The Cost of Corruption in the Italian Solid Waste Industry

The Cost of Corruption in the Italian Solid Waste Industry

Graziano Abrate

(University of Piemonte Orientale and HERMES)

Fabrizio Erbetta

(University of Piemonte Orientale, Ceris-CNR and HERMES)

Giovanni Fraquelli

(University of Piemonte Orientale, Ceris-CNR and HERMES)

Davide Vannoni*

(University of Torino and Collegio Carlo Alberto)

Abstract

The paper investigates the link between corruption and efficiency using a rich micro-level dataset on solid waste collection activities in 529 Italian municipalities observed over the years 2004-2006. To test the impact of corruption on cost efficiency, we estimate a latent class stochastic frontier model accounting for technological heterogeneity across units. The results of our estimates show that corruption significantly increases inefficiency, a finding that is robust to the inclusion of alternative local corruption indicators and other control variables such as geographical, demographic and political factors. Finally, we find that the impact of corruption tends to be greater in the southern regions of the country and in those municipalities that are less involved in recycling activities.

Keywords: corruption, cost inefficiency, latent class stochastic frontier, solid waste.

JEL codes: C33, D24, D73, Q53

**Corresponding Author*: University of Torino, Department of Economics, Mathematics and Statistics, Corso Unione Sovietica 218 bis, 10124 Torino – Italy. Tel: + 39 + 0116706083. E-mail: davide.vannoni@unito.it.

1. Introduction

According to a recent Transparency International (2010) report, corruption levels around the world have increased in spite of institutions such as the OECD and the EU having escalated their efforts to fight it. On June 2011, the EU commission estimated that corruption entails a cost of around €120 billion each year for the EU economy, i.e., 1% of the EU's GDP. The latest data on the corruption perception index classifies Italy amongst the worst EU countries with a cost of €60 billion each year estimated by the Court of Audit. These figures provide two examples of the huge impact that corruption has on a country's economy.

The scientific community has dedicated considerable research efforts to understanding the relationship between corruption and economic development and performance. Several studies have attempted to identify the channels through which corruption can alter the mechanisms of resource allocation at an aggregate (i.e., country) level. With very few exceptions, these studies provide convincing evidence that corrupt institutional and social environments exert a negative effect on economic activity and growth: corruption is harmful to investments (Mauro, 1995), to attracting foreign direct investments (Wei, 2000; Zurawicki and Habib, 2010), to the productivity of capital stock (Lambsdorff, 2003) and to technological change (Salinas-Jiménez and Salinas-Jiménez, 2007).

However, some recent studies have considered the possibility that when institutional frameworks are weak, corruption may have a positive function. For instance, Banerjee *et al*. (2006) point out that foreign private investors may select economic environments that enable acquiring a certain degree of stability through bribery. Méon and Weill (2010) find evidence that corruption is less detrimental, in terms of aggregate labour productivity, in countries where bureaucracy is slow and cumbersome and therefore more inclined to being 'greased'. Halkos and Tzeremes (2010) find a U-shaped relationship between corruption and economic efficiency, suggesting that an increase in corruption above a certain level can positively affect economic performance.

Most studies use aggregate country level data and several authors have highlighted the need to undertake more fine-grained analyses of the relationship between corruption and different performance dimensions. For example, Svensson (2005) underlines the importance of enriching the micro-evidence, stating "*... As more forms of corruption and techniques to quantify them at the micro level are developed, it should be possible to reduce this mismatch between macro and micro evidence on corruption.*" (p. 40). In a similar vein, Dal Bò and Rossi (2007) outline the crucial role of firm-level investigations. Such studies would allow analysing the relationship between corruption and individual agents and therefore "*help to pin down the ways in which corruption damages the economic performance of nations*" (Dal Bò and Rossi, 2007, p. 941).

This paper aims to contribute to the literature in two ways. First, we seek to provide new evidence on the relationship between corruption and efficiency in the solid waste industry, using micro level data from a large set of Italian municipalities. In relation to Italy, recent evidence has highlighted the phenomenon of corruption in the collection and especially in the disposal of waste. In addition, as the report on "Ecomafia" shows (Legambiente, 2011), in the context of environmental crimes, corrupt private companies, local authorities and supervisory bodies often interact and engender illegal networks that undermine the correct management of waste collection and disposal. For example, these networks exert substantial influence over the system of contracts and subcontracts through which the collection, transportation and disposal of waste are managed, or otherwise organizing extortion activities. Second, we use a latent class cost frontier specification (Orea and Kumbhakar, 2004; Greene, 2005a) where the cost frontier allows modelling inefficiency as a function of a set of variables, corruption amongst these, and the latent class model takes into account technological heterogeneity by estimating different parameters for different sub-samples of observations. This approach seems particularly suitable in the context of the waste industry since the different intensities of recycling activities across the various areas of the country could constitute a technological factor that significantly affects costs.

The remainder of this paper is organized as follows. Section 2 describes the few empirical papers that address the corruption-performance link using firm-level data and explains the mechanisms through which corruption can affect efficiency. Section 3 illustrates the latent class cost frontier approach. Section 4 presents the empirical model and describes the dataset, highlighting in particular the main characteristics of the waste management industry in Italy and how the level of corruption is measured. Section 5 discusses the results, while conclusions and policy remarks are provided in Section 6.

2. Corruption and efficiency

Very few studies to date use firm-level data and essentially focus on public utilities, analysing corruption alone or in connection with reforms such as the establishment of independent regulation authorities and openness to private capital investments. For instance, Dal Bò and Rossi (2007) estimate a labour requirement function on a set of 80 electricity distribution companies active in 13 Latin American countries and show that firms operating in more corrupt environments tend to be less efficient in terms of labour use. Estache and Rossi (2008) and Wren-Lewis (2011) extend the analysis by considering the interaction between corruption and political reforms, finding that efficiency losses due to high rates of corruption could be effectively counterbalanced by the activities of well-governed regulatory authorities[1].

These studies use firm-level data to compute productivity and other performance measures, but are still confined to a cross-country framework and, most importantly, use country-level corruption indices. Their data on the perceived levels of corruption, based on expert assessments and opinion surveys, may however raise concerns about perception biases (Reinikka and Svensson, 2005). Yan and Oum (2014) provide a single country-firm level study investigating the effect of corruption on a sample of 55 commercial airports in the US from 2001 to 2009. They indicate that corruption is detrimental to efficiency and that in more corrupt environments airports tended to contract-out and replace in-house labour. These findings confirm the predictions of their theoretical model, according to which corruption diverts managerial efforts from production control to unproductive rent-seeking activities. Yan and Oum (2014) use a state-level measure of corruption, notably the number of government officials convicted of corrupt practices for each of the 50 US states. In their words: *"Compared to the country-level corruption index, this state-level corruption measure is more objective because country-level corruption index is constructed via subjective survey evidence"* (p.16). In our study, we follow a similar approach to Yan and Oum (2014) and use disaggregated data at the council level, relating inefficiency levels to "more objective" local measures of corruption.

2.1. Efficiency measurement and heterogeneity

The scant literature that makes use of firm-level data unequivocally highlights a negative effect of corruption on efficiency. The interpretation of this result requires understanding how inefficiency is modelled. Dal Bò and Rossi (2007), Estache and Rossi (2008) and Wren-Lewis (2011) consider inefficiency as any increase in labour use due to environmental factors, the given outputs and the technology adopted. They do not however use a stochastic frontier or other model that is able to directly estimate inefficiency and test the effects of corruption in a comparative statics framework. In contrast, Estache and Kouassi (2002) estimate a stochastic production frontier model and introduce a second stage to regress inefficiency on a set of variables including corruption. Yan and Oum (2014) define a model with a set of institutional variables in a cost frontier specification to capture excess cost (interpreted as technical inefficiency) and input allocative distortions without however providing an estimate of firm-specific inefficiency levels.

[1] See also Estache and Kouassi (2002) who investigate the institutional determinants of inefficiency in the water industries of several African countries, and Estache *et al*. (2009) who explore the effect of corruption on several performance dimensions such as access, affordability and quality in the telecommunication, electricity and water industries in a set of developing countries. Overall, the regressions provide mixed (and sometimes unexpectedly negative) evidence on the role of political reforms in efficiently countering the corruption plague.

