The Efficiency of Local Public Service Production: The Impact of Size and Institutions

Peter Bönisch, Peter Haug, Annette Illy, Lukas Schreier*

Abstract

Bigger and more centralized municipal entities have been the trend in Europe for the past few decades. Expenditure reductions and efficiency gains are often expected to follow from municipal mergers, but evidence from the economic literature in this regard is far from unambiguous. Therefore, we analyze the effect of municipal size and institutional structure at the local level on efficiency in public goods production. An exceptionally detailed data set on the municipalities in the German state of Saxony-Anhalt allows us to identify the impact of different institutional settings. In contrast to previous studies, we choose the aggregate budget of municipal associations as the object of our analysis since important competences are settled at the joint administrative level. We perform a non-parametric efficiency analysis (DEA) and provide bias-corrected efficiency scores. Furthermore, we test for economies of scale and the impact of environmental variables using the bootstrap approach suggested by Simar and Wilson (2007). To our knowledge, this is the first study to use these techniques in studying this issue. The findings reveal a technology with variable returns to scale; a substantial number of observations operate under decreasing returns to scale. A centralized institutional setting is shown to reduce efficiency.

Keywords: efficiency, local institutions, local government, DEA, second stage bootstrap

JEL Classification: H11, H72
1. Introduction

The eastern part of Germany began experiencing continuous and massive amalgamations among counties, towns, and municipalities in the mid-1990s, very similar to what happened in the western part of Germany in the late 1960s and early 1970s. Indeed, over the last few decades, bigger and more centralized municipal entities has been the trend throughout Europe.

This trend has been especially pronounced in the Nordic countries. The Swedish municipal reforms of 1952 reduced the number of municipalities from 2,500 to about 1,000 and further reforms in 1974 reduced that number to 278 (Bäck 2005). Norway also reduced the number of its municipalities from about 750 in 1930 to the current 430 (Wikipedia contributors 2011). Radical local government reform in the United Kingdom about 25 years ago left that country with the largest average local government size in Europe (in terms of population). This phenomenon is not confined to Europe. For example, municipal amalgamations have taken place in Israel (Reingewertz 2010), Australia (Dollery et al. 2008), in several Canadian provinces (Kushner and Siegel 2000), and in Japan (Yokomichi n.d.). Fox and Gurley (2006) mention “government consolidation” reforms in Jordan, Sudan, and Zimbabwe.

The political goals of municipal concentration include achieving reduced service delivery costs, more equitable provision of services, and better planning (Fox and Gurley 2006). In the case of the Nordic countries, Steineke (2010) concludes: “In all Nordic countries, a central argument in promoting municipal mergers is that public welfare services are more efficiently produced in larger municipalities.”

From the perspective of fiscal federalism theory, (local) government centralization is a mixed blessing. The expected scale economies in the provision of municipal services might be cancelled out by agency problems, increased transaction costs, and fiscal illusion. The increase in average municipality size resulting from the political reforms mentioned above is often accompanied by institutional centralization, which can intensify the problems associated with bigger and more centralized municipal bureaucracies. Thus, in this paper we empirically test the impact of municipal size and institutional structure on municipal efficiency in public goods provision.

Previous studies on the efficiency of municipal service provision can be divided into two broad branches. The first branch is comprised of numerous analyses of single public services: solid waste and sewage disposal (Worthington and Dollery 2001), water (Picazo et al. 2009; Byrnes et al. 2010; Zschille et al. 2010) and energy provision (von Hirschhausen et al. 2006), hospitals (e.g., Aksezer and Benneyan 2010; Blank and Valdmanis 2010), municipal savings banks (Conrad et al. 2009; Bresler 2007), public libraries (De Witte and Geys 2009), road maintenance (Kalb 2009), fire protection (Lan 2009 et al.), care for the elderly (Borge and Haraldsvik 2009), local police services (García-Sánchez 2009), public transportation (Walter and Cullmann 2008), and pre-school education (Montén 2009;
However, a separate analysis of particular municipal services (except for those organized in separate and independent organizational units such as municipal savings banks, public utilities, or public transportation) could lead to biased results. The decision-making unit (DMU) provides a broad variety of local public goods ranging from public safety and fire protection to elementary schools, childcare facilities, and road maintenance. The main problem for efficiency analysis is the impossibility of assigning certain inputs (or expenditures) to different municipal tasks, especially inputs of central administration units, which represent overhead costs. Furthermore, the political decision-making process regarding organizational and institutional reforms at the municipal level is based on its overall effect on municipal service provision.

