Journal of Environmental Management 348 (2023) 119141

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

The optimal size of territorial units in the outsourcing of waste service: An empirical approach

Roberto Benedetti, Marialisa Mazzocchitti *, Alessandro Sarra

Department of Economics, "G. d'Annunzio" University, Chieti, Pescara, Italy

ARTICLE INFO

Handling Editor: Jason Michael Evans

Keywords:
municipal solid waste
Regulation
stochastic frontier analysis (SFA)
economies of scale
competition for the market
Transaction Cost Theory

ABSTRACT

This study provides new insights into the relation between the size of the territorial units for which the provision of waste services is entrusted to external operators and their efficiency. The need for a deeper investigation of this relation arises from the fact that on the one hand, the European Union is pushing its Member States towards the adoption of laws aimed at ensuring competition for the market whenever competition in the market is possible; on the other, both the theoretical and the empirical literature cautions policymakers against the risks associated with contracting out. This problem is addressed here through an efficiency analysis of Italian municipalities in the organization of waste services. The stochastic frontier analysis (SFA) is applied to estimate efficiency scores for the municipalities. Then, a regression analysis is carried out to investigate the relation between the efficiency scores and the size of the municipality. This analysis has been carried out using data on a sample of 6,916 Italian municipalities (87.39% of the entire population) for the year 2019. The method adopted here can also support regulatory authorities in defining the size of the territorial units in which other types of local public service should be outsourced.

1. Introduction

During the last two decades, the European Union has been exercising continual pressure towards a re-organisation of municipal waste services based on the withdrawal of the local governments from their direct provision, and to their assignment to external operators (public, private or mixed).[1] The basic idea was that in this way, a process of industrialization of the service would start with the gradual birth of new operators and an effective competition for the market would eventually emerge. Such changes would bring great advantages to local communities in terms of the quality of the service, improved efficiency and lower costs.

The ability of this strategy to fully exploit its beneficial potential, anyway, is strictly connected to the ease with which it allows selecting the most efficient service providers and to its effectiveness in minimizing the regulatory costs. Both these components seem to be influenced by scale factors. The first because the size of the population to be served has a significant role in attracting bigger operators, the ones which are presumably organized to fully exploit technical economies of scale, and which are usually less likely to operate in small contexts. The second because bigger awarding procedures give bargaining advantages to the contracting entity and allow saving on the regulatory and monitoring costs. This point is not new in either the theoretical or the empirical literature.

Neo-institutional economic theory (Williamson, 1991, 1997, 1999) suggests that a fundamental decision in an economic organization is whether to *make or buy*. The decision to internally produce a service or to contract it out to an external provider is influenced by the relative costs of direct production and the transaction costs, which are the result of bounded rationality, asymmetrical information and opportunistic behaviour by the agents (Williamson, 1997).

In the case of local public services like waste collection, for example, when an external service provider (be it private, public or mixed) has more information about its activities and performance than the contracting authority (local government), it can behave opportunistically to increase profits (or internal benefits in the case of publicly owned operators). In such a case the local government must engage in more precontract preparation and post-contract oversight—high transaction

* Corresponding author.
E-mail addresses: roberto.benedetti@unich.it (R. Benedetti), marialisa.mazzocchitti@unich.it (M. Mazzocchitti), alessandro.sarra@unich.it (A. Sarra).

[1] In Italy, waste service can be a) independently managed by the municipality, or b) entrusted to an external operator. In case b), a public, or private, or mixed company can: b1) obtain the in-house award of the service; or b2) obtain the entrustment of the service by winning a tender.

https://doi.org/10.1016/j.jenvman.2023.119141
Received 17 May 2023; Received in revised form 9 September 2023; Accepted 22 September 2023
Available online 11 October 2023

costs—to mitigate the risks and improve compliance (Brown and Potoski, 2003).

This particular stream of the literature has provided substantial empirical evidence proving that small municipalities obtain less benefit from individually outsourcing the service. Bel and Fageda (2006) empirically support the hypothesis that small municipalities have little bargaining power and incur relatively high monitoring costs. In another work (Bel and Fageda, 2008) they show that the service-related transaction costs for contracting out to an external operator are higher, in proportion, than those incurred by bigger municipalities. Perfectly in line with what is sustained in the present paper, Zafra-Gómez et al. (2013) suggest that some postulates of the New Public Management paradigm, such as the preference for the outsourcing of services, might need to be reconsidered, as they do not always allow achieving better results regarding the cost of the services provided. They argue that this could be due to the fact that smaller local authorities enter into poorly specified contracts or incur high service monitoring costs. This gives rise to high transaction costs, which is then aggravated by a lack of competition among the private suppliers, who show little interest in the relatively small contracts offered by these local authorities. Finally, Bel et al. (2014) show that inter-municipal cooperation leads to lower service provision costs, so that it is the rational choice for municipalities with a suboptimal size.

Since these conclusions seem to confirm the existence of scale advantages in the outsourcing of the service, they would have suggested organizing municipalities in joint powers authorities, entrusting them to manage the outsourcing and monitoring. This solution has been adopted in some national contexts (e.g., Italy and Spain). Unfortunately, even if the empirical literature is quite clear about the existence of economic disadvantages for small municipalities, it is less so about whether there is an optimal population size for the territorial unit[2] which is to outsource the service to a single provider, and, if such an area exists, how to determine it. A similar question has been already treated in a number of papers using Data Envelopment Analysis (DEA), a well-established non-parametric technique for efficiency benchmarking. With that purpose, Sarra et al. (2020a, 2020b) implemented a multi-stage DEA procedure aimed at measuring the influence of population (treated as an environmental variable with respect to the collection process) on the efficiency of waste services, measured from the municipal budget point of view. The conclusion they obtain with reference to the Italian case, seems somehow counterintuitive. In fact, using different techniques, they provide evidence not for the existence of a minimum efficient level of population served, but for an optimal upper population threshold, beyond which no efficiency gain can be obtained through the externalisation to a single provider.

Their approach, anyway, used a notion of efficiency (based on the joint consideration of environmental and economic factors) which seems no longer suitable for empirical studies given the significant changes that have occurred during the last few years.

In particular, they build a DEA model treating waste from separate collection as the output of a municipal "production process", in which municipal costs are the input, and unsorted waste is an undesired output to be minimized (Sarra et al., 2017). That approach was justified by the fact that, at the time, almost none of the municipalities considered in the analysis had reached the thresholds of separate collection laid down by national legislation (Legislative Decree no. 152/2006, the so-called Environmental Code), and that significant differences in the amount of separate collection and in the total expenditure could be observed. In fact, implementing an effective system of separate collection required significant investments that most municipalities were not able to make, given the tight budgetary constraints to which they were subject. Different operational solutions were then adopted, that led to different combinations of sorted and unsorted waste, depending on the financial resources available at the local level and the local political priorities.

This situation has significantly changed in recent years. A significant share of Italian municipalities has reached the regulatory thresholds, so that the use in the DEA model of an undesirable output is no longer justified. When there is no need to deal with an undesirable output, the best analytical option becomes the use of a parametric method for efficiency benchmarking such as Stochastic Frontier Analysis (SFA) in association with econometric tools aimed at measuring the impact of external variables – in the first place the population served – on the efficiency scores (Aigner et al., 1977; Meeusen and van Den Broeck, 1977). The choice fell on SFA mainly because it is the method used in the past in studies on the optimal size of local governments (Vidoli et al., 2023), passenger bus service concessions (Albalate et al., 2023), municipal tax departments (Niaounakis and Blank, 2017), banks (Kumbhakar and Peresetsky, 2013), and water distribution utilities (Filippini et al., 2008). The approach adopted in the present paper, ensures a significant improvement of the analysis for two reasons: it allows measuring the statistical significance of the results obtained, and it allows avoiding the problem of the "separability" of the environmental variables, a typical problem of DEA which prevents the use of efficiency scores in second stage regressions.

In this study, the method of SFA is applied to estimate a production frontier which will be used to estimate the efficiency of a sample of 6916 Italian municipalities (87.39% of the entire population) in the organization of waste management services. In the second step of the analysis, the relation between the efficiency scores and the population served has then been investigated through a regression analysis, with the aim of identifying any threshold value of the dimensional variable beyond which the efficiency scores stably decrease. The population served is considered as the variable representing the size of the territorial unit in which the service should be subject to a single process of outsourcing.