The way heterogeneity is treated is also of relevance. As outlined in Greene (2005a and b) heterogeneity can exert a strong effect on cost differentials and should therefore be properly modelled to avoid estimation bias. Moreover, heterogeneity can be only partially summarized in one or more variables and this is commonly captured by introducing individual effects in the estimation[2]. In spite of this, firm-specific effects have only been explicitly modelled in two cases (Estache and Rossi, 2008; Wren-Lewis, 2011), while Dal Bò and Rossi (2007) only include country-specific effects in their empirical model.

Literature on heterogeneity separates the effects on firm costs of environmental factors that are beyond managerial control and those that are due to the lack of appropriate incentive schemes or inadequacy in managing production processes. Some exogenous environmental factors may however affect the structure of incentives and the way these are transferred to managers. For instance, Olson *et al*. (2000) argue that country-specific governance characteristics such as the quality of bureaucracy, the diffusion of bribery and rule of law can alter incentives inherent in policy regimes and institutions, and thus limit the attainable economic performance.

The empirical studies generally support the evidence that corruption can harm firm productivity by redirecting the use of available resources. According to production theory, a firm can be interpreted as an economic agent that uses traditional inputs coordinated through managerial efforts. Any exogenous factor that prevents exerting this effort damages efficiency. In this regard, Yan and Oum (2014) and Dal Bò and Rossi (2007) provide analytical models aiming to explain the channels through which corruption can actually divert managerial efforts. Both studies essentially build on the idea that corruption leads to weak incentives and hence low levels of efficiency but differ in the underlying mechanisms (*external* versus *internal*) at stake.

Yan and Oum (2014) observe that in a more corrupt environment, policy-makers and bureaucrats tend to reduce public policy-making accountability to be in a better position to extract some private benefits. In this context, characterized by high opacity in the results of political decision-making, the advantage for governments trying to push public service providers to pursue productivity goals is rather weak. As a result, these firms, whether directly managed by officials appointed by local governments or by more independent boards, will not obtain appropriate incentives, will be less sensitive to performance and more eager to follow strategies dictated by the personal agenda of their managers. In this perspective, the underlying assumption is that the diversion of managerial

[2] In many empirical applications, the number of variables reflecting observed individual characteristics is often limited by available data. Furthermore, the search for such variables is frequently frustrated by difficulties in accessing data and by the objective inability to collect full and comprehensive information. Panel data techniques based on the introduction of fixed or random effects have therefore been developed to account for the effect of unobserved time-invariant individual characteristics.

efforts depends on external factors such as the interruption of flows of incentives along a chain of control.

Contrary to this view, Dal Bò and Rossi (2007) argue that the profitability potential of an enterprise is conditional on the control of production processes and participation in (and on the active management of) lobbying activities. The amount of effort dedicated to these activities is contingent on the internal decision of top managers. According to this approach, in more corrupt environments, the marginal return on lobbying activities would increase, so that managers devote more time to these type of unproductive activities to the detriment of efficiency such as the careful control over inputs.

Both these arguments can be applied to the waste industry case. In Italy, waste collection and disposal are mainly undertaken by publicly-owned firms under the control of local governments and, ultimately, citizens. Although in principle they should be interested in the efficient management of the mandated tasks due to the impact this has on the tax burden, the assumption that they have complete information on the technology and are able to make an informed assessment of economic performance seems somewhat unrealistic. This is especially true in contexts plagued by widespread corruption and the entrenched presence of criminal organizations. As discussed in D'Amato *et al*. (2010), the entry of organized crime in the waste disposal cycle is mainly aimed at creating non-market rents (i.e., through fraud, by manipulating tendering processes, by sub-contracting parts of the works to 'friendly' firms). In this context, the diffusion of collusive relationships among managers and suppliers aiming to overcharge firms and seeking illegal sources of profit is well documented. Furthermore, in more corrupt environments, managers are more likely to engage in negotiating activities with local governments to establish more favourable tariffs and service obligations, thereby diverting managerial efforts away from cost monitoring and productive tasks.

Despite the theoretical predictions that corruption implies a general slackness in controlling costs, the empirical evidence aimed at measuring the direct impact of corruption on overall cost efficiency is very scarce. Very few examples exists of production frontier estimations in cross-country studies (Salinas-Jiménez and Salinas-Jiménez, 2007; Halkos and Tzeremes, 2010; Méon and Weill, 2010) while in the few studies that use firm-level data, only Yan and Oum (2014) estimate a cost function frontier. Similar to Yan and Oum (2014), the approach we follow in this study enables us to measure the direct impact of corruption on overall cost efficiency and simultaneously control for latent heterogeneity.

3. Latent class cost frontier model

Our stochastic cost frontier approach assumes that producers face input prices and seek to minimize costs for the production of a certain level of output given the available technology. The deviation of a producer's observed cost from the minimum attainable level is the result of random noise and inefficiency (Kumbhakar and Lovell, 2000). Several models have been proposed to adapt stochastic frontier approaches to the panel data structure, where the availability of repeated observations over time for each unit allows capturing the impact of persistent effects on costs, which have been differently interpreted as unobserved heterogeneity and cost inefficiency.

The baseline stochastic panel data cost frontier, after log transformation, for unit i (i =1, …, N) in year t is as follows:

$$\ln C_{it} = f(y_{it}, w_{it}) + u_i + v_{it} \qquad [1]$$

where C is the production cost, y and w are respectively vectors of outputs and input factor prices, and f(.) represents the minimum attainable cost. The composite error term $\varepsilon_{it} = u_i + v_{it}$ encompasses a two-sided random noise (v_{it}), which is assumed to be independent of a non-negative error component (u_i) representing the time-invariant cost inefficiency.

A key assumption of equation [1] is that it forces all time-invariant factors affecting costs to be interpreted as inefficiency. As a result, this model has the tendency to underestimate producer performance. A number of different stochastic cost frontier models for panel data have been proposed in an attempt to separate unobserved heterogeneity from inefficiency. These models include a set of individual effects (which can be estimated using fixed or random effects econometric techniques) that should capture time-invariant heterogeneity in addition to a time-varying one-sided random term representing cost inefficiency. Since all factors that remain unchanged over time (including any persistent sources of inefficiency characterising the operational routines) are captured by the estimated individual effects, efficiency is overestimated. However, Greene (2005a) and Orea and Kumbhakar (2004) propose tackling this issue by combining the stochastic frontier approach with a latent class structure.[3] The Latent Class Stochastic Frontier Model (LCSFM) assumes that the sample units can be grouped into different classes and that the unobserved heterogeneity is modeled according to class rather than individual characteristics.

[3] Compared to the two aforementioned alternative methodological assumptions, the latent class frontier specification can be regarded as an "in-between" option that may help solve the problem of efficiency overestimation or underestimation. See Abrate et al. (2011) for an empirical application that compares efficiency scores derived from the alternative models.

Essentially, in this approach, the units are classified into different groups based on class membership probabilities directly estimated with the model. The estimation of technological parameters and cost efficiency is thus conditional on latent class membership. The LCSFM model differs substantially from a more traditional approach where the units are classified in advance into several groups and different frontiers, one for each class, are subsequently estimated.

The LCSFM specification is as follows:

$$\ln C_{it}\big|j = \ln C(y_{it}, w_{it}, \beta_j) + u_{it}\big|j + v_{it}\big|j \qquad [2]$$

where j (j = 1,…, J) denotes the class, β_j is a vector of class-specific parameters that reflect technological heterogeneity, and the $v_{it}|j$ term is a conditional normally distributed random noise with zero mean and variance σ^2_{vj}.

The class-conditional inefficiency term $u_{it}|j$ is modeled as the product of a parametric function of time, t, and a time-invariant component $u_i|j$, as specified in [3]:

$$u_{it} \mid j = g(t, \eta_j) \cdot u_i \mid j \qquad [3]$$

where η_j, based on Battese and Coelli's (1992) specification, are parameters that reflect the impact of time in each class and $u_i\ |j$ is assumed to originate from a non-negative truncated normal distribution $N^+(\mu_{ij},\ \sigma^2_{uj})$. The mean value μ_{ij} is dependent on a set of explanatory time-invariant (specific for each unit) variables, z_i, and a vector of class-specific parameters α_j, as follows:

$$\mu_{ij} = \alpha'_j z_i \qquad [4]$$

A set of time-invariant variables (among which corruption) that explain the persistent part of inefficiency reflect the panel nature of the approach. Moreover, since this model incorporates the impact of unobserved heterogeneity through a change in technological coefficients (i.e., in the shape of the frontier), it should generate more precise estimates of cost efficiency.

