The Costs of Corruption in the Italian Solid Waste Industry

Graziano Abrate

(University of Eastern Piedmont and HERMES)

Fabrizio Erbetta

(University of Eastern Piedmont, Ceris-CNR and HERMES)

Giovanni Fraquelli

(University of Eastern Piedmont, Ceris-CNR and HERMES)

Davide Vannoni*

(University of Torino and Collegio Carlo Alberto)

Abstract

The paper investigates the link between corruption and efficiency by using a rich micro-level dataset concerning solid waste collection activities in 529 Italian municipalities observed over the years 2004-2006. In order to test the impact of corruption on cost efficiency we estimate a stochastic latent class frontier approach, which accounts for technological heterogeneity across units. The results of our estimates show that corruption significantly increases inefficiency, a finding which is robust to the inclusion of alternative local corruption indicators and of 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 for those municipalities which 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 and Statistics, Corso Unione Sovietica 218 bis, 10124 Torino – Italy. Tel: + 39 + 0116706083. E-mail: vannoni@econ.unito.it

1. Introduction

According to a recent report of Transparency International (2010), corruption levels around the world are seen as increasing over the last years, despite institutions such as OECD and EU are declaring an increasing effort in fighting corruption. On June 2011, the EU commission has roughly estimated that corruption implies a cost for the EU economy of about €120 billion each year, i.e. 1% of the EU's GDP. As to Italy, latest data on the corruption perception index classify the country among the worst positions within the EU, and the Court of Audit has estimated a cost of €60 billion each year. While the above figures are purely anecdotal, they are a clear signal of the huge impact that corruption may exert on the economy of a country.

The scientific community too has devoted a considerable research effort to understand 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 only very few exceptions, these studies have provided convincing evidence that corrupt institutional and social environments exert a negative impact on economic activity and growth: corruption is found to be harmful for investment (Mauro, 1995), for the attraction of foreign direct investments (Wei, 2000; Zurawicki and Habib, 2010), for the productivity of the capital stock (Lambsdorff, 2003), and for technological change (Salinas-Jiménez and Salinas-Jiménez, 2007).

However, some recent studies considered the possibility that corruption may operate efficiently in situations where the institutional framework is weak. Banerjee *et al*. (2006) pointed out that foreign private investors may select economic environments where it is possible to acquire a certain degree of stability through bribery. Méon and Weill (2010) found evidence that corruption is less detrimental, as far as the aggregate productivity of labour is concerned, in countries where bureaucracy is slow and cumbersome and therefore more prone to be greased. Halkos and Tzeremes (2010) found a U-shaped relationship between corruption and economic efficiency, suggesting that an increase in corruption beyond a certain level can impact positively on economic performance.

Since the large bulk of the literature made use of aggregate country level data, several authors have highlighted the need to undertake more fine grained investigations of the relationship between corruption and the different dimensions of performance. For example, Svensson (2005) underlines the importance of enriching the micro-evidence as follows: "*The link between the macro literature on how institutions provide a more-or-less fertile breeding ground for corruption and the micro literature on how much corruption actually occurs in specific contexts is weak. 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) outlined the crucial role of firm-level investigations. Such studies, indeed, 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).

To date, indeed, there are very few papers using firm level data. The latter, which are essentially focused on public utilities, analyse corruption alone or in connection with reforms such as the establishment of independent regulation authorities and the opening to private capital investments. For instance, Dal Bò and Rossi (2007) estimated a labour requirement function on a set of 80 electricity distribution firms active in 13 Latin America countries, and showed that firms operating in more corrupt environments tended to be less efficient in terms of labour use. Estache and Rossi (2008) and Wren-Lewis (2011) extended the above 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 activity of well-governed regulatory authorities[1].