This paper is organized as follows. Section 2 gives a brief overview of earlier studies about economies of scale in the organization and management of waste services. Section 3 describes the methods chosen for this research. Section 4 introduces the sample, data, variables and empirical models used in this work. Section 5 shows the results obtained. Section 6 discusses the main results and suggests policy implications. Some conclusions are drawn in the final section.

2. Literature review

In the past, quite an extensive literature has dealt with the role of scale factors in the economics of urban waste collection. This literature was mainly focused on the optimal operational size of service providers (as a consequence of the exploitation of technical economies of scale) and, therefore, relied on data retrieved from waste utilities. Different methods were applied, which led to very different, and sometimes conflicting, results.

Stevens (1978), for example, collected data on waste utilities operating in 340 American cities in the two-year period 1974–1975, in order to estimate a logarithmic total cost function. The research found economies of scale for cities up to 20,000 inhabitants and constant returns to scale for populations greater than 50,000. Carvalho et al. (2015) estimated a translog cost function using the SFA from data on 184 Australian waste utilities for the years 2000–2001 and 2005–2006; the empirical research showed that the optimal size lay in the range of 12,000–20,000 served inhabitants. Simões et al. (2010) applied a double bootstrap DEA model to a cross sectional sample of 29 Portuguese waste utilities for the year 2007 and found that the optimal operational size was around 300,000 inhabitants served. Carvalho and Marques (2014), also investigating Portugal, derived a translog cost function by applying

[2] A territorial unit is a geographical area that can correspond to a municipality, or to a sub- or supra-municipal area. In this paper, the term *territorial unit* is preferred to *municipality* because the optimal size is not necessarily that of a municipality but can be larger or smaller. Correspondingly, the contracting authority responsible for a territorial unit can be a municipality or another government body, for example, an intermunicipal authority.

the SFA to a sample of 37 waste utilities using panel data from 2006 to 2010. The research showed that the optimal size lies in an interval of 400,000–550,000 inhabitants.

These, and many others similar approaches, gave a fundamental contribution to the study of the organizational efficiency of service providers; but some significant regulatory problems need to be addressed from a different perspective. Once the service is outsourced, waste utilities can reach the optimal scale by operating in a number of different cities. What is important from the regulatory point of view, instead, is determining the optimal size of a single territorial unit for which the service should be entrusted to an external operator, the one which allows minimizing the costs borne by the municipalities. This size should be such as to maximize the probability of selecting a utility which fully exploits any economies of scale (a prerequisite to minimizing the operational costs that the awarded utility transfers to the municipalities), and to minimize the regulatory and control costs. At any rate, this kind of analysis should be based on data collected from the municipal budgets and not from the utilities' financial accounts. Unfortunately, in this regard, the available literature does not offer a broad spectrum of investigations.

A number of studies focused on the effects of inter-municipal cooperation, trying to ascertain whether there were economies of scale in the production function of waste collection services and how significant they are through the estimation of a cost function. Bel and Costas (2006) used a sample of 186 Spanish municipalities for the year 2000 to estimate a logarithmic cost function. They determined that inter-municipal cooperation had a visible effect on cost savings for municipalities below 20,000 inhabitants. In a similar study considering 56 municipalities in the Aragon region for the year 2003, Bel and Mur (2009) found a positive effect of inter-municipal cooperation on costs savings for municipalities with populations between 5000 and 10,000 inhabitants. Bel and Fageda (2010) used a sample of data of 65 Galician municipalities for the year 2005. They found that municipalities up to 50,000 inhabitants can exploit economies of scale, therefore implying that cooperation between small municipalities can lead to cost savings. Bel et al. (2014) aimed to determine whether small municipalities can reduce the costs of municipal waste services through cooperation. They estimated two cost functions to control for the effects of both direct provision of the service and inter-municipal cooperation. Their results showed that smaller municipalities can obtain significant cost savings through cooperation. Quite different results were obtained by Abrate et al. (2014). They considered 529 Italian municipalities for the period 2004–2006 in order to run a composite cost function model and look for economies of scale and scope economies in the waste disposal and recycling services. They concluded that in municipalities up to 42,500 inhabitants the refuse collection industry exhibits aggregate constant returns to scale and modest economies of scope. As the population increases, economies of scope expand up to a population of 300,000 inhabitants, but moderate overall diseconomies of scale appear.

The common feature of such studies consists in the fact that they limit their attention to data on inputs and outputs of the collection service considered as a municipal process (fuel consumption, workforce, tons of waste, etc.), directly including population as the scale factor in the regressions. This approach would give accurate results only if each municipality in the sample organizes the service on its own, or it is served by an external provider which is active only in that specific municipality. But if the same service provider serves more than one municipality, the specific costs the municipality bears could have little relation with the size of its population. Instead, they depend on the size of the service provider and on how it allocates its own costs among the municipalities in which operates. Moreover, this approach does not consider the regulatory and control costs the municipalities bear, which are, instead, the main variables under their control.

With a broader perspective, a few studies tried to estimate the optimal population size using non-parametric methods for measuring the performance. This approach has some advantages over the econometric estimation of a cost function. First of all, it allows properly taking into consideration the population, which is not an input of the collection process nor a purely environmental variable but is strictly correlated to the total amount of waste produced by the community. Furthermore, it allows specifying the notion of efficiency to adopt, and to include the consideration of dimensions, such the environmental one, which cannot be part of a pure cost function.

Along this line of research, Bohm et al. (2010) collected data on 299 Belgian municipalities for the year 2003 and applied DEA, to find that municipalities that were part of a supra-local joint venture were the most efficient. Considering a notion of joint economic and environmental efficiency, Sarra et al. (2017) applied a DEA-based method to a sample of 289 Italian municipalities for the years 2011–2013 and concluded that the entrustment to an operator of waste service regarding supra-municipal territories can lead to improved environmental and economic efficiency. Sarra et al. (2020b) applied a multi-stage DEA procedure to data relative to a set of 377 Italian municipalities for the years 2013, 2015 and 2017 with the aim of defining the optimal territorial size of the supra-municipal territories, which they determined to be approximately 57,000 inhabitants served.

Scholars are well aware of the drawbacks of the use of conventional DEA. They are mainly due to the impossibility of measuring the statistical significance of the results unless a stochastic DEA is performed, and to the fact that the scores obtained cannot be used in second stage regressions unless the separability of the regressors be empirically ascertained. For these reasons, when possible, SFA has often been used to obtain more robust results, also in the evaluation of the municipal waste management sector. However, the research conducted so far has answered different questions from those in this study.

For example, Fan et al. (2020) conducted an assessment of the efficiency of 30 Chinese provinces in the management of urban waste in the period 2008–2017, also trying to determine which could be the cost drivers of the service. Vishwakarma et al. (2012) applied SFA to evaluate the efficiency of 22 cities in the state of Madhya Pradesh, India, using data from the year 2005. De Groot et al. (2011) used SFA to evaluate the effect of the policy choices on the efficiency in the provision of waste services using data for 1238 Dutch municipalities from 2005 to 2008. They showed that contracting out has a positive effect on costs, however the form of management, public or private, does not affect the efficiency.

A few papers used SFA to investigate the usefulness of inter-municipal cooperation in the field of waste services. They deal with a pre-requisite of our research question, which relates, instead, to the optimal extent of such a cooperation. For example, Fusco and Allegrini (2020) investigated the effect of spatial interdependence in 4250 Italian municipalities for the year 2015 through a Cobb-Douglas cost frontier derived by applying a Spatial SFA model. Since density economies and spatial dependence were observed, they concluded, among other things, that inter-municipal cooperation can be considered a valid practice. Finally, with objectives similar to ours, but investigating eco-efficiency instead of cost efficiency, Molinos-Senante and Maziotis (2021) applied a SFA translog cost frontier model to data relative to 298 Chilean municipalities in the year 2018. Their results suggest that smaller municipalities are less eco-efficient than those of medium and large size, so that increasing the area in which the service is provided could improve performance.