We adopt a maximum likelihood estimation procedure to obtain the overall set of parameters β_j, α_j and η_j. However, since class membership of each unit is subject to uncertainty in the latent class model, prior class probabilities are also estimated (contextually with the rest of the model), based on a multinomial logit specification:

$$P_{ij}\left(\delta_j\right)=\frac{\exp\left(\delta_j^{'}q_i\right)}{\sum_{j=1}^{J}\exp\left(\delta_j^{'}q_i\right)} \qquad [5]$$

where q_i is a vector of time-invariant characteristics for each producer, potentially able to influence such probabilities and δ_j are the class-specific parameters to be estimated. More specifically, any change in variables q_i with respect to the sample mean is expected to increase or decrease the probability of belonging to a given class j.

The likelihood function for unit i conditional on class j depends on β_j parameters, $LF_{ij}(\beta_j)$. Since each unit may have non-zero prior probabilities of belonging to the J classes, the individual likelihood function can be expressed as a weighted average depending not only on parameters β but also on parameters δ:

$$LF_i\left(\beta,\delta\right)=\sum_{j=1}^{J}LF_{ij}\left(\beta_j\right)P_{ij}\left(\delta_j\right) \qquad [6]$$

Therefore, the contribution of class j to the individual likelihood will depend on estimated prior probabilities and on the conditional (on class j) likelihood. In relative terms, this can be written as:

$$P\left(j,i\right)=\frac{LF_{ij}\left(\beta_j\right)P_{ij}\left(\delta_j\right)}{\sum_{j=1}^{J}LF_{ij}\left(\beta_j\right)P_{ij}\left(\delta_j\right)} \qquad [7]$$

where $P(j,i)$ is the posterior class probability, which combines the likelihood measure with the prior class membership probability.

The overall log-likelihood function to maximize, resulting from the aggregation of individual likelihood measures formalized in [6], is obtained as:

$$\ln LF\left(\beta,\delta\right)=\ln\prod_{i=1}^{N}LF_i\left(\beta,\delta\right)=\ln\prod_{i=1}^{N}\left\{\sum_{j=1}^{J}LF_{ij}\left(\beta_j\right)P_{ij}\left(\delta_j\right)\right\} \qquad [8]$$

where it can be noted that the maximization process involves both the set of structural parameters of the cost frontier (β_j) (and, implicitly, the parameters α_j relative to the inefficiency term in the conditional mean model) as well as the set of parameters reflecting any regularity in the classification on the sample (δ_j).

Finally, following Greene (2002), the estimation of the conditional cost efficiency for unit i can be obtained as:

$$CE_i = \exp(-\hat{u}_i) = \exp\left[-\sum_{j=1}^{J} E(u_i|\varepsilon_{it}, j)P(j,i)\right] \quad [9]$$

where the technology of reference of every class is taken into account.

4. Data description and model specification

The Italian Municipal Solid Waste (MSW) industry has several characteristics that make it ideal to test the existence of a relation between corruption and inefficiency. First, the heterogeneous provincial levels of corruption (see Section 4.2) mirror the local character of the industry. Indeed, despite the recent reforms aimed at discouraging in-house service provision by municipalities and favoring competitive tendering, a strict relation between operators and municipalities remains.[4] As a result, most firms still operate in a single municipality or, in any case, within its neighborhood, especially in the collection phase. Second, in spite of using firm level balance sheet data, we are able to observe the total cost of the integrated waste cycle at the municipality level.[5] Finally, some recent cases of bad MSW management have emerged. The media widely reported the recent waste crisis in Naples and surrounding area, while a number of books and movies have clearly depicted the connection between waste management and illegal practices.[6]

The database refers to a balanced panel of 529 Italian municipalities observed from 2004 to 2006, representing more than one third of the Italian population. Table 1a presents the summary statistics of the variables used for the cost frontier specification.

The technology behind the multi-product nature of the MSW service is described using three outputs, two input prices and one fixed input constraint as determinants of total cost. In terms of outputs, we consider two distinct measures of waste produced (tons of MSW disposed, Y_D, and tons of MSW sent for recycling, Y_R) as well as the number of buildings (Y_B) that proxies the number of collection points. This specification allows distinguishing between returns to density (occurring when the tons produced increase but the number of buildings remains unchanged) and returns to scale. Given the congestion problems that can arise in MSW management (especially in relation to finding disposal sites), we also account for the municipality surface area (S), which acts as a constraint linked to the availability of land for new urbanization. This allows a further interpretation of the returns to scale in the long-run, when it may be possible to think of municipality mergers, in

[4] In Italy, a province is an administrative division on an intermediate level between a municipality and a region. A province is composed of many municipalities and several provinces usually form a region.
[5] The information on the costs and the amount of waste collected was gathered from annual MUDs (i.e., annual declarations of municipal solid waste collection).
[6] Among these, the best-seller book *Gomorrah* (Saviano, 2006) reached a large international audience and contributed to sensitizing public opinion against the plague of environmental crimes.

which case the surface area increases along with the three measures of outputs described above. Total cost (C) is the sum of labor, capital and fuel costs borne by the council. Input prices are computed by integrating the information available in the MUDs with additional information drawn from questionnaires sent to the firms (or organizational structures) managing the service in the municipalities. The price of labor (w_L) is given by the ratio of total salary cost to the number of full-time equivalent employees. Capital price (w_K) is obtained by dividing depreciation costs by the capital stock. Similarly to Antonioli and Filippini (2002), the price of fuel is assumed to be the same for all municipalities.

Following Christensen *et al*. (1973), we use a flexible translog functional form, which is quite common in empirical cost and production function studies.[7] The empirical cost frontier, for class j, takes the following expression (where i denotes the municipality and t = 1,…, T time):

$$\begin{aligned}\ln\left(C_{it}/w_{Kit}\right) = \beta_{0j} + \sum_{m\in(D,R,B)} \beta_{mj} \ln Y_{mit} + \beta_{Sj} \ln S_{it} + \beta_{Lj} \ln\left(w_{Lit}/w_{Kit}\right) + \frac{1}{2}\sum_{m\in(D,R,B)}\sum_{n\in(D,R,B)} \beta_{mnj} \ln Y_{mit} \ln Y_{nit} \\ + \frac{1}{2}\beta_{SSj}\left(\ln S_{it}\right)^2 + \frac{1}{2}\beta_{LLj}\left(\ln w_{Lit}/w_{Kit}\right)^2 + \sum_{m\in(D,R,B)} \beta_{mSj} \ln Y_{mit} \ln S_{it} \\ + \sum_{m\in(D,R,B)} \beta_{mLj} \ln Y_{mit} \ln\left(w_{Lit}/w_{Kit}\right) + \beta_{SLj} \ln S_{it} \ln\left(w_{Lit}/w_{Kit}\right) + u_{it}|j + v_{it}|j \qquad [10]\end{aligned}$$

where

$$u_{it}|j = \exp\left(-\eta_j (t-T)\right)\cdot u_i|j \qquad [11]$$

For the empirical analysis, the variables were standardized on their geometric mean. In order to impose homogeneity of degree one in input prices, C_{it} and labor price were normalized over the price of capital. Symmetry conditions $\beta_{mnj} = \beta_{nmj}$ were also imposed.

The time-varying inefficiency component in [11] is modeled according to Battese and Coelli's (1992) specification, where the η_j parameters capture the direction and magnitude of changes in cost inefficiency over time. The time-invariant cost inefficiency component $u_i|j$ is modeled on a set of z_i variables, amongst which corruption (see [4]).

[7] However, as discussed by Abrate et al. (2014a), most empirical studies on the costs of refuse collection are still based on very simple specifications such as the Cobb-Douglas model.