The above studies used firm-level data to compute productivity and other performance measures, but they are still confined within a cross-country framework and, most importantly, they use country-level indices of corruption. The latter exploit data on perceived level of corruption within countries and, by relying upon expert assessments and opinion surveys, may well be affected by potential distortions (Reinikka and Svensson, 2005). In contrast, a single country-firm level study is provided by Yan and Oum (2011), who investigated the effect of corruption on a sample of 55 US commercial airports observed from 2001 to 2009. The airports can be managed through three different institutional arrangements (i.e. they can be directly operated by local government branches of indirectly via airport or port authorities). The authors found that the detrimental effect of corruption on efficiency is confirmed only in the case of airports operated by airport authorities, while for the other two categories corruption seems to have no discernible effects. Moreover, airport authorities were found to be more efficient in managing airports only in less corrupt environments. Therefore, the airport authority, which was in principle better suited to pursue productivity improvements and self-financing goals, turned out to be also the institutional arrangement which was the most exposed to corruption, which acted to divert the managerial efforts from production control towards unproductive rent-seeking activities.

Yaun and Oum (2011) used a state level measure of corruption, notably the number of government officials convicted for corrupt practices for each of the 50 US states. In their words: *"Compared to*

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

the country level corruption index used by Dal Bò and Rossi (2007), this state-level corruption measure is more objective because country-level corruption index is constructed via subjective survey evidence" (p.121). In this paper we follow an approach similar to Yaun and Oum (2011), since, on the one hand, we use disaggregated date at the municipality level to estimate efficiency and, on the other hand, we relate inefficiency levels to "more objective" local measures of corruption.

The scant literature which makes use of firm-level data unequivocally highlights the presence of a negative impact of corruption on efficiency. In order to accurately interpret such a result, it is important to understand the way the inefficiency term has been modelled. Dal Bò and Rossi (2007), Estache and Rossi (2008) and Wren-Lewis (2011) considered as inefficiency any increase in labour use due to environmental factors, given outputs and the adopted technology. In fact, they did not use a model, such as the stochastic frontier, capable to directly estimate inefficiency, but tested the effects of corruption in a comparative statics framework. On the contrary, Estache and Kouassi (2002) estimated a stochastic production frontier, and introduced a second stage to regress inefficiency on a set of variables, among which corruption. Yan and Oum (2011) defined a model where a set of institutional variables is used in a cost frontier specification to capture excess cost (interpreted as technical inefficiency) and input allocative distortions, but they did not provide an estimate of firm-specific inefficiency levels.

It is also important to note how individual heterogeneity has been treated. As outlined in Greene (2005a and b) heterogeneity may exert a strong effect on cost differentials and, therefore, should be properly modelled in order to avoid biased estimates. Moreover, since heterogeneity can be only partially summarized in one or more variables, the way commonly pursued to capture it is to introduce individual effects in the estimation[2]. Despite this, firm-specific effects were explicitly modelled only in two cases (Estache and Rossi, 2008; Wren-Lewis, 2011), whereas Dal Bò and Rossi (2007) included in their empirical model country-specific effects only.

These introductory remarks show that the evidence on the relationship between corruption and performance needs further investigation, especially for what it concerns the use of micro-data and measures of corruption at the local level, as well as the use of a methodology able to quantify the impact of corruption on cost efficiency, adjusted to take into account individual heterogeneity. In this sense, this paper aims to contribute to the above literature in two ways. First, it seeks to provide new evidence on the relationship between corruption and efficiency in the – so far unexplored –

[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 to access data and by the objective inability to collect a full and comprehensive information set. Panel data techniques based on the introduction of fixed or random effects have been therefore developed, so as to account for the impact of time-invariant individual characteristics.

solid waste industry, using a micro level data sample derived from a large set of Italian municipalities. As for the case of Italy, recent anecdotal evidence has highlighted the phenomenon of corruption in the collection and especially in the disposal of waste. In addition, as the report on "Ecomafia" revealed (Legambiente, 2011), in the context of the environmental crimes, corrupt private companies, local authorities and supervisory bodies often interact and give rise to illegal networks that undermine the correct management of waste collection and disposal. For example, the networks exert a substantial influence over the system of contracts and subcontracts through which the processes of collection, transportation and disposal of waste are managed, or organize extortion activities. Secondly, the paper uses a latent class cost frontier specification (Orea and Kumbhakar, 2004; Greene, 2005a) which allows, on the one hand, to model inefficiency as a function of a set of variables, among which corruption, and, on the other hand, to account for the presence of technological heterogeneity. This approach seems particularly suitable in the context of the waste industry, since the different levels of recycling activities achieved in various areas of the country, even under the impulse of local political institutions, could represent a technological factor that can impact significantly on costs, regardless of mere efficiency reasons.