3. Methods

In order to identify the optimal size (in terms of population served) of the territorial units for which the provision of public services is entrusted to external operators, a two-step analysis has been performed. In the first step, a production frontier has been estimated using SFA with the aim of calculating efficiency scores for Italian municipalities. In the second step, the relation between the efficiency scores and the population served has been analysed.

SFA is a parametric stochastic method designed for the approximation of both production and cost frontiers. Basic SFA models were introduced simultaneously but independently by Aigner et al. (1977) and Meeusen and van Den Broeck (1977). The work of Aigner et al. (1977) is rooted in the first studies on the estimation of parametric frontier production functions by Aigner and Chu (1968), Afriat (1972) and Richmond (1974). It presents the canonical model formulation (1) that serves as the basis for other variations. Let y be the n observations on variable y, the column vector $\boldsymbol{x}_i$ be the n observations on variable x_i with $i = 1,\ldots,p$, and let β_i be a vector of the parameters corresponding to x_i, while $\boldsymbol{\varepsilon}$ is the column vector containing the n error terms. The base model looks as follows:

$$\boldsymbol{y} = f(\boldsymbol{x}_i;\beta_i) + \boldsymbol{\varepsilon}. \tag{1}$$

The vector $\boldsymbol{\varepsilon}$ can be decomposed as follows:

$$\begin{aligned} \varepsilon &= v + u, \\ v &\sim_{\text{iid}} \mathcal{N}\left(0,\sigma_v^2\right), \\ u &\sim_{\text{iid}} \mathcal{N}^+\left(0,\sigma_u^2\right), \end{aligned} \tag{2}$$

where v is a normally distributed disturbance which represents the measurement and specification error and u is a positive disturbance which represents the technical inefficiency which follows a truncated normal distribution at zero and guarantees inefficiency to be positive only ($\mathcal{N}^+$). Both the components of the error term are assumed to be statistically independent from each other and from x_i. If $u > 0$, then the production unit is to some extent inefficient. Otherwise, if $u = 0$, then the production unit is fully efficient.

The maximum likelihood estimation method was used to estimate the parameters of the stochastic frontier model. Using the composed error terms of the stochastic frontier model (1), the total variation in output from the frontier level of output, attributed to technical efficiency, is defined by

$$\gamma = \frac{\sigma_u^2}{\left(\sigma_u^2 + \sigma_v^2\right)}. \tag{3}$$

The SFA approach requires an assumption about the functional form of the production function. The choice of functional form brings a series of implications with respect to the shape of the implied isoquants and the values of the elasticities of factor demand and factor substitution (Fried et al., 2008). The most widely used forms of a production function are the first-degree flexible Cobb–Douglas (Cobb and Douglas, 1928) and the second-degree flexible transcendental logarithmic production functions, called for short "translog production functions" (Christensen et al., 1971, 1973; Griliches and Ringstad, 1971).

This study adopts the first-degree flexible Cobb–Douglas production function, which, in its earliest form, modelled output (y) as a function of (two inputs) physical capital (k) and labour (l) as $f(\boldsymbol{k};\boldsymbol{l};\beta_i) = \beta_0 k^{\beta_1} l^{\beta_2}$, where the homogeneous parameters ($\beta_0, \beta_1, \beta_2$) are interpreted as the technology (β_0) and the elasticities of output with respect to the inputs (β_1 and β_2), respectively. In order to describe production processes involving more than two inputs, it can be rewritten as follows:

$$f(\boldsymbol{x}_i;\beta_i) = \beta_0 x_1^{\beta_1} \ldots x_p^{\beta_p} = \beta_0 \prod_{i=1}^{p} x_i^{\beta_i}, \tag{4}$$

where $f(\boldsymbol{x}_i;\beta_i)$ is the output of a production process that requires p inputs (which are denoted by x_i; $i = 1,\ldots,p$). Such a production function has universally smooth and convex isoquants.

With the aim of reducing the computational complexity, the function (4) can be converted into an equation that is linear in the logarithms of the associated variables,

$$\log\left(\beta_0 \prod_{i=1}^{p} x_i^{\beta_i}\right) = \log\beta_0 + \sum_{i=1}^{p} \beta_i \log x_i. \tag{5}$$

Therefore, under the assumption that $f(\boldsymbol{x}_i;\beta_i)$ is of Cobb-Douglas type, the stochastic frontier model in equation (1) can be written in logs as

$$\log\boldsymbol{y} = \log\beta_0 + \sum\nolimits_{i=1}^{p} \beta_i \log\boldsymbol{x}_i + \boldsymbol{\varepsilon}. \tag{6}$$

A first empirical model has been estimated using formula (6) under the following specifications: $\boldsymbol{y}$ is a vector containing n observations of the single output of the production process, $\boldsymbol{x}_i$ is a vector containing n observations of the i-th input of the production process, with $i = 1,\ldots,p$, and the error term is specified as in (2).

Since some of the variables considered as inputs in the model (6) could also be intended as exogenous variables that might affect the performance of the observed units, a second empirical model has been estimated using the following specifications: $\boldsymbol{y}$ is a vector containing n observations of the single output of the production process, $\boldsymbol{x}_i$ is a vector containing n observations of the single input of the production process (therefore, with $p = 1$), and each single element of the vector $\boldsymbol{\varepsilon}$ can be written as

$$\begin{aligned} \varepsilon &= v + u, \\ v &\sim_{\text{iid}} \mathcal{N}\left(0,\sigma_v^2\right), \\ u &\sim_{\text{iid}} \mathcal{N}^+\left(\mu,\sigma_u^2\right), \end{aligned} \tag{7}$$

where μ is a function of a set of variables affecting the performance of the municipalities. In more detail, the second empirical model parametrizes the pre-truncation mean of the truncated normal distribution of u as follows,

$$\mu = \delta_0 + \delta_h \boldsymbol{z}_h^{'}, \tag{8}$$

where the column vector $z_h^{'}$ contains n observation of the h-th exogenous variable which is supposed to affect the performance of the municipalities, with $h = 1,\ldots,v$, δ_0 is the intercept, and δ_h is the parameter corresponding to the h-th exogenous variable.

Once the frontier has been estimated, the efficiency score for each municipality is calculated as follows:

$$eff = \exp(-u) \tag{9}$$

so that *eff* ranges between 0 (minimum efficiency) and 1 (maximum efficiency).

In the second step of the analysis, the relation between the efficiency scores calculated at the end of the first step and the variable expressing the size of the supra-municipal areas has been investigated using two of the most common interpolation techniques: polynomial regression and spline regression. The idea is to find a value of the size variable beyond which the efficiency score steadily decreases. Cialani and Mortazavi (2020) show that the collection and recycling of waste has the highest cost advantage when the quantity of waste to be recycled increases, but does not exceed a certain value (500 billion kg). Niaounakis and Blank (2017) find that scale effects are pronounced, especially for municipalities or inter-municipal cooperations with less than roughly 30,000 properties (roughly 60,000 inhabitants). Therefore, with the aim of identifying a threshold value of the size, after finding the best fitting function for the original data points, its maxima have been studied.

At first, a second-degree polynomial regression model has been estimated, in which the efficiency scores just calculated (*eff*) have been regressed on the dimensional variable p. Let $\boldsymbol{eff}$ be the n observations of the variable *eff*, $\boldsymbol{p}$ be n observations of the variable p, $\beta_0, \beta_1, \beta_2$ be the regression parameters, and $\boldsymbol{\varphi}$ be the column vector containing the n error terms. The second-order polynomial model in one variable is given by

$$\boldsymbol{eff} = \lambda_0 + \lambda_1\boldsymbol{p} + \lambda_2\boldsymbol{p}^2 + \boldsymbol{\varphi}. \tag{10}$$

If the estimated quadratic function fits well the original data points, then the threshold value of the dimensional variable corresponds to the x-coordinate of the vertex of the function (10). Therefore, it is calculated using the formula $-\lambda_1/2\lambda_2$.[3]

If the estimated quadratic function is a poor fit for the original data points, then a better fitting function is found by adopting a spline regression (Hastie and Tibshirani, 1986). This is a non-parametric regression technique. It divides the dataset into bins at intervals or points called *knots* and each bin has its separate fit. In other words, instead of fitting all data points with a single polynomial curve, with a spline regression one attempts to fit segments to different parts of the data, with breakpoints (knots) at pre-determined places, requiring that the segments have to be connected.