4.1. Determinants of latent class prior probabilities and cost inefficiency specification

The latent class approach has the advantage of estimating separate technological parameters for different groups of observations by specifying a given number of classes. Groups are determined through the modeling process rather than being forced into predefined categories, and a set of determinants (q_i) of *prior* latent class probabilities are considered to capture any regularity in the definition of classes. In our study, we use a two-class specification and compare it to a model using a single class. According to Orea and Kumbhakar (2004), goodness of fit measures such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) can be used in case of ambiguity of the appropriate number of classes. The adoption of these criteria led us to focus on the two-class specification, which present the lowest values of both AIC and BIC (AIC = -0.74 and BIC = -0.56 compared to AIC = -0.58 BIC = -0.50 for the one-class model). In addition, the estimation with more than two classes failed to converge, thus suggesting that the discrimination ability of the model can lead, at best, to a clear identification of two classes.[8]

Table 1b describes, for different geographical areas, some key characteristics that may be important technological sources of heterogeneity between operators. The per-capita production of waste (*QPC*) is on average around 470 kilograms per year (almost 1.3 kg per day), with a share of recycling (*SHREC*) of around 20 percent. There is great variance in the sample, especially in terms of waste collection policies: while some municipalities had not yet undertaken a serious recycling program during the years in question, others had already reached significant targets, with a maximum share of almost 75 percent. Moreover, there is a remarkable gap between North (with over 36 percent of recycling on average) and other regions of the country (only 7.1 percent in the South). Since increasing the share of recycling is a political decision, we also collected data on the type of municipal government during the observed period. The data indicates that left-wing (as opposed to centre and right-wing) political orientation is prevalent in around 30 percent of municipalities, with a lower penetration in the South. Moreover, this tends to be associated with a slightly higher recycling share (on average 22.2 percent against 18.7 percent), which may indicate greater attention to environmental policies.

Another important variable is population density. In this respect, we split the usual density measure (*DENS* = population per square kilometer) into two variables, used as proxies for the vertical and horizontal degree of urbanization:

$$DENS = \frac{Population}{Km^2} = \frac{Population}{Y_B} \times \frac{Y_B}{Km^2} = VDENS \text{ x } HDENS \qquad [12]$$

[8] For a similar approach, see Abrate et al. (2011) and Cullmann (2012).

In [12], *VDENS* captures the presence of tall buildings (with a high number of floors and flats) and is typically associated with largest municipalities. Coherently, the highest value is registered in the North, which has relatively fewer small municipalities (around 39 percent of cases below the median level of 18,550 inhabitants). Since the incidence of small municipalities exceeds 61 percent in the South, a remarkably lower vertical degree of urbanization is observed in this region. Conversely, *HDENS*, as a measure of horizontal congestion, is less correlated with small and large municipalities, and a higher value of this indicator could imply more severe constraints on land use (with consequences, for example, on availability of areas for waste disposal).

Exploiting the features of the latent class model, the variables *QPC*, *SHREC*, *VDENS*, *HDENS* have been included as determinants of prior class probabilities in the following multinomial logit model:

$$P_{ij}\left(\delta_j\right)=\frac{\exp\left(\delta_{0j}+\delta_{1j}HDENS_i+\delta_{2j}VDENS_i+\delta_{3j}SHREC_i+\delta_{4j}QPC_i\right)}{\sum_{j=1}^{J}\exp\left(\delta_{0j}+\delta_{1j}HDENS_i+\delta_{2j}VDENS_i+\delta_{3j}SHREC_i+\delta_{4j}QPC_i\right)} \qquad [13]$$

from which can easily be noted that given $J = 2$, $P_{i2} = 1-P_{i1}$.

Finally, our time-invariant cost efficiency term $u_i|j$ is regressed on different measures of corruption (*CORR*). To test the robustness of the effect of corruption we included additional variables that may be related to municipal inefficiency. First, regional efficiency differentials have been often documented in Italy, particularly between northern and southern regions (Erbetta and Petraglia, 2011). Second, municipality size may be related to local public service efficiency beyond pure scale effects captured by returns to scale (e.g. De Borger et al., 1994). Finally, we accounted for a possible relationship between efficiency and political stance (on this topic, see for example De Borger and Kerstens, 1996). We thus tested the impact of the dummy variables representing small municipalities (*DSMALL*) and municipalities ruled by left-wing political parties (*DLWPOL*), together with two geographical dummies (*DNORTH* and *DCENTER*, respectively corresponding to the northern and central regions of the country). More specifically, the time-invariant cost inefficiency function for class *j* is modeled as follows:

$$\begin{aligned}\mu_{ij} &= \alpha_{rj}\mathrm{CORR}_i + \alpha_{\mathrm{NORTH}j}\mathrm{DNORTH}_i + \alpha_{\mathrm{CENTRE}j}\mathrm{DCENTER}_i \\ &+ \alpha_{\mathrm{SMALL}j}\mathrm{DSMALL}_i + \alpha_{\mathrm{LWPOL}j}\mathrm{DLWPOL}_i\end{aligned} \qquad [14]$$

To summarize our approach, Table 1c provides the full list of variables included in the three sub-models: 1) the cost frontier [10]; 2) the prior class probabilities model [13]; 3) the inefficiency term [14]. In any case, all equations are estimated simultaneously.

4.2. Measures of corruption

A crucial point of the analysis concerns the complex assessment of corruption. As argued by Golden and Picci (2005), directly measuring corruption is "*an enterprise that is not possible since corruption is a complex set of variable interactions, processes and phenomena with no single metric*" (p. 37)[9]. We therefore decided to follow three alternative approaches. A first index was obtained using publicly available data from the National Institute of Statistics (ISTAT), a second index was derived directly from economic literature (Golden and Picci, 2005), and a third index was computed borrowing from sociological literature (Block, 1980). All variables are defined at the provincial level while the presence of around 100 provinces ensures the high diversification of the indices across municipalities.

$CORR_1$ represents the number of crimes against the State, public governments and social institutions (per 100,000 inhabitants), and consists of an aggregate indicator that includes crimes such as embezzlement, extortion, conspiracy and those against faith and public order. This measure, which is used by Del Monte and Papagni (2007), has the drawback of underestimating the true phenomenon as it only considers those crimes reported to the police and therefore does not reflect corruption crimes in their entirety.[10] As expected, given that the South's public policies and social system are greatly penalized by the actions of powerful and long-standing criminal organizations, the value is higher than that recorded in the North, even if the difference is not large and is associated with a higher standard deviation (see Table 1d). On average, the level of corruption does not appear to be affected by the size of operations although there are some different territorial tendencies: in particular, small municipalities in the Center have the worst level, while they appear to be slightly better than large municipalities in the northern areas. In terms of political stance, there is a slightly lower level of corruption in municipalities ruled by left-wing parties, although this evidence is only confirmed in the Center.

Golden and Picci (2005) propose and compute $CORR_2$, which is available for all Italian provinces and reflects "*the difference between the amount of physically existing public infrastructure (roads, schools, hospitals, etc.) and the amount of money cumulatively allocated by government to create this public works. Where the difference between the two is larger, more money is being lost to fraud, embezzlement, waste and mismanagement*" (p. 37). The underlying concept is that corruption

[9] We are indebted to an anonymous referee for suggesting a more in-depth discussion of the corruption measures used in this paper.

[10] In particular, this measure can be questioned to the extent that the level of enforcement differs across municipalities. However, Del Monte and Papagni (2007) defend the quality of this variable as an indicator of corruption by showing that it is highly correlated to the traditional corruption perception index (CPI), which being based on survey data, is not subject to this criticism.

increases the cost of building public infrastructures and $CORR_2$ is computed as the ratio between the building cost and the actual value of the public investment.

As recognized by Golden and Picci (2005), the measure captures some inefficiencies as well as various illegal activities that constitute genuine corruption. However, the fact that public sector costs are found to vary geographically in ways quite different from private sector costs *"offers persuasive, if indirect, confirmation that the measure of corruption we create is valid, reflecting the extensive graft and fraud that are especially common to public sector contracting, rather than mere inefficiencies."* (p. 38).

Table 1d shows that in the South, this measure takes on values that are almost double those of the other regions. As to the differences associated with size and political stance, these are very much in line with those in $CORR_1$. Indeed, the pair-wise correlation between $CORR_1$ and $CORR_2$ is equal to 0.39 and is strongly significant (Table 1e).

Finally, building on the sociological approach, we use two indices ($CORR_{ES}$ and $CORR_{PS}$) that are more closely linked to the societal penetration of organized crime.[11] Block (1980), analyzing the case of New York, distinguished between two types of criminal organizations, the first (*enterprise syndicates*) operate in economic and enterprise networks managing illegal affairs while the second (*power syndicates*) aim for power to control the community. In his words, "*its forte is extortion not enterprise [...]"* and infiltrates *"the industrial world specifically in the labor-management disputes*" (p. 129). The European Commission (2010) analyzed the links between the above measures of organized crime and several indices of corruption, finding that the two phenomena are closely intertwined, especially in Italy, with very high correlation coefficients.[12]

The Fondazione RES (2011) report provides alternative measures of Italian criminality, disaggregated at the provincial level. The power syndicate index is based on 5 indicators: number of properties seized from organized crime, number of city council dissolutions, mafia association crimes, number of mafia murders and number of extortions. As to enterprise syndicates, a further 5 types of crime are considered: criminal conspiracy, drug trafficking, exploitation of prostitution, robbery and usury. For each indicator, a ratio is calculated between provincial and national data. Thus, a ratio greater than 1 for a province would indicate a criminal activity higher than that prevailing at the national level. Accordingly, two dummy variables are constructed for power ($CORR_{PS}$) and enterprise syndicates ($CORR_{ES}$), which take on the value of 1 when the average between the five ratios is higher than the unity. As shown in Table 1d, power syndicates have a high

[11] For a comparison with the other two measures, $CORR_1$ and $CORR_2$, we also keep the label "corruption" for $CORR_{ES}$ and $CORR_{PS}$, while it would be more correct to consider the latter as "criminality" indices.