The remainder of this paper is organized as follows. Section 2 explains more in details, drawing from the existing literature, the mechanisms through which corruption may impact on efficiency. Section 3 illustrates the latent class cost frontier approach. Section 4 presents the specification of the empirical model and describes the dataset, highlighting in particular the main characteristics of the waste industry in Italy and the way 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: theoretical background

The literature on heterogeneity separates the effects on firms' cost due to environmental factors which are beyond managerial control from those due to the lack of appropriate incentive schemes or to inadequacy in managing productive processes. Some exogenous environmental factors, however, may still have an impact on the structure of incentives and on the ways in which they are transferred to managers. As argued by Olson *et al*. (2000), for instance, country-specific characteristics of governance, such as quality of bureaucracy, diffusion of bribes and rule of law, may alter the incentives inherent in policy regime and institutions, therefore limiting the attainable economic performance.

In general, the empirical studies support the evidence that corruption can harm firms' productivity by distorting the use of available resources. Borrowing from production theory, a firm may be interpreted as an economic agent who uses traditional inputs and managerial effort to coordinate them. Any exogenous factor that prevents to exert this effort is hurting efficiency. To that regard, Yan and Oum (2011) and Dal Bò and Rossi (2007) provided analytical models aimed at explaining the channels through which corruption may actually divert managerial effort. Basically, both studies build on the idea that corruption leads to weak incentives and therefore to low efficiency levels, but they are different as to the underlying mechanisms (*external* versus *internal*) at stake.

Yan and Oum (2011) observe that in a more corrupt environment, policy-makers and bureaucrats tend to reduce the accountability of public policy-making, so as to be in a better position to extract some private benefits. In such a context, characterized by high opacity about the results of political decision-making processes, the advantage for governments trying to push public services 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 receive the appropriate incentives, will be less sensitive to performance and more keen to follow strategies dictated by the personal agenda of their managers. In this perspective, the underlying assumption is that the diversion of managerial effort depends on an external factor such as the interruption of flows of incentives along a chain of control.

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

Both these arguments could be well applied to the case of the waste industry. In Italy, waste collection and disposal are mainly carried out by publicly-owned firms under the control of local governments and, ultimately, of citizens. Although the latter should be, in principle, interested in an efficient management of mandated tasks, due to the impact this would have on the tax burden, the assumptions that they have complete information about the technology and that they are able to make an informed assessment of the economic performance seem quite unreasonable. This is especially true in contexts plagued by situations of widespread corruption and entrenched presence of criminal organizations. As discussed in D'Amato *et al*. (2010), the entry of organized crime in the waste cycle is mainly aimed at creating shadow circuits for the illegal transport and disposal. In this context, the diffusion of collusive relationships among managers and suppliers aimed at overcharging the firms and at seeking illegal sources of profits is an uncontested matter of fact. Also, in more corrupt environments, managers are more likely to engage in negotiating activities with local governments in order to establish more favourable tariffs and service obligations, thereby diverting the managerial efforts away from cost monitoring and productive tasks.

Despite the theoretical predictions that corruption implies a general slackness of the ability to control costs, the empirical evidence aimed at measuring the direct impact of corruption on overall cost efficiency is very limited, as discussed in the previous section. Among the cross-country studies, there are very few examples of estimation of production frontiers (Salinas-Jiménez and Salinas-Jiménez, 2007; Halkos and Tzeremes, 2010; Méon and Weill, 2010) while, among the few studies that made use of firm level data, only Yan and Oum (2011) estimated a cost function frontier. Similar to Yan and Oum (2011), the approach followed in this paper enables us to measure the direct impact of corruption on overall cost efficiency, controlling at the same time for the presence of latent heterogeneity.

3. Latent class cost frontier model

Our stochastic cost frontier approach assumes that producers are facing input prices and are seeking 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 in order to adapt stochastic frontier approaches to 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 looks as follows:

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

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 Eq. [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 producers' 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, in this case, all factors that do not change over time – including any persistent sources of inefficiency characterising the operational routines – are captured by the estimated individual effects, efficiency turns out to be overestimated. However, Greene (2005a) and Orea and Kumbhakar (2004) proposed to tackle 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 refers to such classes, rather than being associated to the individual units. This would seem particularly useful in those situations, such as the case of solid waste, where the heterogeneity follows a discrete distribution. In fact, it can be reasonably expected that only sufficiently large variations in the percentages of waste sent for recycling are able to determine differences in collection and transport systems, and then, ultimately, in technology.

