$   | 0.9183 | 1.0000 | 0.2672  | 1.3073  | 17.75 |             |
| Cluster 2   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9923 | 1.0000 | 0.5929  | 1.0813  | 6.15  | 46.30       |
| Technology gap ratio | 0.9200 | 1.0000 | 0.5185  | 1.0181  | 0.1381 |             |
| Metafrontier, $CE^*$   | 0.9144 | 1.0000 | 0.4611  | 1.0503  | 15.36 |             |
| Cluster 3   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9895 | 1.0000 | 0.4945  | 1.4171  | 11.24 | 58.33       |
| Technology gap ratio | 0.9421 | 1.0000 | 0.2766  | 1.0003  | 0.1242 |             |
| Metafrontier, $CE^*$   | 0.9326 | 1.0000 | 0.2768  | 1.2117  | 15.88 |             |
| Cluster 4   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9811 | 1.0000 | 0.2493  | 1.5218  | 18.63 | 32.84       |
| Technology gap ratio | 0.9358 | 0.9788 | 0.4822  | 1.0070  | 0.0949 |             |
| Metafrontier, $CE^*$   | 0.9176 | 1.0000 | 0.2344  | 1.4207  | 18.82 |             |
| Cluster 5   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 1.0022 | 1.0000 | 0.4365  | 1.2965  | 10.10 | 42.31       |
| Technology gap ratio | 0.9187 | 0.9985 | 0.3577  | 1.0045  | 0.1449 |             |
| Metafrontier, $CE^*$   | 0.9209 | 1.0000 | 0.3049  | 1.1232  | 16.58 |             |
| Cluster 6   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9557 | 1.0000 | 0.3818  | 1.2284  | 13.82 | 13.61       |
| Technology gap ratio | 0.9190 | 1.0000 | 0.4005  | 1.1682  | 0.1547 |             |
| Metafrontier, $CE^*$   | 0.8805 | 1.0000 | 0.3115  | 1.2286  | 20.37 |             |
| Cluster 7   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9987 | 1.0000 | 0.8312  | 1.1401  | 4.04  | 39.44       |
| Technology gap ratio | 0.9090 | 1.0000 | 0.3203  | 1.0829  | 0.1680 |             |
| Metafrontier, $CE^*$   | 0.9095 | 1.0000 | 0.3203  | 1.1193  | 17.56 |             |
| Cluster 8   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9655 | 1.0000 | 0.3860  | 1.0712  | 13.03 | 31.82       |
| Technology gap ratio | 0.8718 | 0.9760 | 0.3092  | 1.1522  | 0.1856 |             |
| Metafrontier, $CE^*$   | 0.8476 | 0.9791 | 0.3092  | 1.1568  | 22.54 |             |

Table 9 Order-$m$ efficiencies and metatechnology ratios, clusters based on size (levels of powers)

| Group       | Mean | Median | Minimum | Maximum | SD  | % Eff. obs. |
|-------------|------|--------|---------|---------|-----|-------------|
| Cluster 1   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9011 | 1.0000 | 0.2428  | 1.6983  | 20.08 | 10.68       |
| Technology gap ratio | 0.8972 | 0.9773 | 0.3659  | 1.1949  | 0.1541 |             |
| Metafrontier, $CE^*$   | 0.8904 | 1.0000 | 0.2344  | 1.4207  | 20.48 |             |
| Cluster 2   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9665 | 1.0000 | 0.3955  | 1.6893  | 13.88 | 59.79       |
| Technology gap ratio | 0.9683 | 1.0000 | 0.5640  | 1.1364  | 0.0749 |             |
| Metafrontier, $CE^*$   | 0.9417 | 1.0000 | 0.3772  | 1.1364  | 13.43 |             |
| Cluster 3   |      |        |         |         |     |             |
| Group efficiency, $CE^g$ | 0.9788 | 1.0000 | 0.6376  | 1.0020  | 0.73  | 90.14       |
| Technology gap ratio | 0.9866 | 1.0000 | 0.7564  | 1.0000  | 0.0491 |             |
| Metafrontier, $CE^*$   | 0.9787 | 1.0000 | 0.6376  | 1.0000  | 7.30  |             |