Let the segment $[a, b]$ be divided into partial segments by the points $c_1, c_2, \ldots, c_k$, so that $a = c_1 < c_2 < \ldots < c_k = b$.

Assuming that cubic polynomials will be used to interpolate the data, a spline regression model is constructed starting from a piecewise function of the form

$$S(c) = \begin{cases} s_1(c) & \text{if} \quad c_1 \le c \le c_2, \\ s_2(c) & \text{if} \quad c_2 \le c \le c_3, \\ \quad \vdots & \\ s_{k-1}(c) & \text{if} \quad c_{k-1} \le c \le c_k, \end{cases} \tag{11}$$

where s_i is a third-degree polynomial defined by

$$s_i(c) = \psi_1 + \psi_2(c - c_k) + \psi_3(c - c_k)^2 + \psi_4(c - c_k)^3. \tag{12}$$

The above model will be a smooth curve within the intervals bounded by the knots, but the "connection" between the segments will not be smooth. To force smoothness over the entire fitted curve, two additional constraints must be added to the continuity of the function: the continuity of the first derivative, and continuity of the second derivative. This will guarantee that the fitted curve is differentiable, with no sharp changes in the direction. This is called a cubic spline. We can have a basis of polynomials of any degree, but cubic splines are flexible enough for most circumstances.

Once a cubic spline function has been estimated, in order to find any maximum points, the trend of that function is studied by examining its derivatives.

4. Data, variables, and empirical models

The analyses have been performed using data on a sample of 6916 Italian municipalities. At the end of the 2019, Italy was divided into 7914 municipalities. After outlier detection and, in general, data cleaning, 998 municipalities have been removed (see Supplementary material, to gain further insights into data cleaning). Therefore, the sample used in this study represents 87.39% of the entire population.

A two-stage analysis has been performed. Sources of information and variables used in the analyses are displayed in Table 1.

In the first stage a production frontier has been estimated. As already said, it has been formulated following the model (6) and the specification (2) for the error term. Specifically, is has called SFAnz and has been estimated using one output and six inputs and assuming that both the elements which compose the error component have zero mean and sigma square variance. The first production frontier is specified as

$$\begin{aligned} \log(\boldsymbol{pc\_waste}) &= \beta_0 + \beta_1 \log(\boldsymbol{pc\_currexp}) + \beta_2 \log(\boldsymbol{pt\_inc}) \\ &+ \beta_3 \log(\boldsymbol{mun\_height}) + \beta_4 \log(\boldsymbol{cap\_af}) + \beta_5 \log(\boldsymbol{dis\_pop}) \\ &+ \beta_6 \log(\%\boldsymbol{sorwaste}) + \boldsymbol{v} - \boldsymbol{u}, \end{aligned} \tag{13}$$

with $v \sim_{\text{iid}} \mathcal{N}(0, \sigma_v^2)$ and $u \sim_{\text{iid}} \mathcal{N}^+(0, \sigma_u^2)$.

The output of the production frontier is the quantity of waste collected in 2019 per person, measured in tons of waste collected per person per year (data retrieved from the Italian Institute for Environmental Protection and Research, ISPRA), called *pc_waste*. The inputs of the production frontier are: a) current expenditure for waste collection services entered in the municipal budget of the year 2019 per person, measured in Euro (data extracted from a financial database of Italian local authorities compiled by Bureau Von Dick, called AIDA PA), called *pc_currexp*; b) the percentage of the waste from separate collection in the total waste collected in 2019 (source: ISPRA), called *%sorwaste*; c) per taxpayer income, calculated using data from tax declarations submitted by taxpayers in 2020 (such data are made available by the Italian Ministry of Economics), called *pt_inc*; d) the number of beds available in tourist accommodation facilities located in the municipality in 2019 (data collected form the Italian National Institute of Statistics, ISTAT), called *cap_af*; e) the elevation of the municipality above sea level, calculated as the average between the lowest and the highest points of the municipality (data gathered from ISTAT), called *mun_height*; and f) an index of the dispersion of the population over the municipal territory, calculated as the ratio of the number of residents living in scattered housing clusters and the number of residents living within the administrative boundaries of a municipality in 2011 (data retrieved form the 2010–2011 census made by ISTAT), called *dis_pop*.

The municipality-specific technical efficiency (eff^{nz}) of the single municipality has been estimated using the expectation of u conditional on the random variable ε. It follows that

$$eff^{nz} = \exp(-u) \tag{14}$$

so that $0 \le eff^{nz} \le 1$.

In the second stage a regression analysis has been carried out. The efficiency scores estimated with the first empirical model (eff^{nz}) have been regressed on the size variable using the empirical model

$$\boldsymbol{eff}^{nz} = \beta_0 + \beta_1 \boldsymbol{mun\_pop} + \beta_2 \boldsymbol{mun\_pop}^2 + \boldsymbol{\delta}. \tag{15}$$

The size variable, called *mun_pop*, is the number of inhabitants of the municipality in 2019 (data on population are taken from ISTAT). As already explained in Section 3, the idea of fitting a polynomial is supported by some empirical studies that suggest an initially increasing and then decreasing trend of this function (Cialani and Mortazavi, 2020; Niaounakis and Blank, 2017).

The production frontier has also been formulated following the model (6) and the specification (7) for the error term. Given that this second attempt has led to statistically non-significant efficiency scores, its empirical models and results are exhibited in Appendix B.

Table 2 presents descriptive statistics for all the variables used in this study.

The computational implementation has been performed in the R programming language (Team R Core, 2015) and based on the "frontier" package (Coelli and Henningsen, 2013).

5. Results

The stochastic production frontier *SFAnz* – which has been estimated using the empirical model (13) – shows statistically significant results. The maximum likelihood estimates of the *SFAnz* frontier are displayed in Table 3. The estimate of the gamma parameter is highly significant, and it takes a value of 0.69. This can be interpreted as saying that 69% of the variation in output among the municipalities is due to the differences in technical efficiency; the remaining part is due to disturbance. The estimate of the σ^2 parameter (0.1087) is different from 0, indicating a good fit and correctness of the assumption of half-normal distributions.

All the coefficients in the model are extremely significant. Specifically, each one is significant at 0.1%. In line with expectations, the coefficients for *pc_currexp*, *pt_inc*, and *cap_af* are positive, meaning that an increase in per capita current expenditure, the average income per taxpayer and the number of beds in accommodation facilities should induce an increase in per capita waste collected, and vice versa. For

[3] The formula to find the coordinates of the vertex of a parabola is $(x; y) = (-\lambda_1/2\lambda_2; \lambda_0 - (\lambda_1^2/4\lambda_2))$.

Table 1
Sources, definitions and measurement of variables.

Name	Abbreviation	Description	Year	Source	Unit of measure	Model (*) – Type of variable (**)	Expectation
Per-capita waste generation	pc_waste	Amount of municipal waste generated per person	2019	Own elaboration from ISPRA and ISTAT data	Tons/year	SFAnz and SFAwz– y	
Current expenditure on waste management	pc_currexp	Current expenditure on the entire cycle of waste management (from collection to treatment/disposal) per person	2019	AIDA PA	Euros/year	SFAnz and SFAwz – x	+
Pro taxpayer income	pt_inc	Average income earned per taxpayer	2019	Own elaboration from MEF data	Euros/year	SFAnz – x SFAwz – z	+
Population dispersion index	dis_pop	Ratio of number of persons living in scattered housing to the total number of residents	2011	Own elaboration from ISTAT data	%	SFA_nz – x SFA_wz – z	–
Capacity of accommodation facilities	cap_af	Number of beds in accommodation facilities	2019	ISTAT	Number	SFA_nz – x SFA_wz – z	+
Height	mun_height	Municipality height above the sea level	2015	Own elaboration from ISTAT data	Metres	SFA_nz – x SFA_wz – z	–
Percentage of sorted waste	%sorwaste	Ratio between the amount of waste from separate collection and the total amount of waste collected	2019	ISPRA	%	SFA_nz – x SFA_wz – z	–
Efficiency score	eff	Efficiency score returned by SFA	2019	Own elaboration		Regression – dependent v.	
Population	mun_pop	Count of persons who have their usual residence in the territory of the municipality	2019	ISTAT	Number	Regression – independent v.	+/– (***)

Note: * SFAnz = SFA model where no exogenous variables that might affect the performance of the municipalities have been considered; SFAwz = SFA model including exogenous variables. ** y = output of the production process; x = input of the production process; z = exogenous variable. *** The efficiency score is expected to increase as the population increases to a certain value, after which it is expected to decrease.