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

The LCSFM specification is as follows:

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

where j (j = 1,…, J) denotes the class, β_j is a vector of class-specific parameters which 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 Eq. [3]:

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

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

where η_j, based on Battese and Coelli's (1992) specification, are parameters that reflect the impact of time within 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} may be seen as dependent on a set of explanatory time-invariant (and specific for each unit) variables, z_i, and a vector of class-specific parameters α_j, as follows:

$$\mu_{ij} = \alpha_j' z_i \tag{4}$$

The presence of a set of time-invariant variables (among which corruption), explaining the persistent part of inefficiency, reflects 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 the cost efficiency.

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

$$P_{ij}(\delta_j) = \frac{\exp(\delta_j' q_i)}{\sum_{j=1}^{J} \exp(\delta_j' q_i)} \tag{5}$$

where q_i is a vector of time-invariant characteristics, for each producer, potentially able to influence such probabilities and δ_j are class-specific parameters to be estimated. More specifically, any change in variables z_i with respect to the sample mean is expected to increase or decrease the probability to belong 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 to belong to the J classes, the individual likelihood function may be, finally, expressed as a weighted average depending not only on parameters β but also on parameters δ:

$$LF_i(\beta, \delta) = \sum_{j=1}^{J} LF_{ij}(\beta_j) P_{ij}(\delta_j) \tag{6}$$

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, it can be written as:

$$P(j,i) = \frac{LF_{ij}(\beta_j) P_{ij}(\delta_j)}{\sum_{j=1}^{J} LF_{ij}(\beta_j) P_{ij}(\delta_j)} \tag{7}$$

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 Eq. (6), is obtained as:

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

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 may be obtained as:

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

where reference technology of every class is taken into account.

4. Data description and model specification

The Italian Municipal Solid Waste (MSW) industry exhibits several characteristics making it an ideal field of application for testing the existence of a relation between corruption and inefficiency. First, the heterogeneous provincial levels of corruption (see Section 4.2) mirror the local nature of the industry. Indeed, despite recent reforms were aimed at discouraging in-house provision by municipalities and at favoring competitive tendering procedures, a strict relation between operators and municipalities remains. As a result, the large majority of 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.[4] Finally, some recent cases of bad MSW management have emerged as topical. The media have widely reported on the recent waste crisis in Naples and the

[4] The information concerning the costs as well as the amount of waste collected were gathered from annual MUDs (i.e. annual declarations concerning municipal solid waste collection).

surrounding, while a number of books and movies have clearly informed about the connection between waste management and illegal practices.[5]

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 multi-product nature of MSW service has been described using three indicators: tons of MSW disposed (Y_D), tons of MSW sent for recycling (Y_R) and number of buildings (Y_B), which proxies the number of collection points. Moreover, given the congestion problems that may arise in MSW management (especially in terms of finding disposal sites), we also account for the municipality surface (S), which acts as a constraint linked to the availability of land for new urbanization. As to the inputs, the disaggregation was limited to labor and capital, also because of difficulties in getting data (see Antonioli and Filippini, 2002, for a similar approach). In fact, input prices have been 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 (P_L) is given by the ratio of total salary expenses to the number of full-time equivalent employees. Capital price (P_K) is obtained by dividing depreciation costs by the capital stock.

Following Christensen *et al.* (1973), we use a flexible translog functional form, which is quite common in empirical studies of cost and production functions.[6] The empirical cost frontier, for class j, takes the following expression (where i denotes the municipality and $t = 1,\dots, T$ time):

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

and

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

[5] Among the latter, the best-seller book *Gomorrah* (Saviano, 2006) reached a large international audience and contributed to sensitize the public opinion against the plague of environmental crimes.

[6] However, as discussed by Abrate et al. (2011b), most empirical studies on the costs of refuse collection are still based on very simple specifications, such as the Cobb-Douglas model.

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 considered.