economic activity variables, not only industry ($ENV4$) and construction ($ENV8$) but also agriculture ($ENV9$)—measured by the number of agricultural vehicles—, are also the lowest. In the case of cluster 1, the tourist index ($ENV2$) is also the lowest. In these dismal environments one may reasonably expect that the levels of unemployment will be among the highest, especially for cluster 5 (5.7 %), and that population growth ($ENV7$) will either be negative (cluster 1) or moderate (cluster 5). The levels of mean efficiency in these clusters (compared with the metafrontier) are among the highest (0.9183 and 0.9204 for clusters 1 and 5, respectively), and clearly higher than those found for clusters 6 and 8, whose main environmental conditions are in direct contrast. Therefore, the reasons that might explain the relative underperformance of clusters 6 and 8 apply here, with the opposite sign.

4.2 Densities for each specific cluster

Figure 2 shows densities estimated using kernel smoothing of unconditioned and cluster-specific frontiers. The tighter
probability mass at unity shows that observations are indeed much closer to those in their groups than to observations in other groups. These differences are also significant, as indicated by the $p$ values in Table 10, which were obtained by applying the Simar and Zelenyuk (2006) test, whose details are provided in “Appendix”. The substantial amount of probability mass found at the upper tail of the output mix distribution (Fig. 2a) indicates that controlling for group membership contributes in a more modest way of explaining efficiency differentials than in the case of clusters formed using environmental variables.

Figures 3, 4 and 5 provide analogous information to that in Fig. 2, revealing that the differences exist for most groups, especially when conditioning either for output mix or environmental conditions. This is shown by tighter densities when conditioning either for output mix or environmental group membership—solid lines in Figs. 3, 4.

Table 10 Distribution hypothesis tests (Li 1996; Simar and Zelenyuk 2006)

| Efficiency distributions compared          | $t$-Statistic | $p$ Value |
|--------------------------------------------|---------------|-----------|
| Metafrontier vs. output mix clusters       | 3.4814        | 0.0002    |
| Metafrontier vs. environmental clusters    | 7.8215        | 0.0000    |
| Metafrontier vs. size (levels of powers)   | 2.7929        | 0.0026    |
5 Concluding remarks

Over the last few years, a relevant area of research in the field of public economics and regional science and urban economics has been the analysis of the efficiency of lower layers of government such as regional governments or, as in the present study, local governments. The topic is not only relevant per se, but also because of recent events such as the economic and financial downturn. In some euro-area countries, economic activity has stalled, resulting in a sharp reduction of tax revenues and a simultaneous rapid increase in public spending. In these circumstances, the efficient management of public resources at all levels of government becomes even more important. This is particularly pertinent from a viewpoint of European integration and the Stability and Growth Pact (SGP), in order to facilitate and maintain European Economic and Monetary Union.

Interest is even higher in certain European countries such as Spain, where municipalities face tighter budget constraints since the approval of the law on budget stability (“Ley General de Estabilidad Presupuestaria”, 2001), which establishes mechanisms to control public debt and public spending in pursuit of a balanced budget. In addition, the current economic and financial scenario that has emerged since 2008 has led this country into a recession of unprecedented depth and duration. A critical consequence...
has been the deterioration of the country’s fiscal position, which has worsened dramatically since 2007, contrasting with the regular and significant improvement since the recession of 1993—from a surplus of 1.9 % of GDP in 2007, the fiscal balance moved to a deficit of 11.1 % in 2009. Since most fiscal deterioration is structural, stronger measures to cut expenditures or raise taxes are necessary to improve the fiscal balance. These measures are contemplated in the update of the stability program for Spain (“Plan de Acción Inmediata (immediate action plan) 2010” and “Plan de Austeridad (austerity plan) 2011–2013”), which combines a tight control of public expenditures with a modest increase of revenues, in line with empirical research which shows that expenditure-driven consolidations tend to be more sustainable than revenue-driven consolidations (Guichard et al. 2007).