That is why the second branch of this literature—global municipal efficiency—is more appropriate for our undertaking. Recent work in this area covers Belgium (Geys and Moesen 2009; De Borger and Kerstens 1996), Finland (Loikkanen and Susiluoto 2005), Brazil (Sampaio de Sousa et al. 2005), Spain (Balaguer-Coll and Prior 2009; Gimenez and Prior 2007, Prieto and Zofio 2001), Portugal (Afonso and Fernandes 2008), Japan (Nijkamp and Suzuki 2009), and Germany (Kalb et al. 2012; Geys et al. 2010; Kalb 2010a; Kriese 2008; Geys et al. 2007). The main results of these papers as to the relationship between size and performance or between institutional structure and performance are rather mixed—if they even deal with these aspects. Hence, the question of whether there are scale effects and whether there are effects of decentralized organizational forms on municipal efficiency is far from answered.

In our analysis, we focus on municipal efficiency in the German state of Saxony-Anhalt. Since this state’s municipal system includes a broad variety of institutional settings at the local government level, we are able to identify the impact of certain local institutional forms on municipal efficiency. Recent studies on German municipal efficiency usually ignore the impact of different institutions. In studies on municipal service provision, this creates another methodological problem. In efficiency analysis it is of utmost importance to identify the correct decision-making unit (DMU) in order to avoid biased efficiency estimates. In the case of German local governments, it is necessary to know whether the municipality in question is a member of a municipal association, where inputs for service provision are usually commonly controlled. If such is the case, all inputs and outputs of the production process can be observed only at the level of the municipal association. This is why we calculate an aggregate budget for municipal associations, a deviation from other German studies (Kalb et al. 2012; Geys et al.

---

1 For a survey of earlier studies, see De Borger and Kerstens (2000) or Worthington and Dollery (2000). Also, Kalb (2010b) contains an extensive list of efficiency studies of different public services.


3 Geys et al. (2010), Kalb (2010a), and Kriese (2008) study municipal efficiency at the local level using data from Baden-Württemberg or Saxony. They do not control for different institutional settings at the local level, although in Baden-Württemberg, as well as in Saxony and most other German states, small communities in rural areas are obliged to form municipal associations (called Verwaltungsgemeinschaft, Verwaltungsverband, Samtgemeinde, Amtsgemeinde, etc.) to benefit from economies of scale without losing their status as independent municipalities. We discuss this crucial issue in Section 2.
Finally, most of the aforementioned studies on global municipal efficiency employ non-parametric efficiency techniques such as DEA or FDH and control for external effects on the estimated efficiency scores by using a simple two-stage approach where environmental determinants are regressed on the efficiency scores in a Tobit model. Simar and Wilson (2007) show that this procedure results in biased second-stage efficiency estimates. Thus, we account for the serial correlation of the efficiency scores by applying the two-stage bootstrapped DEA procedure suggested by Simar and Wilson (2007) to calculate non-parametric measures of global (technical) municipal efficiency and to control for the influence of certain environmental variables. Furthermore, we use a modified version of the bootstrap procedure proposed by Simar and Wilson (2002) to test for economies of scale of the production technology. To the best of our knowledge, this is the first study to employ these bootstrap techniques in an analysis of global municipal efficiency.

The paper is organized as follows. Section 2 opens with a brief overview of the institutional framework at the local level in Germany. In Section 3 we develop the theoretical background for our efficiency analysis and the second-stage parameter estimates. In Section 4 we discuss DEA and the bootstrap techniques used in our analysis. The data set is described in Section 5 and the results are discussed in Section 6. We conclude and discuss further research perspectives in Section 7.

2. Institutional Framework

Our study is based on data from the German state of Saxony-Anhalt for 2004. In this section, we describe the institutional setting relevant for our analysis. The local government level in Saxony-Anhalt is divided into 21 rural districts and three district-free towns that carry out both municipal and district tasks. The rural districts consist of 1,115 district-affiliated municipalities. Due to the great difference between district tasks and municipal tasks, we restrict our efficiency analysis to the district-affiliated municipalities.