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

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

The latent class approach has the advantage of estimating separate technological parameters for different groups of observations, by specifying a given number of classes. In our study we used a two-class specification and we compare it to a model with a single class.[7] Groups are determined through the modeling process, rather than being forced into predefined categories, and a set of determinants (q_i) of latent class *prior* probabilities were considered, in order to capture any regularity in the definition of the classes.

Table 1b describes, for different geographical areas, some key characteristics that may represent 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 a great variability in the sample, especially in terms of policies for waste collection. In fact, while some municipalities had not yet started, over the years in question, to undertake a serious recycling program, others had already reached important targets in this direction, with a maximum share of almost 75 percent. Moreover, a remarkable gap exists between North (with more than 36 percent of recycling on average) and the other regions (only 7.1 percent in the South). Since increasing the share of recycling is a political decision, we also collected data concerning the type of municipal government during the observed period. 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, it tends to be associated with a slightly higher recycling share (on average 22.2 percent against 18.7 percent), which may indicate a greater attention to environmental policies.

[7] According to Orea and Kumbhakar (2004), goodness of fit measures like AIC and BIC can be used in case of ambiguity about the appropriate number of classes. The adoption of these criteria led us to focus on the two-class specification. In addition, estimation with more than two classes failed to converge, thus suggesting that the discrimination ability of the model can lead to, at best, a clear identification of two classes.

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

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

In Eq. [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 is characterized by a lower presence of relatively 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 such a region. On the contrary, *HDENS*, as a measure of horizontal congestion, is less correlated with the presence of small and large municipalities, and a higher value of this indicator could imply more severe constraints on land use (with consequences, for example, in terms of 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:

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

from which it is easy to note 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*). In order to test the robustness of the effect of corruption, we used as control variables the dummy variables representing small municipalities (*DSMALL*) and municipalities ruled by left-wing political parties (*DLWPOL*), together with two geographical dummies (*DNORTH* and *DCENTRE*). More specifically, the time-invariant cost inefficiency function for class *j* is modeled as follows:

$$\begin{aligned} \mu_{ij} &= \alpha_{rj} CORR_i + \alpha_{NORTHj} DNORTH_i + \alpha_{CENTREj} DCENTRE_i \\ &+ \alpha_{SMALLj} DSMALL_i + \alpha_{LWPOLj} DLWPOL_i \end{aligned} \tag{14}$$

4.2. Measures of corruption

A crucial point of the analysis clearly concerns the complex assessment of corruption. We followed 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 the economic literature (Golden and Picci, 2005), and a third index was computed borrowing from the sociological literature (Block, 1980). All variables have been defined at the provincial level, and the presence of around 100 provinces ensures a high diversification of the indexes across municipalities.[8]

Accordingly, $CORR_1$ is represented by 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 other crimes against the faith and public order. As expected, given the presence in the South of a public policy and a social system greatly penalized by the action of powerful and long-standing criminal organizations, the value is higher than the one recorded in the North, even if the difference is not big and is associated with a higher standard deviation (see Table 1c). On average, the level of corruption does not seem to be affected by the size of operations, nonetheless there are some different territorial tendencies: in particular, small municipalities are worst in the Center, while they appear to be slightly better than large ones in the northern areas. As to the political side, there is a slightly lower level of corruption in municipalities ruled by left-wing parties, though this evidence is confirmed only in the case of the Center.

$CORR_2$ has been proposed and computed by Golden and Picci (2005). The measure, which is available for all Italian provinces, 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*" (p. 37). The underlying idea is that corruption raises the costs of building public infrastructures, so that $CORR_2$ is computed as the ratio between the building cost and the actual value of public investment. From Table 1c we can see that in the South this measure takes on values which are almost double with respect to the other regions. As to the differences associated with size and politics, they are very much in line with those emerged from $CORR_1$. Indeed, the pair-wise correlation between $CORR_1$ and $CORR_2$ is equal to 0.39 and is strongly significant (Table 1d).

Finally, building on the sociological approach, we used two indexes ($CORR_{ES}$ and $CORR_{PS}$), which are more strictly linked to the societal penetration of organized crime.[9] Block (1980), analyzing the case of New York, distinguished between two types of criminal organizations, the first (*enterprise syndicate*) operating in the economic and enterprise networks managing illegal affairs, while the

---

[8] In Italy, a province is an administrative division of intermediate level between a municipality and a region. A province is composed of many municipalities, and usually several provinces form a region.