Table 2
Descriptive statistics of the variables used in the SFA model and in the linear regression model (see Table 1 for a detailed description of the variables).

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	St.Dev.
pc_waste	0.124	0.347	0.426	0.455	0.511	2.969	0.188
pc_currexp	30.066	99.562	131.173	149.992	174.595	943.270	4223.885
%sorwaste	0.000	0.532	0.680	0.630	0.764	0.975	0.124
pt_inc	6419.050	15385.603	18709.060	18692.085	21689.428	49556.460	3251.334
dis_pop	0.000	0.019	0.061	0.106	0.144	0.974	30.095
cap_af	0.000	14.000	59.000	650.345	253.000	97477.000	421.373
mun_pop	0.032	1.019	2.455	7.412	6.203	1395.980	85.084
mun_height	0.360	145.457	316.055	448.738	621.365	2590.760	0.183

Table 3
Maximum Likelihood Estimates for the SFAnz frontier (see Table 1 for a detailed description of the variables).

	Coefficient: Symbol	Coefficient: Value	Standard Error
Intercept	β_0	−8.8119***	0.1445
Log(pc_currexp)	β_1	0.2533***	0.0074
Log(pt_inc)	β_2	0.6856***	0.0139
Log(mun_height)	β_3	−0.0308***	0.0026
Log(cap_af)	β_4	0.0291***	0.0014
Log(dis_pop)	β_5	−0.1997***	0.0291
Log(%sorwaste)	β_6	−0.3091***	0.0257
σ^2		0.1087***	0.0034
γ		0.6915***	0.0202
Log likelihood value		*−77.1026*	
No. of observations		*6916*	
Mean efficiency		*0.8155*	

Note: the symbol "***" indicates that the probability that $Z \leq z$ is less than 0.001. Therefore, it denotes a very high level of significance.

example, the coefficient for the per taxpayer income of 0.68 indicates that per capita waste collected is elastic to changes in the per taxpayer income. This is not surprising, since when income increases, consumption expenditure also increases (even if by a smaller amount), and accordingly the production of waste.

On the contrary, the coefficients for *mun_height*, *dis_pop*, and *%sorwaste* and the intercept are negative. This is again not surprising. For example, regarding *%sorwaste*, a negative coefficient indicates that an increase in the percentage of sorted waste should induce a decrease in per capita waste collected, and vice versa. This point can be explained by the fact that separate collection activities (which allows to increase the percentage of sorted waste) lead to virtuous behaviours that lead to a reduction of the quantity of waste produced. Reusing of certain parts of composed objects that would be trashed, if the composed object were thrown away as unsorted waste, can be counted among them. Let us think of an expired jar of beans. In case of undifferentiated waste collection, it would be thrown away in full and, consequently, its total weight would be included in the total amount of waste collected. In the case of separate waste collection, it could occur that only the contents are thrown away, while the jar could be washed and reused for other purposes. Therefore, in this case, just the weight of the contents would contribute to the total amount of waste collected.

The efficiency scores eff^{nz} have been calculated using (14). The mean of the scores gained by the Italian municipalities is very high at 0.81. Therefore, in general, the Italian municipalities show a satisfactory performance in the organization of waste management services in 2019. In more detail, no municipality obtains a very low efficiency score (0.01–0.20). Just 10 municipalities get a low efficiency score (0.21–0.40). The number of municipalities which obtain a moderate efficiency score (0.41–0.60) is quite low (168) in comparison to the sample size. Among the remaining municipalities, 2160 achieve a high efficiency score (0.61–0.80) and 4578 a very high efficiency score (0.81–1.00).

The relation between eff^{nz} and the size of the supra-municipal area (population) has been studied by estimating a second-order polynomial regression model (15). The results of the regression analysis are displayed in Table 4.

The adjusted R-squared is rather low at 0.06. The main reason for this result is that the performance of a municipality in the organization of waste management services cannot be totally explained by its size. Anyway, looking at the estimated parameters, it can be immediately noted that the coefficient of the squared term – which is extremely significant from a statistical standpoint – is negative, meaning that the parabola opens downwards. Therefore, the efficiency scores appear to increase as the size of the municipality increases only up to a certain value, beyond which they decrease. The resulting parabolic curve is represented in Fig. 1.

The estimated function takes its highest value when the size of the municipality is 62,251 inhabitants. This value can be interpreted as the greatest size of the territorial unit in which the service should be entrusted to an external operator.

The results of the analysis performed estimating the production frontier formulated following the model (6) and the specification (7) for the error term and the spline regression for the investigation of the relation between the efficiencies and the size of the municipality – which are described in detail in Appendix B because this second attempt has led to statistically non-significant efficiency scores – are not all that different from those just shown. In more detail, the threshold value of the population beyond which the efficiency scores stably decrease is around 90,000.

6. Discussion and policy implications

The most remarkable result to emerge from the data is that the optimal population level that should be taken into consideration when defining the boundaries of the territorial units for the organization of waste collection services is about 62,000 inhabitants. It is emphasized that this result is only slightly higher than prior findings. A possible cause of the discrepancy is due to the fact that, unlike other research carried out in this area – which evaluated jointly the environmental and cost performance of municipalities in the organization of waste management services (Sarra et al., 2020a, 2020b) – in this study only the economic efficiency is investigated, given that most of the municipalities achieve the target sets by the Italian legislation in terms of thresholds of separate collections to be reached for the year 2019.

At first sight, this result is apparently at odds with the literature confirming the existence of technical economies of scale in the provision of waste services in such a way that the optimal operational size of service providers is around 300,000 inhabitants served (Simões et al., 2010) or more (Carvalho and Marques, 2014), but this is not so. Actually, the main implication of the abovementioned finding is not that

Table 4
Results of the best fitting polynomial regression model to test for the relationship between efficiency scores (estimated using the SFAnz model) and the municipality's size (see Table 1 for a detailed description of the variables).

	Coefficient:Symbol	Coefficient: Value	Standard Error
Intercept	β_0	0.8032***	0.0012
Log(mun_pop)	β_1	0.0204***	0.0010
Log(mun_pop)^2	β_2	−0.0025***	0.0003
Multiple R^2		*0.0664*	
Adjusted R^2		*0.0661*	

Note: the symbols "***" indicates that Pr $(>|t|) < 0.001$, implying a high level of significance.

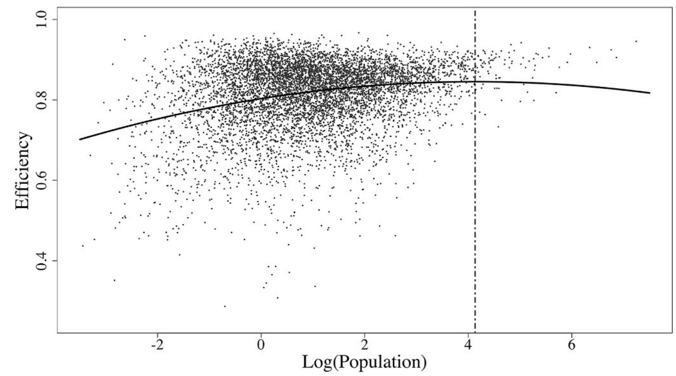

Fig. 1. Threshold value of the size when efficiency scores are estimated using the SFAnz model.

each service provider should operate at a scale corresponding to 62,000 inhabitants served, but that the efficiency of the authority that entrusts the management of the service to an external operator seems to be maximized when it refers to supra-municipal areas – or sub-municipal areas, when the municipality's population is higher than that value, – of about 62,000 inhabitants. As aforementioned, for service providers this does not imply an obstacle to reaching their optimal operational size. In fact, they can achieve the scale of production which allows them to exploit technical economies of scale by getting entrusted for the service by more than one municipality.