Saxony-Anhalt is well-suited for the efficiency analysis of district-affiliated municipalities. First, the range of municipal tasks is rather homogeneous. Second, we are able to investigate the effect of three different municipal governance forms on global efficiency. On the one hand, there are independent

---

4 Furthermore, the existing studies on municipal efficiency in Germany apply highly aggregated cost data that either are not corrected for double bookings nor for expenditures that are not costs in the usual sense, or contain expenditures that do not correspond to the output indicators. The fact that the authors neglect municipal enterprises or single-purpose municipal associations in their calculations is, under some circumstances, only a minor problem if certain expenditure categories were deducted from the current expenditures. For a more detailed discussion on this issue, see Section 5.

5 There are no district-affiliated towns with a special legal status (Große Kreisstadt) carrying out district duties as, for example, in Saxony or in Baden-Württemberg.
municipalities that perform the whole spectrum of municipal tasks on their own. On the other hand, especially the smaller municipalities are obliged by municipal law to join an association with several other municipalities, but do not lose their legal status as an independent municipality by doing so. In 2004, two different forms of municipal association existed in Saxony-Anhalt. In the first type, members transfer the planning, organization, provision, and control of several municipal services to a central administration unit with a separate budget. This joint administration office acts on behalf of the member municipalities. In the second type, one (usually the largest) of the member municipalities provides own administration facilities to carry out the tasks transferred to the municipal association by its members, but without separate budgeting. In the following, we refer to the first as “Type A associations” and to the second as “Type B associations.” In both types, members make contributions or pay compensation to fund the tasks transferred to the joint administration. In contrast to the expenditure side, tax revenues as well as transfers received from the municipal fiscal equalization system remain at the member level.

Figure 1: Local government structure in Saxony-Anhalt, 2004

Source: Authors’ illustration.

---

6 In contrast to the expenditure side, tax revenues as well as transfers received from the municipal fiscal equalization system remain at the member level.
3. Municipal Size, Institutions, and Efficiency of Local Public Good Production

Our analysis has two aims. First, we provide insight into the shape of the production technology of local governments in Saxony-Anhalt. Second, we test whether different municipal institutional settings offer different incentives to local governments to reach the technological frontier.

3.1. The Production Possibility Set of Local Governments

The shape of the technological frontier of municipal service production has important policy implications since, according to Oates’s (1972) “decentralization theorem,” more centralized service provision makes sense only if it leads to lower unit costs of provision. Therefore, unit cost reductions resulting from economies of scale in the production technology may justify amalgamation. Moreover, lower costs per inhabitant—even for production technologies with constant returns to scale—could be the result of non-rivalry in consumption of the relevant good (“economies of sharing”) (for further details, see Reiter and Weichenrieder 1997). Both effects usually decrease with increasing municipal size; economies of sharing erode due to fixed resources—especially land—which have to be consumed to benefit even from pure public goods. At the same time, it is reasonable to assume that in large municipalities increasing complexity of service provision resulting from heterogeneous preferences or managerial difficulties could lead to rising unit costs.

Nevertheless, the existence of economies of scale and economies of sharing does not per se justify amalgamations of small municipalities to larger units. Gordon Tullock (1969) points out that small municipalities can benefit from economies of scale by contracting out at least some of their services—either to private suppliers, municipal enterprises, or other (local) governments. Another way of “contracting out” is the association of small, but independent municipalities. These associations could be either multi-purpose associations like the Type A or Type B described above or single-purpose units for the provision of, e.g., water or sewage disposal.

To sum up, the technology underlying the production of local public services determines whether there is a potential for increased productivity from amalgamations at the local level. Increasing returns to scale at least locally is the necessary condition for justifying further consolidation.

---

7 We have to abstract from the other two implication of Oates’s theorem, since we cannot say anything about the allocative efficiency of service provision in welfare economic terms. The more heterogeneous the preferences of the population and the less likely spillover effects, the more decentralized service provision should be.

8 Examples are the German Zweckverband or U.S. school districts.

9 However, even if our empirical analysis rejects global constant returns to scale, we do not calculate an optimal size (measured in population or area) for municipalities since there is one optimal size for each possible output mix.
3.2. Determinants of Efficiency

It is reasonable to assume that single municipalities and multi-purpose associations are faced with the same production possibility set; however, the incentives and ability of local politicians or bureaucrats to reach the production frontier could vary between institutional settings, which is why empirically different efficiency levels are possible. There are several reasons this might occur.