There is now a relatively large body of literature devoted to the analysis of municipality efficiency. However, although the number of theoretical and applied contributions is high, some decisive issues still remain unsolved. One refers to the definition of municipality output. This is a thorny question, especially if we take into account that municipalities are facing budget constraints while at the same time the law requires them to provide a minimum amount of services and facilities. A related issue is the remarkably varied environmental conditions—in our case defined as socioeconomic variables—that each municipality faces, which may have a marked effect on their performance.

We deal with these issues using a three-stage procedure. In the first stage, efficiency is assessed unconditionally. In the second stage, municipalities are classified into groups using cluster analysis taking into account variables based on their output mix and environmental conditions. We identify groups of municipalities that share some important features in these fields, and the differences among them were statistically significant. In the third stage we assess how results vary when comparing each municipality should be compared with those in its group rather than with all municipalities.

The efficiency literature has dealt with the existence of groups and varying environmental conditions following different criteria. Several approaches now exist to handle this and related issues (see, for instance, Banker and Morey 1986a, b; Ray 1991; Daraio and Simar 2007b; De Witte and Kortelaainen, 2008, among others). We focus on one which has gained popularity over the last few years, namely that proposed by Battese and Rao (2002), Battese et al. (2004), or O’Donnell et al. (2008), who compares the decision making units in different groups under the assumption that technology differs across groups and, therefore, one should estimate both group frontiers and a metafrontier. For our particular interests, one of the advantages of these methods is, precisely, their ability to handle different observations in different groups. We deal with these concepts in the context of the order- \( m \) frontiers proposed by Cazals et al. (2002), in order to tackle relevant problems such as the existence of outliers, or the curse of dimensionality (see 20 Simar and Wilson 2008, p.441).

Our results indicate that both hypotheses are relevant, especially that referring to the importance of variations in environmental conditions—i.e., it is essential to control for the environment surrounding each municipality. This is particularly important in contexts, such as Spain, with a very high number of municipalities, which differ in several dimensions that go beyond the definition of inputs and outputs, and that cannot be considered as determinants of efficiency. If we factor in the financial strains some municipalities are now experiencing, especially due to the drop in municipal revenues brought about by the burst of the housing bubble (Spanish municipalities are responsible for approving construction permits) we realize the importance of having an additional instrument to measure their performance more accurately.

Although the literature on efficiency has proposed some ways to deal with environmental conditions, our methods are more robust to outliers, and they alleviate the curse of dimensionality. However, we must admit that the groups constructed, both using output mix and environmental variables, were somewhat subjective in terms of number of groups and composition because of the cluster analysis technique employed in their formation. Ideally, the research agenda should address how to define groups more objectively when natural boundaries between them do not exist.

Acknowledgments We are grateful for the helpful comments by Léopold Simar, Philippe Vanden Eckaut, Javier Perote, Esther del Brío, Knox Lovell, Christopher O’Donnell, Valentin Zelenyuk, Prasada Rao, Antonio Peyrache, Alicia Rambaldi and other participants at the XI European Workshop on Efficiency and Productivity Analysis (Pisa, June 2009), 49th European Congress of the Regional Science Association International ( Lodz, 2009), XVII Encuentro de Economía Pública (Murcia, 2010), University of Salamanca Seminar (Salamanca, Spain), and CEPA Seminar (University of Queensland, Brisbane, December 2010), as well as those by the three referees which contributed substantially to the overall improvement of the paper. All three authors acknowledge the financial support of Ministerio de Ciencia e Innovación (ECO2010-18967/ECON and ECO2011-27227). Maria Teresa Balaguera-Coll and Emili Tortosa-Ausina also acknowledge the financial support of Fundació Caixa Castelló-Bancaixa (P1.B2009-54) and Generalitat Valenciana (PROMETEO/2009/066), respectively. The usual disclaimer applies.