With a view to the greater use of tender procedures (sought by the European Union), one of the most important aspects which arises when the abovementioned finding is placed in the context of the Italian reality is the noteworthy number of tenders that should be launched throughout the national territory, and, in particular, the number of those that should be managed by each single contracting authority. In fact, at the time of writing, the organization of the municipal waste management service in Italy has a total of 65 Optimal Territorial Areas (OTAs) – which are territorial partitions composed of several adjacent municipalities within which there is an authority (which can be called the OTA authority) which is in charge of managing the process of outsourcing and monitoring the results. In other words, in Italy the law states that, when the waste sector reform is fully implemented, the contracting authority is the OTA authority (no longer the individual municipality). Given that the population of Italy is around 60 million, each OTA corresponds to about 923,000 inhabitants on average. Therefore, a single OTA's authority (the contracting authority) will have to launch around 15 tender procedures, on average.

If a single entity manages several tender procedures considerable benefits can be obtained. In the first place, this allows avoiding a duplication of fixed transaction costs (so reducing their incidence on the tariffs for the service) and developing specialized competences in the managing of the tenders.

Additionally, given that tenders will refer to territorial units that will be roughly the same size, the intensity of mimicking effects should arise both on the part of the external operators and on that of the municipalities, and this should lead to an alignment in both the performance of the external operators and of the tariffs on waste set by neighboring municipalities. Moreover, managing a great number of tender procedures helps the contracting authority to fight bid rigging or collusive tendering, i.e. those circumstances in which businesses, that would otherwise be expected to compete, secretly conspire to raise prices or lower the quality of goods or services for purchasers who wish to acquire products or services through a bidding process (OECD, 2009). It is well known that bid rigging may not be evident from the results of a single tender. Often a collusive scheme is only revealed when one examines the results from a number of tender procedures over a period of time. In order to tackle bid rigging effectively, an authority should be able to collect historical information on bidding behavior, to monitor bidding activities, and to perform analyses of the data about the bids. In this way, managing several tender procedures helps the contracting authorities (and competition authorities) to compare and monitor the performance of the service providers in competition among them reducing information asymmetries and helping to identify troubling situations.

From a sustainability perspective, if the optimal size for tendering procedures is reached, a general improvement in economic efficiency is expected for the contracting authority and, consequently, monetary resources are freed up that can be invested in improving the qualitative-environmental performance of the whole waste cycle. This can help to overcome the shrinkages of resources which hinder the growth of investments needed for a better implementation of the principles of circularity (technological platforms, recycling, and so on). In this regard, it should be noted that a mismatch between the optimal size of territorial units in the outsourcing of waste service and the optimal scale-size of treatment facilities from the perspective of both economic and environmental efficiency (Koley, 2023; Morelli et al., 2020) can occur. How to overcome this problem remains an open research question that requires further investigations.

7. Conclusions

In this study the problem of the optimal size of the territorial units where the provision of public services is entrusted to external operators has been addressed. An empirical analysis aimed at investigating the effect of the size of the municipality on its performance in organizing waste collection services has been presented. The evidence from this analysis suggests that the efficiency scores increase as the size of the municipality increases only up to approximately 62,000 inhabitants. As the municipality's population grows beyond this value, the efficiency scores decrease steadily. This finding would seem to imply that the optimal population level that should be taken into consideration by the regulatory authorities in defining the boundaries of the territorial partitions for entrusting waste service to an external operator is about 62,000 inhabitants.

This result is somewhat in line with those of the few earlier studies on this topic. A strong point of this work is the considerable reliability of the method applied in the empirical analysis – it is a parametric technique which allows estimating a production frontier for the measurement of efficiency scores, unlike the others which are based on non-parametric techniques – which makes the results gathered from this study trustworthy. Another strong point of this work is that the empirical analysis was performed on a very large sample, which covers a wide range of sizes.

The approach used here would lend itself well for use by the regulatory authorities struggling to find the optimal size of the territorial units where the provision of waste service or any other public service will be entrusted to external operators. From the perspective of the greater use of tender procedures, in the specific case of waste services in Italy, the results of the analysis imply that each contracting authority should manage 15 tenders on average. Among the main benefits deriving from it, the focus was placed on a) an intensification of mimicking effects, which should lead to an alignment in both the performance of the external operators and in the tariffs on waste set by the neighboring municipalities, and b) a greater ease for the contracting authority in fighting bid rigging or collusive tendering.

One weakness of this study is that the empirical application is limited to data for one year. Research into solving this problem is already in progress. Data are being collected to extend the dataset to subsequent years. On this point, it has to be added that public bodies have recently started to make available data which was hard to find until now, among which are data that reveal the specific characteristics of the agreements that regulatory authorities conclude with external operators. The use of this type of data allows extending the forthcoming analyses on the current topic to the consideration of the behaviour and the performance of external operators.

Another important problem to resolve in future studies is the treatment of missing data. Relatively few studies which propose methods for the imputation of missing data have been published. They will be the starting point for redesigning an analysis where missing data should not be eliminated.

Finally, future work will also try to investigate if there is some heterogeneity in the input elasticities of the estimated production frontier. The feeling is that input elasticities may vary from one region to another, giving rise to frontiers that are also very different from one another.

Credit author statement

Roberto Benedetti, Marialisa Mazzocchitti, and Alessandro Sarra contributed equally to this work.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A

The production frontier has been formulated following the model (6) and the specification (7) for the error term. More precisely, it has been estimated using one output and one input, and assuming that all the other variables considered as inputs in the empirical model (13), in addition to the *mun_pop*, affect the distribution of the element of the error component which represents the technical inefficiency. The second empirical model estimated (hereinafter, SFAwz) can be written as follows:

$$\log(\boldsymbol{pc\_waste}) = \beta_0 + \beta_1 \log(\boldsymbol{pc\_currexp}) + \boldsymbol{v} - \boldsymbol{u} \tag{A.1}$$

where $v \sim_{\text{iid}} \mathcal{N}(0, \sigma_v^2)$ and $u \sim_{\text{iid}} \mathcal{N}^+(\mu_i, \sigma_u^2)$, with

$$\mu_i = \delta_0 + \delta_1 \log(\boldsymbol{pt\_inc}) + \delta_2 \log(\boldsymbol{mun\_height}) + \delta_3 \log(\boldsymbol{cap\_af}) + \delta_4 \log(\boldsymbol{dis\_pop}) + \delta_5 \log(\%\boldsymbol{sorwaste}) + \delta_6 \log(\boldsymbol{mun\_pop}) + \delta_7 \log(\boldsymbol{mun\_pop})^2 + \boldsymbol{W} \tag{A.2}$$

where $\delta_0, \ldots \delta_7$ are parameters to be estimated and W is a random error.

After the production frontier has been estimated, the efficiency score for each municipality (eff^{wz}) has been calculated, as follows:

$$eff^{wz} = \exp(-u) \quad (A.3)$$

so that $0 \le eff^{wz} \le 1$.

Finally, the efficiency scores estimated with the second empirical model (eff^{wz}) have been regressed on the dimensional variable (*mun_pop*) using both a second-order polynomial regression model (10) and a spline regression model (11).

As regards the stochastic production frontier, called SFAwz, the maximum likelihood estimates (displayed in Table A1) are found to be not statistically significant. Unfortunately, the estimate of the gamma parameter – which is not statistically significant – is close to 0, indicating that there is (almost) no inefficiency. Basically, there is a lack of convergence. Thus, there are problems which negatively affect the validity and reliability of the statistical tests. A model misspecification is considered to be the first and main problem behind it all.