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

second (*power syndicate*) aiming at the 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 report by Fondazione RES (2011) provides alternative measures of Italian criminality, disaggregated at the provincial level. The power syndicate index is based on 5 types of 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 the enterprise syndicates, other 5 types of crime were considered: criminal conspiracy, drug trafficking, exploitation of prostitution, robbery and usury. With reference to each indicator, a ratio was calculated between provincial and national data. Thus, a ratio greater than 1 for a province would indicate a criminal activity higher than the one prevailing at the national level. Accordingly, two dummy variables have been 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 1c, power syndicate is highly present in southern municipalities, especially in the 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, with prevalence in the North. It is also interesting to notice that the enterprise syndicate in small municipalities does not vary so much across geographical areas, while it appears to be stronger (less strong) in northern (central) Italy municipalities ruled by left-wing political parties. Finally, Table 1d indicates the absence of significant correlation between power and economic syndicate, confirming that they refer to different aspects of criminality; conversely, both measures are significantly correlated with the alternative corruption indicators used in the paper ($CORR_1$ and $CORR_2$).

5. Results

As discussed in the previous sections, our empirical strategy requires to simultaneously estimating three types of class-conditional parameters referred to the stochastic cost frontier model, the prior probability model and the cost inefficiency model. There are three main key research questions we plan to test:

- Is the underlying technological heterogeneity actually acting in the direction of identifying separate classes?
- Are such classes polarized on the base of output composition, and, in particular, according to the incidence of recycling activities?
- Is cost inefficiency, measured as the residual from the stochastic cost frontier of each class, differentially affected by corruption?

Table 2 compares the results on the stochastic cost frontier parameters for the baseline model with two latent classes (columns 2 and 3) with the ones stemming from a model which assumes a unique class (column 1).[10] All first order parameters are strongly significant and have the expected sign and, given the log-log transformation and the data normalization, can be interpreted as cost elasticities for the "average" municipality. The coefficient associated to the relative price of labor (β_L) can be interpreted also as the cost share of labor and is equal, on average, to 0.58 (Column 1). However, Class 1 has a significantly lower labor share (0.40) than Class 2 (0.71). From Column 1, the cost output elasticities (ε) equal to $\beta_D = \varepsilon_{CY_D} = 0.73$, $\beta_R = \varepsilon_{CY_R} = 0.19$ and $\beta_B = \varepsilon_{CY_B} = 0.17$, respectively. Moreover, the negative sign of β_{DR} suggests the presence of scope economies between disposal and recycling activities.[11] Returns to scale (RTS) can be easily computed for the average municipality using the following formula: RTS = $1/(\beta_D + \beta_R + \beta_B)$. All three columns show evidence of diseconomies of scale (RTS ranges from 0.88 to 0.91), suggesting that costs increase more than proportionally with respect to increases in the amount of waste disposed, of waste sent for recycling, as well as in the number of buildings in a given area. It is worth noting the negative sign associated with the surface, which may be interpreted as a fixed factor acting as a constraint, in terms of land availability, for each municipality. Even if the surface represents an environment constraint which cannot be modified in the long-run, in principle it may be possible to think of municipality mergers, whereby S increases along with Y_D, Y_R and Y_B. Accepting the latter hypothesis, and including the coefficient β_S in the computation of scale economies, returns to scale turns out to be constant, suggesting that the increase of the area size partially counteracts the effect of a disproportionate rise in refusal collection costs that occurs when Y_D, Y_R and Y_B increase in congested areas.

For each observation, it is possible to measure the relative performance, in terms of efficiency, with respect to the cost frontier (maximum efficiency =1): the higher the distance from the frontier, the lower is efficiency. 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). This is due to the better ability of the two-class model to account for technological heterogeneity, thereby providing a more reliable measurement of efficiency.

Table 3a shows the results of the multinomial logit model represented by Eq. [13], that aims at disentangling the determinants of prior probabilities, while Table 3b describes the main characteristics of the two classes identified according to posterior probabilities. All variables

[10] The baseline model corresponds to the case in which $CORR_1$ has been included as the unique determinant of inefficiency. Subsequently, we will present in table 4 the results of regressions in which variants of the inefficiency model have been tested.