1. Common production or contracting out of local services can increase the transaction costs of service provision. Services that are less than hardly tangible and difficult to measure services might be especially prone to moral hazard, adverse selection, or low-service quality. Moreover, coordination costs between member municipalities increase disproportionally with an increasing number of municipalities. Therefore, municipal associations having a large number of members might not be efficient.

2. Information costs are expected to be higher for higher levels of government compared to lower levels, as well as for larger local government units compared to smaller local governments. Prohibitively high information costs may lead to over- or underprovision of certain public goods in large communities—even if benevolent politicians and bureaucrats are assumed. In this case, decentralized provision would be the better option. Since the member municipalities of associations are still legally independent (and have their own tax revenues), they may have incentives to provide the common administration with high-quality information about needs of their populations compared to districts in a larger municipality. This may lessen information costs in associations compared to single municipalities of similar size.

3. Information asymmetries in principal-agent relations between voters and the municipal council as well as between the municipal council and bureaucrats or external providers might increase the agents’ opportunities to follow their own interests. For given output quantities, politicians or bureaucrats have incentives for inefficient input employment. Vote-maximizing politicians might prefer inefficient capital use for prestigious investments (“state-of-the-art technology”) or overmanning for social reasons. Furthermore, local bureaucrats might also follow their own interests, for example, “budget maximization” (Niskanen 1971) or “slack maximization”10 (e.g., Wyckhoff 1990). It is most likely that the bureaucrats of single municipalities have more decision-making discretion than bureaucrats of municipal associations since the former are controlled by just one council, instead of by many.

Furthermore, there are likely differences between Type A and Type B associations. For example, in the former, the head of the common administration is elected by the mayors of the member municipalities, who are themselves ineligible for the office; in Type B associations,

10 “Slack” refers to the budget surplus the sponsor (the local government) is willing to finance over the actual minimum costs.
however, the head of the common administration is simply the mayor of the largest member municipality (see Kregel 2005; Bogner 2007). In the second case, the head of the administration is an elected politician and therefore is more willing to restrain the slack-maximizing tendency of bureaucracies. This may lead to more efficient service production for Type B associations compared to Type A associations, where the head of the common administration is a civil servant and is controlled only by the mayors of the member municipalities and has not been voted into office by the general population.

4. However, there are two sanctioning mechanisms that could prevent local politicians and bureaucrats from abusing their information advantages in an environment of interjurisdictional competition: exit and voice of the citizens (Hirschman 1970). It seems reasonable to assume that interjurisdictional competition is more effective between large numbers of small municipalities than between, ceteris paribus, few large municipal units: citizens’ information costs and migration costs tend to be lower in small municipalities with decentralized organization.

Most theoretical studies of the “second-generation theory” (Oates 2005; Weingast 2009) of fiscal federalism that include election processes and exit options conclude that a decentralized provision of local public goods is superior to centralized provision. In Besley and Coate’s (2003) model, the central government discriminates against certain regions as a result of so-called pork-barrel politics and the budget externalities caused by the simultaneous access of the political agents to a “common pool”. Similar suboptimal decisions in the provision of local public goods (or inefficiencies in production) are more likely in independent municipalities compared to municipal associations of equal size.

Since municipal associations have no elected common council or mayor (councils and mayors are elected only in each member municipality), it seems likely that citizens of associations can sanction inefficient service production more effectively compared to single municipalities of a comparable size. Similarly, moving within the borders of a municipal association may incur fewer costs than leaving a large single municipality.

Our theoretical considerations on the efficiency effects of the different institutional settings of local government organization are summarized in Table 1. Since the local institutions observed in the real world are multidimensional structures, it is not straightforward to derive the net effect of the different aspects discussed above. Thus, the overall effect of political institutions on the efficiency of the production of local public services remains an empirical question. However, based on the summary of our arguments in Table 1, we derive the following two hypotheses to be tested empirically:
### Table 1: Determinants of efficiency and local institutions