Table A1
Maximum Likelihood Estimates for the SFAwz frontier (see Table 1 for a detailed description of the variables)

	Symbol	Coefficient	Standard Error	Statistical significance
Intercept	β_0	−2.5001	0.0523	***
Log(pc_currexp)	β_1	0.2471	0.0072	***
Intercept	δ_0	−6.7336	0.1593	***
Log(pt_inc)	δ_1	0.7643	0.0160	***
Log(mun_height)	δ_2	−0.0521	0.0028	***
Log(cap_af)	δ_3	0.0493	0.0018	***
Log(dis_pop)	δ_4	−0.3136	0.0306	***
Log(%sorwaste)	δ_5	−0.3020	0.0270	***
Log(pop)	δ_6	−0.0749	0.0039	***
Log(pop)^2	δ_7	0.0034	0.0011	**
σ^2		0.0569	0.0010	***
γ		0.0640	0.0657	
Log likelihood value		*104.1337*		
No. of observations		*6916*		
Mean efficiency		*0.6578*		

Note: the symbols "***" and "**" indicate different levels of statistical significance. Specifically, each symbol corresponds to a range of values for the probability that $Z \le z$, which are: Pr $(Z \le z) < 0.001$, and $0.001 <$ Pr $(Z \le z) < 0.01$, respectively.

The coefficient of the single input of the production frontier is significant at 0.1% and takes a positive value, as expected. More specifically, the coefficient for *pc_currexp* of 0.24 indicates that per capita waste collected is elastic to changes in per capita current expenditure.

The coefficients of the variables explaining the mean of the inefficiency component of the error term are all statistically significant at 0.1% with the only exception being the squared *mun_pop*, which is significant at 1%. Their signs are in accordance with expectations.

Despite the poor reliability of the maximum likelihood estimates for the SFAwz model, the efficiency scores eff^{wz} have been calculated using formula (18). The mean of the scores obtained by the Italian municipalities is high, reaching 0.65. In more detail, no municipality obtains a very low efficiency score (0.01–0.20). Just 48 municipalities obtain a low efficiency score (0.21–0.40). The number of municipalities which obtain a moderate efficiency score (0.41–0.60) is 2589. Among the remaining municipalities, 3125 obtain a high efficiency score (0.61–0.80) and 1154 a very high efficiency score (0.81–1.00).

The relation between eff^{wz} and the size of the supra-municipal area has been studied by estimating both a polynomial regression model and a spline regression model. The results of the regression analysis are displayed in Table A2.

Table A2
Results of both the best fitting polynomial regression model and the spline regression model to test for the relation between efficiency scores estimated using the SFAwz and the size of the municipality (see Table 1 for a detailed description of the variables)

	Coefficient	Standard Error	t value	Statistical significance
Polynomial regression				
Intercept	0.6794	0.0019	354.637	***
Log(mun_pop)	−0.0064	0.0016	−3.945	***
Log(mun_pop)^2	−0.0057	0.0005	−10.268	***
Multiple R^2	*0.0512*			
Adjusted R^2	*0.0510*			
Spline regression				
Intercept	0.6485	0.0386	16.802	***
bs(Log(mun_pop), 5)1	−0.0454	0.0551	−0.823	
bs(Log(mun_pop), 5)2	0.0928	0.0356	2.607	**
bs(Log(mun_pop), 5)3	−0.0765	0.0440	−1.737	.
bs(Log(mun_pop), 5)4	−0.0369	0.0433	−0.853	
bs(Log(mun_pop), 5)5	−0.3158	0.0781	−4.040	***
Multiple R^2	*0.0567*			
Adjusted R^2	*0.0560*			

Note: the symbols "***", "**", "*", and "." indicate different levels of statistical significance. Specifically, each symbol corresponds to a range of values for the probability that Pr $(>|t|)$, which are: Pr $(>|t|) < 0.001$, $0.001 <$ Pr $(>|t|) < 0.01$, $0.01 <$ Pr $(>|t|) < 0.05$, and $0.05 <$ Pr $(>|t|) < 0.1$ respectively.

The adjusted R-squared is somewhat low at 0.05 for both regression analyses. As regards the polynomial regression model, it can be immediately noted that the coefficient of the squared term is negative, meaning that the parabola opens downwards. Therefore, in this case again the efficiency scores appear to increase as the size of the municipality increases, but only up to a certain value beyond which they decrease. The resulting parabolic curve is represented in Figure A1 panel a. Looking at the parabola it can be noted that it takes its highest value when the municipality is very small, 568 inhabitants, to be precise. This result is difficult to explain. In other words, it is difficult to justify in economic terms why having 568 inhabitants corresponds to the greatest size of the geographical area in which the service should be awarded to a single operator.

This is the main reason why the relation between the efficiency scores eff^{wz} and the size of the supra-municipal areas has been studied estimating a spline regression model too. The estimated spline function (which has been represented in Figure A1 panel b) has a global (absolute) maximum for a population of 980, which is, again, too low to be the greatest size of a geographical area in which the service should be provided by a single operator. Nevertheless, after that point, the function doesn't decrease stably. The first derivative of the function decreases until it takes a value equal to 0.04. Then, there is a change in the sign of the second derivative, and consequently of the concavity of the function, indicating an inflection point (it can be found at 18,356). After that point, the spline function steadily decreases, and it takes values less than 0.04 from the point at which population is equal to 91,835. This value can be interpreted as the greatest size of the geographical area in which the service should be provided by a single operator.

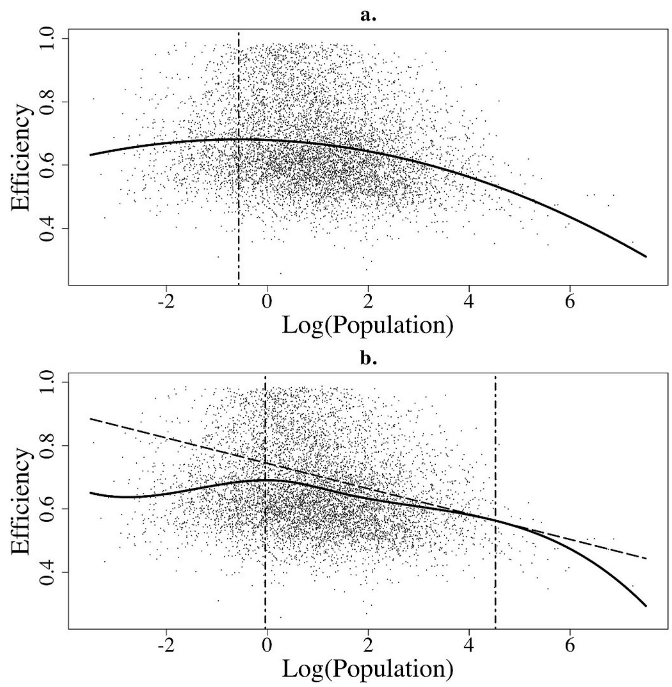

Fig. A1. Threshold values of the size when efficiency scores are estimated using the SFA model with zeds: a. polynomial regression; b. spline regression.
Note: solid lines are the estimated functions; the dash-dot lines indicate the maximum points and other points of interest; the dashed line illustrates the tangent drawn at the point beyond which the first derivative is less than −0.04.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2023.119141.

References

Abrate, G., Erbetta, F., Fraquelli, G., Vannoni, D., 2014. The costs of disposal and recycling: an application to Italian municipal solid waste services. Reg. Stud. 48, 896–909. https://doi.org/10.1080/00343404.2012.689425.

Afriat, S.N., 1972. Efficiency estimation of production functions. Int. Econ. Rev. 13, 324–341.

Aigner, D.J., Chu, S.F., 1968. On estimating the industry production function. Am. Econ. Rev. 58, 826–839.

Aigner, D.J., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. J. Econom. 6, 21–37. https://doi.org/10.1016/0304-4076(77)90052-5.

Albalate, D., Bel, G., Rosell, J., 2023. An evaluation of optimal scale and jurisdiction size to improve efficiency in metropolitan bus systems. Res. Transport. Bus. Manag. 46 https://doi.org/10.1016/j.rtbm.2022.100822.

Bel, G., Costas, A., 2006. Do public sector reforms get rusty? Local privatization in Spain. J. Pol. Reform 9, 1–24. https://doi.org/10.1080/13841280500513084.

Bel, G., Fageda, X., 2010. Empirical analysis of solid management waste costs: some evidence from Galicia, Spain. Resour. Conserv. Recycl. 54, 187–193. https://doi.org/10.1016/j.resconrec.2009.07.015.

Bel, G., Fageda, X., 2008. Local privatization , intermunicipal cooperation , transaction costs and political interests : evidence from Spain. Res. Inst. Appl. Econ. 2008, Working Papers 4, 1–23.