Appendix: Testing the closeness between efficiency distributions

Once efficiency scores have been estimated there are multiple ways to display and compare results (El- Mahgary
and Lahdelma 1995). We consider some methods which provide us with more accurate information (see, for instance Li 1996; Li et al. 2009). If we based our interpretations on a number of summary statistics only, we would miss a considerable amount of relevant information. Most of these methods are based on kernel smoothing to nonparametrically estimate the density functions corresponding to both \( z^m_a \) and \( z^m_{Sc} \) indices.\(^{22}\) The kernel density estimator \( \hat{f} \) for a univariate density \( f \) based on a sample of \( S \) efficiency indices (either \( z^m_a \) or \( z^m_{Sc} \)) is \( \hat{f}(x) = (Sh)^{-1} \sum_{i=1}^{S} K((x - z^m_i)/h) \), where \( x \) is the municipality index, \( z^m_i \) represents its efficiency index, \( x \) is the evaluation point, \( h \) is the bandwidth, and \( K \) is a kernel function satisfying certain properties. Additional decisions concerned the kernel and bandwidth choice. For choice of kernel, we considered the Gaussian kernel for its ease of computation. Regarding bandwidth selection, the plug-in method suggested by Wand and Jones (1994) is increasingly adopted in the literature.\(^{23}\)

We must control for the fact that efficiency indices are bounded between (0,1]. The Silverman (1986), Schuster (1985) reflection method is based on the idea of “reflecting” the probability mass lying beyond unity—where, theoretically, no probability mass should exist. The kernel estimate disregarding the boundary condition can be shown to be biased and inconsistent (Simar and Wilson 1998).\(^{24}\)

After estimating densities, the Li (1996) test shows whether the observed visual differences are statistically significant or not. The test is based on measuring the distance between densities \( f(x) \) and \( g(x) \) for which we consider the mean integrated square error, i.e.:

\[
L = L(f(x), g(x)) = \int \frac{(f(x) - g(x))^2}{x} dx
\]

\[
= \int \frac{(f^2(x) + g^2(x) - 2f(x)g(x))}{x} dx
\]

\[
= \int \frac{(f(x)df(x) + g(x)dg(x) - f(x)dg(x))}{x}
\]

where \( F \) and \( G \) are cumulative density functions for the distribution of \( X \), with density functions \( f(x) \) and \( g(x) \) and \( f \) is the nonparametric kernel estimator referred to above.

Given that \( \hat{f} = (Sh)^{-1} \sum_{i=1}^{S} K((x - z^m_i)/h) \), and \( \hat{g} = (Sh)^{-1} \sum_{i=1}^{S} K((y - z^m_{Sc_i})/h) \) then, in practice, an estimator for \( L \) is:

\[
\hat{L} = \int (\hat{f}(x) - \hat{g}(x))^2 dx
\]

\[
= \frac{1}{S(S-1)h} \sum_{i=1}^{S} \sum_{j=1}^{S} K\left(\frac{z^m_i - z^m_j}{h}\right) + K\left(\frac{z^m_{Sc_i} - z^m_{Sc_j}}{h}\right)
\]

\[
- K\left(\frac{z^m_i - z^m_{Sc_i}}{h}\right) - K\left(\frac{z^m_j - z^m_{Sc_j}}{h}\right)
\]

(4)

The integrated square error is also required for estimating the statistic on which the test is based, the expression for which is as follows:

\[
T = \frac{Sh^{1/2}L}{\hat{\sigma}}
\]

where

\[
\hat{\sigma} = \frac{1}{S^2h^{1/2}} \sum_{i=1}^{S} \sum_{j=1}^{S} \left[K\left(\frac{z^m_{Sc_i} - z^m_{Sc_j}}{h}\right)
\right] + 2K\left(\frac{z^m_i - z^m_{Sc_i}}{h}\right)
\]

(6)

and \( h \) is the bandwidth.