<table>
<thead>
<tr>
<th>Institutional setting</th>
<th>Transaction costs</th>
<th>Information costs</th>
<th>Information asymmetries</th>
<th>Costs of “exit and voice”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent municipalities</td>
<td>low (+)</td>
<td>high (-)</td>
<td>high (-)</td>
<td>high (-)</td>
</tr>
<tr>
<td>Municipal associations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type A</td>
<td>high (-)</td>
<td>low (+)</td>
<td>high (-)</td>
<td>low (+)</td>
</tr>
<tr>
<td>Type B</td>
<td>high (-)</td>
<td>low (+)</td>
<td>low (+)</td>
<td>low (+)</td>
</tr>
</tbody>
</table>

Note: Low and high refer to the strength of the particular effect depending on the institutional setting. The expected relative effect of the institutional setting on efficiency is given in parentheses.

In the first hypothesis, we assume that for independent municipalities, the information costs and the effects of information asymmetries and “exit and voice” outweigh the lower transaction costs in the production of public services at the local level.

**Hypothesis 1:** Municipal associations are more efficient than centralized local governments.

The second hypothesis has to do with the inner structure of municipal associations.

**Hypothesis 2:** Local government efficiency decreases with an increasing number of member municipalities in municipal associations.

Apart from size and institutional setting, there are several other factors that might affect municipal efficiency. To take this heterogeneity among municipalities into account, we include several control variables shown to be relevant in the literature: population density, population change, unemployment, age structure, and fiscal illusion in the form of the effect of intergovernmental grants (“flypaper effect”) and debt illusion.

Other variables could be relevant, too, such as the presence of private nonprofit organizations that offer substitutes for local public goods and services. One might also consider political variables to be relevant (significance of left-wing parties or political fragmentation). However, there are several reasons for omitting them. First, it does not make much sense to calculate an aggregated Herfindahl index or the aggregated share of left-wing seats for a municipal association comprised of many independent municipal councils. Furthermore, in small municipalities, the number of councilors without party affiliation or councilors who are members of independent voters unions is high; indeed, in Saxony-Anhalt, mayors are not even obliged to reveal their party affiliation. Thus, it is difficult or even impossible to calculate and interpret the most popular political indicators for our chosen object of analysis.

---

11 It should be noted that the presence of, e.g., kindergartens run by churches or other private nonprofit organizations, does not necessarily reduce municipal spending on those services. In fact, German municipalities are obliged to provide and finance certain services at a given quality and quantity. If these mandatory goods and services are provided by non-municipal suppliers instead, the local governments have to compensate the external suppliers.
4. Methodology

We employ data envelopment analysis (DEA) to measure the technical efficiency of German municipal associations. In this section, we briefly address some relevant methodological issues. First, we introduce the main concepts of DEA and discuss the measurement of technical efficiency when physical amounts of inputs are not available. Second, we provide a brief overview of the bootstrap algorithm proposed by Simar and Wilson (2007) for testing the impact of environmental variables on the estimated efficiency scores. Finally, we adapt the procedure suggested by Simar and Wilson (2002) to test for global and local scale efficiency.

4.1. Measuring Technical Efficiency

In our analysis, efficiency is defined as the radial distance of a municipal association (DMU) from the technological frontier. Formally, each DMU converts a vector of physical amounts of inputs \( x \in \mathbb{R}^p \) into a vector of \( q \) outputs \( y \in \mathbb{R}^q \). The underlying production process is described by the production set \( \Psi \), which is the set of all feasible points \((x, y)\). That is,

\[
\Psi = \{(x, y) \in \mathbb{R}^{p+q} \mid x \text{ can produce } y \}.
\]

For our analysis, the efficient frontier of \( \Psi \) is of interest. Technically efficient DMUs operate on the technology frontier of \( \Psi \), whereas inefficient firms operate at points in the interior of \( \Psi \). To define technical efficiency, we choose an input-oriented perspective because municipalities are expected to have more discretion in choosing their input mix, whereas the provision of certain outputs is mandatory, as discussed in Section 5. Hence, the Debreu-Farrell technical efficiency score \( \theta(x, y) \) is defined as

\[
\theta(x, y) = \inf \{\theta \mid (\theta x, y) \in \Psi\}
\]