Bel, G., Fageda, X., 2006. Between privatization and intermunicipal cooperation: small municipalities, scale economics and transantion costs. Urbanublic Econ. Rev. 13–31.

Bel, G., Fageda, X., Mur, M., 2014. Does cooperation reduce service delivery costs? Evidence from residential solid waste services. J. Publ. Adm. Res. Theor. 24, 85–107. https://doi.org/10.1093/jopart/mus059.

Bel, G., Mur, M., 2009. Intermunicipal cooperation, privatization and waste management costs: evidence from rural municipalities. Waste Manag. 29, 2772–2778. https://doi.org/10.1016/j.wasman.2009.06.002.

Bohm, R.A., Folz, D.H., Kinnaman, T.C., Podolsky, M.J., 2010. The costs of municipal waste and recycling programs. Resour. Conserv. Recycl. 54, 864–871. https://doi.org/10.1016/j.resconrec.2010.01.005.

Brown, T.L., Potoski, M., 2003. Transaction costs and institutional explanations for government service production decisions. J. Publ. Adm. Res. Theor. 13, 441–468. https://doi.org/10.1093/jopart/mug030.

Carvalho, P., Marques, R.C., 2014. Economies of size and density in municipal solid waste recycling in Portugal. Waste Manag. 34, 12–20. https://doi.org/10.1016/j.wasman.2013.10.004.

Carvalho, P., Marques, R.C., Dollery, B., 2015. Is Bigger Better? an Empirical Analysis of Waste Management in New South Wales. Waste Management. https://doi.org/10.1016/j.wasman.2015.01.024.

Christensen, L.R., Jorgenson, D.W., Lau, L.J., 1973. Transcendental logarithmic production frontiers. Rev. Econ. Stat. 55, 28–45.

Christensen, L.R., Jorgenson, D.W., Lau, L.J., 1971. Conjugate duality and the transcendental logarithmic function. Econometrica 39, 255–256.

Cialani, C., Mortazavi, R., 2020. The cost of urban waste management: an empirical analysis of recycling patterns in Italy. Front. Sustain. Cities 2. https://doi.org/10.3389/frsc.2020.00008.

Cobb, C.W., Douglas, P.H., 1928. Theory of production. Am. Econ. Rev. 18, 139–165.

Coelli, T., Henningsen, A., 2013. Frontier: stochastic frontier analysis. R package version 1.1.

de Groot, H., van Heezik, A., Hollanders, D., Felsö, F., 2011. Cost efficiency of waste management in Dutch municipalities. In: European-American Workshop on Efficiency an Productivity Analysis. EWEPA 2011 - Verona, Italy.

Fan, X., Yu, B., Chu, Z., Chu, X., Huang, W.C., Zhang, L., 2020. A stochastic frontier analysis of the efficiency of municipal solid waste collection services in China. Sci. Total Environ. 743 https://doi.org/10.1016/j.scitotenv.2020.140707.

Filippini, M., Hrovatin, N., Zorić, J., 2008. Cost efficiency of Slovenian water distribution utilities: an application of stochastic frontier methods. J. Prod. Anal. 29, 169–182. https://doi.org/10.1007/s11123-007-0069-z.
Fried, H.O., Lovell, C.A.Knox, Schmidt, S.S., 2008. The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press.
Fusco, E., Allegrini, V., 2020. The role of spatial interdependence in local government cost efficiency: an application to waste Italian sector. Socioecon Plann Sci 69. https://doi.org/10.1016/j.seps.2019.01.003.
Griliches, Z., Ringstad, V., 1971. Economies of Scale and the Form of the Production Function; an Econometric Study of Norwegian Manufacturing Establishment Data. North-Holland Pub. Co.
Hastie, T., Tibshirani, R., 1986. Generalized additive models. Stat. Sci. 10, 354–363. https://doi.org/10.2307/2246134.
Koley, S., 2023. Arsenic Calamity in India's West Bengal: a Critical Review of Mitigation Scenarios. https://doi.org/10.1680/JENSU.21.00066.
Kumbhakar, S., Peresetsky, A., 2013. Cost efficiency of Kazakhstan and Russian banks: results from competing panel data models1. Macroecon. Finance in Emerg. Market Econ. 6, 88–113. https://doi.org/10.1080/17520843.2012.727444.
Meeusen, W., van Den Broeck, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. Int. Econ. Rev. 18, 435–444.
Molinos-Senante, M., Maziotis, A., 2021. The cost of reducing municipal unsorted solid waste: evidence from municipalities in Chile. Sustainability 13. https://doi.org/10.3390/su13126607.
Morelli, B., Cashman, S., Ma, X., Turgeon, J., Arden, S., Garland, J., 2020. Environmental and cost benefits of co-digesting food waste at wastewater treatment facilities. Water Sci. Technol. 82, 227–241. https://doi.org/10.2166/wst.2020.104.
Niaounakis, T., Blank, J., 2017. Inter-municipal cooperation, economies of scale and cost efficiency: an application of stochastic frontier analysis to Dutch municipal tax departments. Local Govern. Stud. 43, 533–554. https://doi.org/10.1080/03003930.2017.1322958.
OECD, 2009. Guidelines for Fighting Bid Rigging in Public Procurement: Helping Governments to Obtain Best Value for Money.
Richmond, J., 1974. Efficiency estimation of production functions. Int. Econ. Rev. 15, 515–521.
Sarra, A., Mazzocchitti, M., Nissi, E., 2020a. A methodological proposal to determine the optimal levels of inter-municipal cooperation in the organization of solid waste management systems. Waste Manag. 115, 56–64. https://doi.org/10.1016/j.wasman.2020.07.024.
Sarra, A., Mazzocchitti, M., Nissi, E., 2020b. Optimal regulatory choices in the organization of solid waste management systems: empirical evidence and policy implications. Environ. Sci. Pol. 114, 436–444. https://doi.org/10.1016/j.envsci.2020.09.004.
Sarra, A., Mazzocchitti, M., Rapposelli, A., 2017. Evaluating joint environmental and cost performance in municipal waste management systems through data envelopment analysis: scale effects and policy implications. Ecol. Indicat. 73 https://doi.org/10.1016/j.ecolind.2016.10.035.
Simões, P., de Witte, K., Marques, R.C., 2010. Regulatory structures and operational environment in the Portuguese waste sector. Waste Manag. 30, 1130–1137. https://doi.org/10.1016/j.wasman.2009.12.015.
Stevens, B.J., 1978. Scale, market structure, and the cost of refuse collection. Rev. Econ. Stat. 60, 438–448.
Team R Core, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria [WWW Document]. URL. http://www.r-project.org/.
Vidoli, F., Auteri, M., Marinuzzi, G., Tortorella, W., 2023. Spatial interdependence in cost efficiency and local government optimal size: the case of Italian municipalities. Socioecon Plann Sci 86. https://doi.org/10.1016/j.seps.2022.101481.
Vishwakarma, A., Kulshrestha, M., Kulshreshtha, M., 2012. Efficiency evaluation of municipal solid waste management utilities in the urban cities of the state of Madhya Pradesh, India, using stochastic frontier analysis. Benchmark Int. J. 19, 340–357. https://doi.org/10.1108/14635771211242996.
Williamson, O., 1999. Public and private bureaucracies: a transaction cost economics perspectives. J. Law Econ. Organ. 15, 306–342. https://doi.org/10.1093/jleo/15.1.306.
Williamson, O.E., 1997. Transaction cost economics and public administration. In: Public Priority Setting: Rules and Costs. Springer Netherlands, Dordrecht, pp. 19–37. https://doi.org/10.1007/978-94-009-1487-2_2.
Williamson, O.E., 1991. Comparative economic organization: the analysis of discrete structural alternatives. Adm. Sci. Q. 36, 269–296.
Zafra-Gómez, J.L., Prior, D., Díaz, A.M.P., López-Hernández, A.M., 2013. Reducing costs in times of crisis: delivery forms in small and medium sized local governments' waste management services. Publ. Adm. 91, 51–68. https://doi.org/10.1111/j.1467-9299.2011.02012.x.