[12] *"When investigations into corrupt activities are launched, investigators usually discover some criminal organisation's involvement. By the same token, when organised crime is investigated, the involvement of corrupt politicians or entrepreneurs often comes to light"* (European Commission, 2010, p. 18).

presence in southern municipalities, especially in areas traditionally dominated by criminal organizations such as the *Mafia* (Sicily), the *'Ndrangheta* (Calabria) and the *Camorra* (Campania). Conversely, the diffusion of enterprise syndicates is widespread all over Italy but prevalent in the North. Also interesting to note is that enterprise syndicates in small municipalities do not vary greatly across geographical areas, while they appear to be stronger (less strong) in the municipalities in northern (central) Italy ruled by left-wing political parties. Finally, Table 1e indicates the absence of a significant correlation between power and economic syndicates, confirming that these refer to different aspects of criminality. Conversely, both measures significantly correlate with the alternative corruption indicators used in the study ($CORR_1$ and $CORR_2$).

5. Results

As discussed in the previous sections, our empirical strategy requires the simultaneous estimation of three types of class-conditional parameters with reference to the stochastic cost frontier model, the prior probability model and the cost inefficiency model. The three key research questions we test are:

- Does the underlying technological heterogeneity actually allow identifying separate classes?
- Are these classes polarized on the basis of output composition, and, in particular, according to the incidence of recycling activities?
- Is cost inefficiency, measured as the residual of the stochastic cost frontier of each class, differently affected by corruption?

Table 2 compares the results of the baseline model with two latent classes (columns 2 and 3) with those stemming from a model that assumes a unique class (column 1).[13] All first order parameters are strongly significant and have the expected sign and, given the log-log transformation and data normalization, can be interpreted as cost elasticities for the "average" municipality. The coefficient associated with the relative price of labour (β_L) can also be interpreted as the cost share of labour and is equal, on average, to 0.58 (Column 1). However, Class 1 has a significantly lower labour share (0.40) than Class 2 (0.71). From Column 1, the cost output elasticities (ε) are respectively equal to $\beta_D = \varepsilon_{CY_D} = 0.73$, $\beta_R = \varepsilon_{CY_R} = 0.19$ and $\beta_B = \varepsilon_{CY_B} = 0.17$. Moreover, the negative sign of β_{DR} suggests the presence of cost complementarities between disposal and recycling activities.[14] Since

[13] The baseline model corresponds to when $CORR_1$ is included as the unique determinant of inefficiency. Table 4 shows the results of regressions where the variants of the inefficiency model were tested.

[14] Since the focus of this paper is the relationship between corruption and cost efficiency, we do not discuss solid waste industry technology in-depth. The interested reader can refer to Abrate *et al*. (2014a).

the cost function includes number of buildings together with tons of waste disposed and recycled, we can separately compute a measure of output density economies and a measure of scale (or size) economies. The former reflects the proportional increase in costs due to an increase in output, keeping the number of buildings fixed, (ODE = $1/(\beta_D + \beta_R)$), while the latter reflects a simultaneous increase of both outputs and number of buildings (RTS = $1/(\beta_D + \beta_R + \beta_B)$). All three columns indicate that the average municipality shows increasing returns to density (ODE ranges from 1.08 to 1.17) and decreasing returns to size (RTS ranges from 0.88 to 0.91). This result is commonly found in literature (see Abrate et al., 2014b), suggesting that costs increase more than proportionally with respect to increases in the amount of waste disposed, waste sent for recycling and the number of buildings in a given area. Worth noting is the negative sign associated with surface area, which can be interpreted as a fixed factor acting as a constraint in terms of land availability for each municipality. Even if the surface area is an environmental constraint that cannot be modified in the long-run, in principle, municipality mergers are conceivable whereby S increases along with Y_D, Y_R and Y_B. Accepting this hypothesis, and including coefficient β_S in computing scale (or size) economies, the returns to scale (size) are constant, suggesting that the increase of surface area partially counteracts the effect of a disproportionate rise in refuse collection costs that occurs when Y_D, Y_R and Y_B increase in congested areas.

For each observation, we can measure relative performance in terms of efficiency with respect to the cost frontier (maximum efficiency =1): the higher the distance from the frontier, the lower the efficiency[15]. Moving from the single-class to the two-class model, the average efficiency in the sample increases by almost 10 percentage points (from 0.594 to 0.683), given that each observation compares its cost performance with respect to its own class frontier.

Table 2 also shows the results of the multinomial logit model represented by [13], which aims to disentangle the determinants of prior probabilities. All variables included in the prior probability model have a significant role in assigning class membership. The negative sign attributed to *SHREC* indicates that municipalities with lower recycling shares are more likely to belong to Class 1. The same is true for towns with low per-capita waste production, and a low (high) degree of vertical (horizontal) urbanization. A hypothetical municipality with average values with respect to all four variables would have a slightly higher prior probability of belonging to Class 1 (around 53 percent). However, any deviation from such average values affects the class membership probabilities.

[15] It should be recognized that cost minimization may not be a goal for municipalities. If this were to occur, they could tend to pay inputs beyond fair market prices. Since the cost frontier specification includes information on input prices, this could in turn inflate the reference cost along the stochastic frontier, leading to a milder target cost. We thank an anonymous referee for having raised this point. In order to check the robustness of our estimates we tested a model without input prices. The results remain substantially the same.

Table 3 describes the main characteristics of the two classes identified according to *posterior* probabilities. We can observe the polarization between classes with respect to several variables (among which those included in the prior probability model). The average posterior probability value indicates that class membership is well identified (respectively 92 and 84 percent). In particular, higher recycling share is confirmed as a distinctive feature of Class 2. This class also encompasses more municipalities in the North – where the recycling programs are, on average, more diffused – as well as larger municipalities (the share of DSMALL is 42 percent). In Class 2, we also observe almost twice as many municipalities ruled by left-wing parties, which tend to declare more explicit "environment-oriented" goals. As to the corruption measures, it is interesting to note that Class 1 tends to have slightly higher average values. However, the difference appears relevant only with respect to the power syndicate index, i.e., the only corruption measure that is linked to geographical factors given its higher level in the Southern regions. To the contrary, the enterprise syndicate indicator is more evenly distributed across classes, with a slight prevalence in Class 2.

We turn now to the core of our analysis, namely, the relationship between corruption and efficiency. Table 2 shows an average efficiency score for the baseline model of 0.683, while Table 3 highlights the remarkable difference between municipalities clustered in Class 1, where the efficiency score is 0.60 and municipalities belonging to Class 2, where it is 0.77. This shows the large potential for efficiency gains especially for Class 1 (i.e., the less virtuous class in implementing recycling programs). Moreover, looking at Table 2, we note that both classes have negative η-parameters, which implies an increase of inefficiency over time; however, the trend is statistically significant only for Class 1. As to structural inefficiency, the lower part of Table 2 reporting the estimates of [4] shows that the corruption index exerts a positive effect on inefficiency, although again the effect is statistically significant only for Class 1. This suggests that the greater recycling-orientated Class 2 also appears to be more virtuous in terms of being less influenced by corruption. For Class 1, we can instead quantify the impact of corruption. If we simulate an increase of $CORR_1$ from the 1st quantile (4.2) to the median value (5.27), the inefficiency increases by almost 10 percent. Moving from the median to the 3rd quantile (6.84), a further 14 percent increase is obtained.[16]

Table 4 tests alternative variants of the cost inefficiency model [14].[17] Several control variables in relation to geographical area, municipality size and political stance were added to $CORR_1$. In fact, a

[16] Also note in the hypothesis of a single class (first column of Table 2) that the coefficient on corruption is positive, strongly significant and of a similar magnitude.