[11] Since the focus of this paper is about the relationship between corruption and cost efficiency, we do not discuss the technology of the solid waste industry into depth. The interested reader may refer to Abrate *et al*. (2011b).

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 exhibiting a low per-capita waste production, and a low (high) degree of vertical (horizontal) urbanization. A hypothetical municipality with average values with respect to all these four variables would have a slightly higher prior probability to belong to Class 1 (around 53 percent). However, any deviation from such average values affects the class membership probabilities.

In Table 3b we can observe the polarization between classes with respect to several variables (among which the ones 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, a higher recycling share is confirmed to be a distinctive feature of Class 2. The latter is also embracing 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 an almost double presence of municipalities ruled by left-wing parties, which usually tend to declare more explicit "environmental-oriented" goals. As to the corruption measures, it is interesting to notice 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. On 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, i.e. the relationship between corruption and efficiency. Table 2 shows an average efficiency score for the baseline model of 0.683, while Table 3b highlights that there is a 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 reveals a large potential for efficiency gains especially for Class 1 (i.e. the class less virtuous in implementing recycling programs). Moreover, looking at Table 2, we can notice that both classes have negative η-parameters, which imply an increase of inefficiency over time; however, the trend is statistically significant only for Class 1. As to structural inefficiency, the bottom part of Table 2, that reports the estimates of Eq. [4], shows that the corruption index exerts a positive impact on inefficiency, though, again, the effect is statistically significant only for Class 1. This suggests that the more recycling-oriented Class 2 seems to be more virtuous also 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), one would get another 14 percent increase.[12]

Table 4 tests several alternative variants of the cost inefficiency model [14].[13] Several control variables relative to the geographical area, the municipality size and the political orientation have been added to $CORR_1$. In fact, one major concern could be that the highlighted effect of corruption on inefficiency is actually hiding 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.[14] We also run some simple simulation in order to estimate a monetary value for the cost inefficiency due to corruption. By using 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 amounting to around €92 million per year. This accounts for around 4.7 percent of refuse collection costs, and corresponds to about €4.2 for 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).[15]

The control variables are often not significant, suggesting that corruption acts indeed as the most important driver of inefficiency. There is however some evidence of higher inefficiency associated with left-wing politics (especially for Class 1). This result may be explained by the fact that left-wing parties are usually more oriented to pursue equity and environmental goals, rather than efficiency gains.

Finally, we tested for the effect of alternative corruption indexes, while keeping 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, but the magnitude of the effect turns out to be substantially similar for both sub-groups. Moving from the first quantile (0.81) to the median value (1.30) and to the 3rd

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

[13] We decided to present only the results for the inefficiency model, since the cost frontier parameters as well as the class characteristics are very similar to the one 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 membership is very limited (from 7.2 to 11.3 percent depending on the pair-wise comparisons). The full set of estimates is available upon request.

[14] To correctly compare the effects in the two groups we should actually take into account eventual differences in the distribution of $CORR_1$. However, Table 3b shows similar means 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 represents a fairly good approximation.

[15] The percentages of cost reduction are computed with reference to the estimated cost frontiers. For what concerns the specific figures given for the cases of 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).

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 exhibit a positive sign though, interestingly, the power syndicate is significant only in Class 1 while the enterprise syndicate is significant in Class 2.[16] Even if this polarized result must be taken with caution, a possible interpretation is the following. In municipalities where recycling shares are low – mainly distributed in the South – efficiency may be more seriously harmed by the presence of power syndicate, which can gain control of disposal sites and force enterprise decisions through extortions. Conversely, in municipalities that recycle more – mainly distributed in the North – the 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 usually perceived as a "social disorder" phenomenon that can negatively impact on growth and performance of nations, empirical studies which try to measure the extent of corruption and to quantify its impact on several dimensions of performance should be welcome.

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

Our paper aims at contributing to this strand of literature by using a stochastic latent class cost frontier approach to evaluate the impact of corruption on cost efficiency in the provision of solid waste services of a sample of Italian municipalities. To that respect, the refusal collection industry, according to some recent anecdotal and judicial evidence, seems to represent a suitable case study.