The Li (1996) test requires some assumptions to be met such as independently distributed efficiency scores in each sub-group (and identically within each sub-group). However, order-\( m \) efficiency estimates are dependent in the statistical sense, since perturbations of observations lying on the estimated frontier will in most cases cause changes in efficiencies estimated for other observations. As indicated by Simar and Zelenyuk (2006), in these circumstances the Li (1996) test has to be modified in several ways. First, we must control for the issue of bounded support in order-\( m \) density estimation. However, as demonstrated by Simar and Zelenyuk (2006), although controlling for the boundary effect is important in density estimation, the statistic based on the reflection method is “essentially the same as the original Li (1996) test, with the difference being a factor of \( \sqrt{2} \) and the fact that the bandwidth used in estimation of the statistic is obtained from the data with reflection rather than the original data”. Therefore, since the independency issue is not negligible, we should then follow Simar and Zelenyuk (2006), who provide a way of adapting the Li (1996) test to the order-\( m \) context via bootstrapping techniques to improve its performance. These authors provide consistent bootstrap estimates of the \( p \) values of the Li (1996) test as follow:

\(^{22}\) Several monographs cover this topic in depth. See, for instance, Silverman (1986), Pagan and Ullah (1999) or, more recently, Li and Racine (2007).

\(^{23}\) It can also be easily accessed, as it is implemented in many software packages such as R.

\(^{24}\) Although we used a variety of statistical procedures to perform all computations, the FEAR package by Paul W. Wilson for the statistical software R provides codes for computing both efficiencies and densities, considering the reflection method. See http://business.clemson.edu/Economic/faculty/wilson/.
\[ \hat{p} = \frac{1}{B} \sum_{b=1}^{B} I\{\hat{L}^b > \hat{L}\}, \]  

where \( b = 1, \ldots, B \) is the number of bootstrap replicates. The \( p \) values are then adapted to our context, where the true efficiency scores are replaced by order-\( m \) estimates. Simar and Zelenyuk (2006) consider somewhat \textit{ad hoc} methods to solve the discontinuity problem generated by the spurious probability mass at unity—recall that, by construction, at least one observation will always be on the frontier, and in most circumstances the number will be quite large. We adopt one of their proposed methods (Algorithm II; see Simar and Zelenyuk 2006), based on computing and bootstrapping the Li (1996) statistic using the sample of order-\( m \) estimates where those equal to unity are “smoothed” away from the boundary. We add a small noise, within, say, 5% of the empirical distribution of \( \hat{z}^m_x \), disregarding those equal to unity, but with an order of magnitude smaller than the noise of the estimation. The smoothing procedure is performed via:

\[ \hat{z}^m_x = \begin{cases} \frac{\hat{z}^m_x + \epsilon_x}{2}, & \text{if } \hat{z}^m_x = 1; \\
\hat{z}^m_x, & \text{otherwise.} \end{cases} \]  

where \( \epsilon_x = \text{Uniform}(0, \min\{S^{-2/(f+1)}, a\}) \) (for the DEA estimator), \( a \) is the \( x \)-quantile of the empirical distribution of \( \hat{z}^m_x \) ignoring those equal to unity.