[17] We decided to present only the results for the inefficiency model since the cost frontier parameters and the class characteristics are very similar to that presented for the baseline model. In particular, considering the alternative cases with two classes (Model II, III and IV with respect to the baseline model), the number of municipalities changing class

major concern could be that the highlighted effect of corruption on inefficiency actually hides other explanatory factors, in particular territorial heterogeneity. The first two columns of Table 4 (Models I and II) show the results of the models with one and two classes respectively. The role of corruption is confirmed and in this case is significant, but less strong, also for Class 2.[18] We also ran some simple simulations to estimate a monetary value for cost inefficiency due to corruption. Using the Model II estimates, for example, we find that a generalized reduction of the corruption level by 10 percent would generate total cost savings in our sample of around €92 million per year. This accounts for around 4.7 percent of refuse collection costs and corresponds to around €4.2 per inhabitant. Moreover, if we take the example of the two biggest Italian municipalities, Milan and Rome, reducing their corruption index to the sample average value would bring cost saving of 10 and 50 million € per year (i.e., 8.8 and 14 percent of total costs respectively).[19]

The control variables are often non-significant, suggesting that corruption is the key driver of inefficiency. There is however some evidence of higher inefficiency associated with left-wing politics (especially for Class 1). This result could be explained by the fact that left-wing parties tend to be more oriented to pursuing equity and environmental goals than efficiency gains.

Finally, we test for the effect of alternative corruption indices while maintaining the same set of control variables. Model III includes $CORR_2$, which is a "missing expenditure" measure of corruption based on public infrastructure spending. Also in this case, the index is positive and significant in both classes and the magnitude of the effect is substantially similar for both sub-groups. Moving from the first quantile (0.81) to the median value (1.30) and to the 3rd quantile (1.88), inefficiency increases by 5-6 percent at each step. The last two columns test the impact of $CORR_{PS}$ and $CORR_{ES}$, which are inspired by the sociological approach to the measurement of organized crime. They both show a positive sign although, interestingly, the power syndicate index is significant only in Class 1 while the enterprise syndicate index is significant in Class 2.[20] Although this polarized result should be considered with caution, a possible interpretation could be the following. In municipalities where recycling shares are low – mainly distributed in the South – efficiency may be more seriously harmed by power syndicates, which can gain control of disposal sites and force enterprise decisions through extortion. Conversely, in municipalities that recycle

---

membership is very small (from 7.2 to 11.3 percent depending on the pair-wise comparisons). The full set of estimates is available on request.

[18] To correctly compare the effects in the two groups, we should take into account eventual differences in the distribution of $CORR_1$. However, Table 3 shows similar mean and standard deviations for the two sub-groups and the same can be said with respect to the quantile values. Therefore, the simple comparison of coefficients is a fairly good approximation.

[19] The cost reduction percentages are computed with reference to the estimated cost frontiers. As concerns the specific figures given for Milan and Rome, the simulation assumes a reduction of their actual corruption index levels (6.62 and 7.73 respectively) to the average sample value (5.49).

[20] Since the indicators are dummies, the coefficients directly give the inefficiency percentage difference between municipalities with higher than average corruption and those with lower than average corruption.

more – mainly distributed in the North – more "economic oriented" criminal organizations may enter the waste chain, for example, through the direct management of recycling activities.

6. Conclusions

Given that corruption is generally perceived as a "social disorder" that can negatively affect the growth and performance of nations, empirical studies that try to measure the extent of corruption and quantify its impact on several dimensions of performance should be welcome.

To date, empirical literature is mostly oriented towards macroeconomic objectives where the aggregates of reference are represented by individual countries. In contrast, there are relatively few studies that use micro-level data to investigate the corruption-performance link.

Our paper aims to contribute to this strand of literature by using a latent class stochastic cost frontier approach to evaluate the impact of corruption on cost efficiency in the provision of solid waste disposal services of a sample of Italian municipalities. The refuse collection industry, according to some recent anecdotal and judicial evidence, is a suitable case study.

Our approach is based on measuring efficiency as the distance from the estimated frontier, while unobserved heterogeneity is controlled for by identifying subsets of observations that constitute separate technological classes. Since recycling programs differ substantially between municipalities, we have treated the volume of waste sent to disposal sites or incinerated and the volume of waste sent for recycling as separate output variables, and included among the regressors the number of collection points and the size of the municipal area. As a final step of our analysis, the inefficiency estimates generated by the cost frontier are regressed on a set of explanatory variables including corruption.

Our results can be summarized as follows. First, the estimated class-specific coefficients show two well-behaved cost frontiers that actually differ in terms of underlying technologies, validating our choice to rely on a latent class stochastic frontier approach. Second, this differentiation depends on density variables and, as expected, on the share of waste sent for recycling. More specifically, we observe that the class with more recycling programs (Class 2) is more labor-intensive, less sensitive to the effect of increasing costs due to demographic pressure and denotes a higher share of left-wing political orientation. Third, and most important, corruption is found to have a negative and significant effect on cost efficiency, thus confirming previous theoretical and empirical evidence. As a novel contribution to literature, which very rarely uses disaggregated corruption indices, we test the impact of several "local" measures of corruption and criminal behavior, and find that the detrimental effect on efficiency is remarkably robust. Moreover, the impact depends on the class type and appears to be less important in Class 2, which has a higher share of municipalities in the North and municipalities involved in major recycling programs. The cost efficiency of this class is, however, negatively affected by the $CORR_{ES}$ index, which accounts for the presence of enterprise syndicates. To the contrary, in the Southern areas of the country, $CORR_{PS}$ seems to be more important, suggesting that illegal networks, and especially criminal organizations that are more involved in controlling the territory, find it easier to infiltrate collection and disposal activities. In this sense, the possibility of managing illegal landfill clearly plays a crucial role.

From a policy standpoint, our results show that fighting corruption and criminal organizations would bring considerable efficiency gains. According to some simple simulations, if it were possible to reduce the corruption level in the two biggest Italian cities, Milan and Rome, moving their corruption index just to the sample average value, yearly cost savings would be in the order of respectively 10 and 50 million €. Moreover, our latent class stochastic frontier approach disentangles two sub-groups of municipalities showing different average values of efficiency and affected differently by corruption. Accepting the higher share of recycling as the most distinctive feature of Class 2, this suggests that pushing municipalities more actively to pursue recycling programs could bring about efficiency improvements and at the same time reduce the adverse (in terms of cost efficiency) effects of corruption.

References

Abrate G., Erbetta, F., Fraquelli G., 2011. Public Utility Planning and Cost Efficiency in a Decentralized Regulation Context: The Case of the Italian Integrated Water Service. *Journal of Productivity Analysis* 35: 227-242.

Abrate G., Erbetta, F., Fraquelli G., Vannoni D., 2014a. The Costs of Disposal and Recycling: An Application to Italian Municipal Solid Waste Services. *Regional Studies*, forthcoming.

Abrate G., Erbetta F., Fraquelli G., Vannoni D., 2014b. Size and Density Economies in Refusal Collection, in Kinnaman T. and Takeuchi K. (eds.), *Handbook on Waste Management*, Edward Elgar.

Antonioli B., Filippini, M., 2002. Optimal Size in the Waste Collection Sector. *Review of Industrial Organization* 20 (3): 239-252.

Battese G., Coelli, T., 1992. Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis* 3: 153-169.

Banerjee S.G., Oetzel, J.M., Ranganathan, R., 2006. Private Provision of Infrastructure in Emerging Markets: Do Institutions Matter? *Development Policy Review* 24 (2): 175-202.

Block A.A., 1980. *East Side-West Side: Organizing Crime in New York*. University College of Cardiff Press, Cardiff.

Christensen L., Jorgensen, D., Lau, L., 1973. Transcendental Logarithmic Production Frontiers. *Review of Economics and Statistics* 55: 28-45.

Cullmann A., 2012. Benchmarking and firm heterogeneity: a latent class analysis for German electricity distribution companies. *Empirical Economics* 42: 147-169.

Dal Bó E., Rossi, M.A., 2007. Corruption and Inefficiency: Theory and Evidence from Electric Utilities. *Journal of Public Economics* 91: 939-962.

D’Amato A., Mazzanti, M., Nicolli, F., 2010. Waste Sustainability, Environmental Policy and Mafia Rents: Analysing Geographical and Economic Dimensions. Mimeo.

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

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

Del Monte A., Papagni E., 2007. The Determinants of Corruption in Italy: Regional Panel Data Analysis. *European Journal of Political Economy*, 23: 379-96.