Our approach is based on the measurement of efficiency as a 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 site or incinerated and the volume of waste sent for recycling as separate output variables, and we included among the regressors the number of collection points and the extension of 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 depict two well-behaved cost frontiers that actually differ in terms of the underlying technologies, validating

[16] In this case, since the indicators are dummies, the coefficients directly give the percentage difference of inefficiency between the municipalities characterised by higher than average corruption and those characterised by lower than average corruption.

our choice to rely on a latent class stochastic frontier approach. Secondly, such a differentiation depends on density variables and, as expected, on the share of waste sent for recycling. More specifically, we observe that the class with a higher presence of recycling programs (Class 2) is more labor-intensive, less sensitive to the effect of increasing costs due to demographic pressure (i.e. with milder diseconomies of scale with respect to Class 1) and denotes a higher share of left-wing political orientation. Thirdly, and most importantly, corruption is found to have a negative and significant effect on cost efficiency, thus confirming previous theoretical and empirical evidence. As a novel contribution in the literature, which very rarely has used 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 in particular appears to be less important for Class 2, which has a higher share of municipalities localized in the North and of municipalities involved in important recycling programs. The cost efficiency of such a class, however, is negatively affected by the index $CORR_{ES}$, which accounts for the presence of enterprise syndicate. On the contrary, in the Southern areas of the country, $CORR_{PS}$ seems to be more important, suggesting that illegal networks, and especially criminal organizations more involved in the control of the territory, find it easier to infiltrate in processes of collection and disposal. In this sense, the possibility to manage illegal landfills clearly plays a crucial role.

From a policy standpoint, our results show that fighting corruption and criminal organizations would bring non trivial efficiency gains. According to some simple simulations, if it would be possible reducing 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 exhibiting different average values of efficiency and resulting differently affected by corruption. Accepting the higher share of recycling as the most distinctive feature of Class 2, this suggests that pushing municipalities to be more active in pursuing recycling programs could bring efficiency improvements and reduce at the same time the adverse (in terms of cost efficiency) effects of corruption.

References

Abrate G., Erbetta, F., Fraquelli G., 2011a. 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., 2011b. The Costs of Disposal and Recycling: An Application to Italian Municipal Solid Waste Services. *Working Paper HERMES*, 1, Collegio Carlo Alberto, Moncalieri.

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.

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

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.

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., 2011. Effects of Government Quality and Institutional Choice on the Efficiency of Airports in the United States. Mimeo, University of British Columbia.

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
P_L	Price of labor (€ / Employee)	36,607	5,735	22,663	62,613
P_K	Price of capital (depreciation rate)	0.087	0.013	0.049	0.124

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

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 holding this characteristic.

Table 1c. 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 bracket)
** Since the variable is a dummy, the reported average values represent the percentage of municipalities holding this characteristic.

Table 1d: 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 parameters of the stochastic cost frontier function

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)
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)
$\ln(P_L/P_K)$	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)
$(\ln P_L)^2$	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)
$\ln Y_D \ln P_L$	β_{DL}	0.0071 (0.0941)	-0.0529 (0.1703)	0.1714* (0.0982)
$\ln Y_R \ln P_L$	β_{RL}	0.0379 (0.0297)	0.0187 (0.0592)	0.0044 (0.0445)
$\ln Y_B \ln P_L$	β_{BL}	0.0845 (0.1573)	0.4555 (0.2924)	-0.3526** (0.1603)
$\ln S \ln P_L$	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)
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 ($CORR_1$)	α_1	0.0896*** (0.0041)	0.1007*** (0.0096)	0.0178 (0.0142)
Average efficiency score[(1)] (standard deviation in brackets)		0.594 (0.156)	0.683 (0.181)	

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

(1) 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 is efficiency.

Table 3a. Multinomial Logit Model: determinants of prior probabilities

Variable		Parameters (s.e.)	
		Latent class 1	Latent class 2
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 (%)		53.42	46.58

Notes: *** Statistically significant at 1%; ** statistically significant at 5%; * statistically significant at 10%
Parameters for latent class 1 represents the differential impact of each factor on the probability to be assigned in class 1 rather than class 2 (for this reason the parameter of latent class 2 are set to 0).

Table 3b. 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

		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 1 latent class only. In the other models, Class 1 is composed, on average, by municipalities with lower share of recycling