References

Andersen P, Petersen NC (1993) A procedure for ranking efficient units in data envelopment analysis. Manage Sci 39(10):1261–1264
Balaguer-Coll MT, Prior D, Tortosa-Ausina E (2007) On the determinants of local government performance: a two-stage nonparametric approach. Eur Econ Rev 51(2):425–451
Balaguer-Coll MT, Prior D, Tortosa-Ausina E (2010a) Decentralization and efficiency of local government. Ann Reg Sci 45:571–601
Balaguer-Coll MT, Prior D, Tortosa-Ausina E (2010b) Devolution dynamics of Spanish local government. Environ Plan A 42(6):1476–1495
Banker RD, Morey RC (1986a) Efficiency analysis for exogenously fixed inputs and outputs. Oper Res 34:513–521
Banker RD, Morey RC (1986b) The use of catégoriues variables in data envelopment analysis. Manage Sci 32:1613–1627
Barankay I, Lockwood B (2007) Decentralization and the productive efficiency of government: evidence from Swiss cantons. J Public Econ 91(5–6):1197–1218
Battese GE, Rao D (2002) Technology gap, efficiency, and a stochastic metafrontier function. Int J Bus 1(2):87–93
Battese GE, Rao D, O’Donnell CJ (2004) A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. J Prod Anal 21(1):91–103
Bennett JT, DiLorenzo TJ (1982) Off-budget activities of local government: the bane of the tax revolt. Public Choice 39(3):333–342
Bonaccorsi A, Daraio C (2008) The differentiation of the strategic profile of higher education institutions. New positioning indicators based on microdata. Scientometrics 74(1):15–37
Bos JW, Kool CJM (2006) Bank efficiency: the role of bank strategy and local market conditions. J Bank Fin 30(7):1953–1974
Bosch N, Pedraja F, Suárez-Pandiello J (2000) Measuring the efficiency of Spanish municipal refuse collection services. Local Govt Surv 26(3):71–90
Bradford D, Malt R, Oates W (1969) The rising cost of local public services: some evidence and reflections. Natl Tax J 22:185–202
Brueckner JK (1981) Congested public goods: the case of fire protection. J Public Econ 15:45–58
Brueckner JK, Wingler TL (1984) Public intermediate inputs, property values, and allocative efficiency. Econ Lett 14:245–250
Camacho AS, Dyson RG (2008) A generalisation of the Farrell cost efficiency measure applicable to non-fully competitive settings. OMEGA 36(1):147–162
Cazals C, Florens J-P, Simar L (2002) Nonparametric frontier estimation: a robust approach. J Econom 106:1–25
Cordero-Ferrera JM, Pedraja-Chaparro F, Salinas-Jiménez J (2008) Measuring efficiency in education: an analysis of different approaches for incorporating non-discretionary inputs. Appl Econ 40(10):1323–1339
Daraio C, Simar L (2005) Introducing environmental variables in nonparametric frontier models: a probabilistic approach. J Prod Anal 24:93–121
Daraio C, Simar L (2007a) Advanced robust and nonparametric methods in efficiency analysis. Methodology and applications. Studies in productivity and efficiency. Springer, New York
Daraio C, Simar L (2007b) Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. J Prod Anal 28(1):13–32
De Borger B, Kerstens K (1996) Cost efficiency of Belgian local governments: A comparative analysis of FDH, DEA, and econometric approaches. Reg Sci Urban Econ 26:145–170
De Borger B, Kerstens K (2000) What is known about municipal efficiency: the Belgian case and beyond. In: Blank J (eds) Public provision and performance: contributions from efficiency and productivity measurement. El, Amsterdam, pp 299–330
De Witte K, Dijkgraaf E (2009) Mean and bold? On separating merger economies from structural efficiency gains in the drinking water sector. J Oper Res Soc 61(2):222–234
De Witte K, Geys B (2011) Evaluating efficient good provision: theory and evidence from a generalised conditional efficiency model for public libraries. J Urban Econ 69(3):319–327
De Witte K, Kortelainen M (2008) Blaming the exogenous environment? Conditional efficiency estimation with continuous and discrete environmental variables. Discussion papers 833, Center for Economic Studies, Leuven, Belgium
Deller SC (1992) Production efficiency in local government: a parametric approach. Public Fin Fin Publiques 47:32–44
Deprins D, Simar L, Tulkens H (1984) Measuring labor-efficiency in public economic and social services: some evidence and reflections. Natl Tax J 22:185–202
El-Gamal M, Inanoglu H (2005) Inefficiency and heterogeneity in Turkish banking: 1990–2000. J Appl Econom 20:641–664
El-Mahgary S, Lahlouma R (1995) Data envelopment analysis: visualizing the results. Eur J Oper Res 83(3):700–710
Färe R, Grosskopf S, Lovell CAK, Pasurka C (1989) Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. Rev Econ Stat 71(1):90–98
Ferrier GD, Lovell CAK (1990) Measuring cost efficiency in banking: econometric and linear programming evidence. J Econom 46:229–245

( Springer)
Wilson PW (1995) Detecting influential observations in data envelopment analysis. J Prod Anal 6(1):27–45
Worthington A, Lee B (2008) Efficiency, technology and productivity change in Australian universities, 1998–2003. Econ Educ Rev 27(3):285–298
Zafra Gómez JL, Prior D, Plata Díaz AM, López Hernández AM (2011) Reducing costs in times of crisis: delivery forms in small and medium sized local governments’ waste management services. Public Administration, forthcoming