Erbetta F., Petraglia C., 2011. Drivers of regional efficiency differentials in Italy: technical inefficiency or allocative distortions? *Growth and Change* 42(3): 351-375

Estache A., Goicoechea, A., Trujillo, L., 2009. Utilities Reforms and Corruption in Developing Countries. *Utilities Policy* 17: 191-202.

Estache A., Kouassi, E., 2002. Sector Organization, Governance and the Inefficiency of African Water Utilities. *Policy Research Working Paper* 2890, World Bank Institute.

Estache A., Rossi M. A., 2008. Regulatory Agencies: Impact on Firm Performance and Social Welfare, *Policy Research Working Paper* 4509, The World Bank.

European Commission, 2010. *Examining the Links Between Organized Crime and Corruption. Center For Study of Democracy*. European Commission, Brussels.

Fondazione RES, 2011. Alleanze nell'ombra. Mafie ed Economie Locali in Sicilia e nel Mezzogiorno. *Rapporto RES 2010*, Donzelli Editore, Roma.

Golden M.A., Picci, L., 2005. Proposal for a New Measure of Corruption, Illustrated with Italian Data. *Economics and Politics* 17 (1): 37-75.

Greene W.H., 2002. Alternative Panel Data Estimators for Stochastic Frontier Models. Working Paper, Department of Economics, Stern School of Business, NYU.

Greene W.H., 2005a. Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model. *Journal of Econometrics* 126: 269–303.

Greene W.H., 2005b. Fixed and Random Effects in Stochastic Frontier Models. *Journal of Productivity Analysis* 23 (1): 7–32.

Halkos G.E., Tzeremes, N.G., 2010. Corruption and Economic Efficiency: Panel Data Evidence. *Global Economic Review*, 39 (4): 441-454.

Kumbhakar S., Lovell, C.A., 2000. *Stochastic Frontier Analysis*, Cambridge University Press.

Lambsdorff J., 2003. How Corruption Affects Productivity. *Kyklos* 56 (4): 457-474.

Legambiente (2011). *Ecomafia 2011. Le Storie e i Numeri della Criminalità Ambientale*. Edizioni Ambiente, Milano.

Mauro P., 1995. Corruption and Growth. *Quarterly Journal of Economics* 110 (3): 681–712.

Méon P.G., Weill, L., 2010. Is Corruption an Efficient Grease? *World Development* 38 (3): 244-259.

Olson M., Sarna, N., Swamy, A.V., 2000. Governance and Growth: a Simple Hypothesis Explaining Cross-Country Differences in Productivity Growth. *Public Choice* 102: 341-364.

Orea L., Kumbhakar, S.C., 2004. Efficiency Measurement Using a Latent Class Stochastic Frontier Model. *Empirical Economics* 29:169–183.

Reinikka R., Svensson, J., 2005. Using Micro-Surveys to Measure and Explain Corruption. *World Development* 34 (2) 359-370.

Salinas-Jiménez M., Salinas-Jiménez, J., 2007. Corruption, Efficiency and Productivity in OECD Countries. *Journal of Policy Modeling* 29: 903-915.

Saviano R., 2006. *Gomorrah. A Personal Journey into the Violent International Empire of Naples' Organized Crime System*. Mondadori, Milano.

Svensson J., 2005. Eight Questions About Corruption . *Journal of Economic Perspectives* 19 (3): 19-42.

Transparency International, 2010. Global Corruption Barometer 2010. www.transparency.org, Berlin, Germany.

Wei S.J., 2000. How Taxing is Corruption on International Investors? *Review of Economics and Statistics* 82 (1): 1-11.

Wren-Lewis L., 2011. Do Infrastructure Reforms Reduce the Effect of Corruption? Theory and Evidence from Latin America and the Caribbean. *ECARES working paper* 26, Université Libre de Bruxelles.

Yan J., Oum, T.H., 2014. Effects of Government Corruption on the Efficiency of U.S. Commercial Airports, *Journal of Urban Economics*, forthcoming.

Zurawicki L., Habib, M., 2010. Corruption and Foreign Direct Investment. What Have We Learned? *International Business & Economics Research Journal* 9 (7): 1-10.

Table 1a. Summary statistics: cost, outputs and prices

Variable	*Description*	Mean	Std. dev.	Min	Max
C	Total cost *(000 €)*	5,436	23,965	46	48,065
Y_D	Waste disposed (t)	17,122	71,196	118.44	1,462,128
Y_R	Waste recycled (t)	3,770	13,044	8.86	210,211
Y_B	Number of buildings	4,960	7,309	353	127,713
S	Surface (Km^2)	83.44	106.16	2	1285
w_L	Price of labor (€ / Employee)	36,607	5,735	22,663	62,613
w_K	Price of capital (depreciation rate)	0.087	0.013	0.049	0.124

Table 1b. Summary statistics: sample characteristics by geographical area

Variable	*Description*	North	Center	South	Total
	Number of municipalities	204	118	207	529
QPC	Total waste per capita (kg) *	446 (142)	521 (124)	464 (148)	470 (143)
SHREC	Share of recycling (%)*	36.3 (15.0)	13.4 (9.8)	7.1 (7.5)	19.8 (17.5)
HDENS	Horizontal density*	126 (90)	73 (64)	108 (123)	107 (102)
VDENS	Vertical density*	9.3 (5.8)	6.8 (3.5)	5.4 (3.9)	7.2 (5.0)
DSMALL	Less than the median value, i.e. less than 18,550 inhabitants **	38.7	49.2	61.4	49.9
DLWPOL	Left wing politics **	34.0	33.3	20.9	28.7

* Average value (standard deviation in bracket)
** Since the variable is a dummy, the reported average values represent the percentage of municipalities with this characteristic.

Table 1c. Model and variables

Sub-model 1: *Stochastic cost frontier* *(Translog)*	*Sub-model 2:* *Prior class probabilities* *(Multinomial Logit)*	*Sub-model 3:* *Inefficiency model* *(Battese-Coelli 1992)*
Dep. Var. = Total cost (TC)	Dep. Var. = $P_{ij}(\delta_j)$	Dep. Var. = Inefficiency (u_{ij})
Waste disposed (Y_D), Output Waste recycled (Y_R), Output N. of buildings (Y_B), Output Surface (S), Fixed Input Labor price (w_L), Input price Capital price (w_K), Input price	Waste per Capita (*QPC*) Share of Recycling (*SHREC*) Horizontal Density (*HDENS*) Vertical Density (*VDENS*)	Corruption ($CORR_1$, $CORR_2$, $CORR_{PS}$, $CORR_{ES}$) Municipal size (*DSMALL*) Left-wing politics (*DLWPOL*) Regional dummies (*DNORTH*, *DCENTER)*

Table 1d. Summary statistics: corruption measures

Variable	*Description*	North	Center	South	Total
$CORR_1$	Crimes against public faith per 100,000 inhabitants (ISTAT)*	5.15 (2.02)	5.74 (1.68)	5.68 (1.91)	5.49 (1.92)
	With DSMALL = 1	5.06 (1.70)	5.96 (1.36)	5.67 (1.64)	5.55 (1.63)
	With DLWPOL = 1	5.29 (2.08)	5.27 (1.81)	5.62 (1.96)	5.38 (1.98)
$CORR_2$	Corruption index based on Golden and Picci (2005)	1.10 (0.63)	1.12 (0.55)	2.15 (1.48)	1.52 (1.15)
	With DSMALL = 1	0.95 (0.54)	1.24 (0.65)	2.06 (1.30)	1.55 (1.12)
	With DLWPOL = 1	1.19 (0.64)	0.88 (0.36)	2.19 (1.64)	1.39 (1.12)
$CORR_{PS}$	High Power Syndicate**	0	0	76.8	30.1
	With DSMALL = 1	0	0	70.9	34.1
	With DLWPOL = 1	0	0	78.5	22.4
$CORR_{ES}$	High Enterprise Syndicate**	50.0	32.2	36.1	41.8
	With DSMALL = 1	45.6	34.5	33.9	37.5
	With DLWPOL = 1	59.1	19.5	39.2	43.2

* Average value (standard deviation in brackets)
** Since the variable is a dummy, the reported average values represent the percentage of municipalities with this characteristic.

Table 1e: Summary statistics: pair-wise correlations among corruption indicators

	$CORR_1$	$CORR_2$	$CORR_{PS}$	$CORR_{ES}$
$CORR_1$	1			
$CORR_2$	0.3926***	1		
$CORR_{PS}$	0.1191***	0.4115***	1	
$CORR_{ES}$	0.3031***	0.3151***	0.0215	1

*** Statistically significant at 1%.

Table 2. Estimated results of: 1) the cost frontier; 2) the prior class probabilities model ; 3) the inefficiency model

Variables		LATENT CLASS MODEL 1 class	LATENT CLASS MODEL 2 classes	
		Parameters (Standard errors in brackets)	Parameters Latent Class 1 (Standard errors in brackets)	Parameters Latent Class 2 (Standard errors in brackets)
Sub-model 1: Cost frontier				
$\ln Y_D$	β_D	0.7320 (0.0187)***	0.7149 (0.0385)***	0.7010 (0.0211)***
$\ln Y_R$	β_R	0.1909 (0.0069)***	0.1926 (0.0146)***	0.1543 (0.0094)***
$\ln Y_B$	β_B	0.1749 (0.0283)***	0.2309 (0.0606)***	0.2548 (0.0339)***
$\ln S$	β_S	-0.0936 (0.0139)***	-0.1149 (0.0277)***	-0.1021 (0.0210)***
$\ln(w_L / w_K)$	β_L	0.5823 (0.0563)***	0.3953 (0.1118)***	0.7120 (0.0723)***
$(\ln Y_D)^2$	β_{DD}	0.3441 (0.0431)***	0.3830 (0.0838)***	0.2625 (0.0389)***
$(\ln Y_R)^2$	β_{RR}	0.0710 (0.0046)***	0.0834 (0.011)***	0.0291 (0.0075)***
$(\ln Y_B)^2$	β_{BB}	0.3702 (0.1052)***	0.3824 (0.1961)**	0.1980 (0.1125)*
$(\ln S)^2$	β_{SS}	0.0497 (0.0217)**	0.0270 (0.0446)	0.1989 (0.0332)***
$(\ln w_L)^2$	β_{LL}	0.1560 (0.2273)	-0.2265 (0.4315)	0.5076 (0.4104)
$\ln Y_D \ln Y_R$	β_{DR}	-0.1001 (0.0101)***	-0.0899 (0.0188)***	-0.0937 (0.0147)***
$\ln Y_D \ln Y_B$	β_{DB}	-0.2948 (0.0613)***	-0.3226 (0.1212)***	-0.1346 (0.0522)***
$\ln Y_R \ln Y_B$	β_{RB}	0.0627 (0.0136)***	0.0489 (0.0273)*	0.0602 (0.0228)***
$\ln Y_D \ln S$	β_{DS}	0.0322 (0.0214)	0.0413 (0.0400)	-0.0516 (0.0300)*
$\ln Y_R \ln S$	β_{RS}	-0.0258 (0.0067)***	-0.0289 (0.0131)**	-0.0221 (0.0106)**
$\ln Y_B \ln S$	β_{BS}	-0.0761 (0.0317)**	-0.0635 (0.0696)	-0.0560 (0.0506)
$\ln Y_D \ln w_L$	β_{DL}	0.0071 (0.0941)	-0.0529 (0.1703)	0.1714 (0.0982)*
$\ln Y_R \ln w_L$	β_{RL}	0.0379 (0.0297)	0.0187 (0.0592)	0.0044 (0.0445)
$\ln Y_B \ln w_L$	β_{BL}	0.0845 (0.1573)	0.4555 (0.2924)	-0.3526 (0.1603)**
$\ln S \ln w_L$	β_{SL}	-0.0765 (0.0771)	-0.2594 (0.1293)**	0.1030 (0.0759)
Constant	β_0	-0.7193 (0.0183)***	-0.8158 (0.0499)***	-0.3791 (0.0217)***
Sub-model 2: Prior class probabilities(1)				
Constant	δ_0	-	2.2126 (0.6141)***	0
HDENS	δ_1	-	0.0025 (0.0015)*	0
VDENS	δ_2	-	-0.1046 (0.0316)***	0
SHREC	δ_3	-	-2.7309 (0.8151)***	0
QPC	δ_4	-	-2.2236 (1.0409)**	0
Prior probabilities at data means (%)		100.00	53.42	46.58
Sub-model 3: Inefficiency term				
Scale factor in time-varying inefficiency	η	-0.0350 (0.0054)***	-0.0692 (0.0109)***	-0.0050 (0.0081)
Variables in mean of truncated distribution:				
Corruption (*CORR₁*)	α_1	0.0896 (0.0041)***	0.1007 (0.0096)***	0.0178 (0.0142)
Average efficiency score(2)		0.594 (0.156)	0.683 (0.181)	

Notes: *** Statistically significant at 1%; ** statistically significant at 5%; * statistically significant at 10%

(1) Parameters for latent class 1 represent the differential impact of each factor on the probability to be assigned in class 1 rather than class 2 (for this reason the parameters of latent class 2 are set to 0).

(2) The efficiency score measures the relative performance of each observation with respect to the cost frontier (maximum efficiency =1): the higher the distance from the frontier, the lower the efficiency.

Table 3. Sample breakdown by class membership based on posterior probabilities

Variable	*Latent class 1*	*Latent class 2*
Number of municipalities	261	268
SHREC	15.2 (15.8)	24.2 (18.0)
HDENS	107 (112)	107 (90)
VDENS	5.9 (3.7)	8.5 (5.7)
QPC	457 (125)	483 (158)
DNORTH	25.3	51.5
DCENTER	21.5	23.1
DSMALL	57.9	42.2
DLWPOL	20.8	36.4
$CORR_1$	5.58 (1.62)	5.41 (1.99)
$CORR_2$	1.64 (1.26)	1.41 (1.04)
$CORR_{PS}$	39.1	21.3
$CORR_{ES}$	38.7	44.8
Average efficiency score	0.596 (0.171)	0.767 (0.147)
Posterior probabilities (%)	92.24	84.44

Table 4. Sensitivity analysis: alternative models for inefficiency (sub-model 3)

		Model I	Model II	Model III	Model IV
			CLASS 1	CLASS 1	CLASS 1
Scale factor in time-varying inefficiency	η	-0.0346***	-0.0779***	-0.0864***	-0.0750***
		(0.0053)	(0.0128)	(0.0143)	(0.0117)
Variables in mean of truncated distribution					
$CORR_1$	α_1	0.0802***	0.0908***		
		(0.0055)	(0.0132)		
$CORR_2$	α_2			0.1156***	
				(0.0415)	
$CORR_{ES}$	α_{ES}				0.1059
					(0.0911)
$CORR_{PS}$	α_{PS}				0.2799***
					(0.1067)
DNORTH	α_{NORTH}	0.0094	-0.2734**	-0.2264	-0.0833
		(0.0457)	(0.1089)	(0.1701)	(0.1562)
DCENTER	α_{CENTRE}	-0.0125	-0.0441	0.1142	0.1784
		(0.0525)	(0.0901)	(0.1242)	(0.1337)
DSMALL	α_{SMALL}	0.0659	0.0846	0.2775***	0.1738*
		(0.0475)	(0.0900)	(0.0996)	(0.1003)
DLWPOL	α_{LWPOL}	0.1543***	0.2078**	0.3410***	0.3375***
		(0.0437)	(0.0864)	(0.1130)	(0.1036)
			CLASS 2	CLASS 2	CLASS 2
Scale factor in time-varying inefficiency	η		-0.0074	-0.0081	-0.0089
			(0.0063)	(0.0062)	(0.0076)
Variables in mean of truncated distribution					
$CORR_1$	α_1		0.0495***		
			(0.0096)		
$CORR_2$	α_{R2}			0.1075***	
				(0.0240)	
$CORR_{ES}$	α_{ES}				0.1759*
					(0.0906)
$CORR_{PS}$	α_{PS}				0.1733
					(0.1221)
DNORTH	α_{NORTH}		-0.0102	0.1410**	0.1081
			(0.0750)	(0.0637)	(0.1167)
DCENTER	α_{CENTRE}		-0.1348	0.0232	0.1421
			(0.0984)	(0.0873)	(0.1203)
DSMALL	α_{SMALL}		-0.1631*	-0.1287	-0.2819**
			(0.0918)	(0.0807)	(0.1179)
DLWPOL	α_{LWPOL}		0.0441	0.0589	0.0227
			(0.0737)	(0.0664)	(0.0818)

Notes: *** Statistically significant at 1%; ** statistically significant at 5%; * statistically significant at 10%
Model I has only 1 latent class. In the other models, Class 1 is on average composed of municipalities with a lower share of recycling.