and can be interpreted as the factor by which inputs can be reduced to achieve technical efficient production of the given outputs \( y \) (keeping the input and output mix fixed). Hence, for an efficient DMU, \( \theta(x, y) = 1 \), and otherwise \( 0 < \theta(x, y) \leq 1 \). In other words, the inputs used by inefficient DMUs are compared to the minimal amount of inputs \( x^{\text{eff}} = x \cdot \theta(x, y) \) technically necessary to produce the given amount of outputs \( y \). Since in empirical studies the true production possibility set \( \Psi \), as well as \( \theta(x, y) \) and \( x^{\text{eff}} \), are unknown we have to rely on estimators, defined below, based on the sample \( \mathcal{X}_n = \{(x_i, y_i, z_i), i = 1, \ldots, n\} \) where \( z_i \) is a vector of observed environmental variables for the \( i \)th municipality that constrain its choice of inputs and outputs.

The estimator for the input-oriented efficiency score \( \theta(x, y) \) is defined by
(3) \[
\hat{\theta}(x, y) = \min \{ \theta > 0 \mid \theta x \geq X \lambda, \ y \leq Y \lambda, \ i \lambda = 1, \ \lambda \geq 0 \}
\]
where \( Y = [y_1 \ldots y_n] \) and \( X = [x_1 \ldots x_n] \), with \( x_i, y_i \ i = 1, \ldots, n \) denoting the observed vectors of inputs and outputs, and \( \lambda \) a \( n \)-dimensional vector of intensity variables that are multiplied by a vector of ones \( i \) and, hence, sum up to one (Banker, Charnes, and Cooper 1984).

Our data set contains no information on the physical amount of inputs employed by the municipalities under consideration.\(^{12}\) However, we have access to cost information for the inputs. Since all municipalities within one state are parties to the same collective wage agreement and have access to the same capital market, we assume that the unobserved input prices are the same for all municipalities in our data set. Under these circumstances, we can replace the unobserved input quantities with the expenditures for the respective inputs and cost and technical efficiency are the same by definition (Färe and Primont 1988). Hence, we use linear programming techniques to solve the following linear program for each municipal association (and for each municipality not belonging to any municipal association) in our data set

(4) \[
\hat{\theta}(x, y) = \min \{ \theta > 0 \mid \theta c \geq C \lambda, \ y \leq Y \lambda, \ i \lambda = 1, \ \lambda \geq 0 \}
\]
where \( c \) is a \( p \)-dimensional vector of expenditures associated with the input vector \( x \) and \( C = [c_1 \ldots c_n] \) denotes the observed vector of costs in the sample.

4.2. Explaining Efficiency Scores

We use a two-stage approach suggested by Simar and Wilson (2007) to assess the impact of certain environmental variables on municipal efficiency. Therefore, we assume a data-generating process (DGP) in which the environmental variables do not affect the frontier of the production possibility set \( \Phi \), but do influence the resulting efficiency scores. For technical reasons we employ the Shephard input distance function \( d(x, y) \) in what follows. The Shephard distance measure is defined as the reciprocal of the efficiency score \( \theta(x, y) \) used above, hence \( d(x, y) = 1/\theta(x, y) \). We assume the following relationship between the efficiency scores and the environmental variables

(5) \[
d(x, y) = \beta' z + e \geq 1
\]
where \( \beta \) is a vector of parameters and \( e \) is a continuous iid random error variable that is normally distributed with \( N(0, \sigma^2_e) \) and left truncated at \( 1 - \beta' z \).

\(^{12}\) That physical input quantities as well as input prices are not available is a problem common to most studies analyzing global efficiency at the municipal level. Kalb et al. (2011), Geys et al. (2010), and Kalb (2010a), for example, assume identical input prices for all observations and estimate a cost frontier function including only output quantities (and environmental variables) as independent variables. However, with this approach it is not possible to implement flexible functional forms like the translog cost frontier since some terms of higher order are not identifiable. Kriese (2008) estimates a multi-product input distance function using a SFA framework and replacing the unobservable input quantities by the corresponding expenditure categories.
Empirically, we replace \( d(x, y) \) in Equation (5) with its consistent estimator \( \hat{d}(x, y) = 1/\hat{\theta}(x, y) \) defined above. Hence, Equation (5) becomes

\[
\hat{d}(x, y) \approx \beta^* z + e \geq 1.
\]

The maximum likelihood estimates for \( \beta \) and \( \sigma_e^2 \) are consistent but inference is misleading.\(^{13}\) This is why the inference presented in this paper is based on a bootstrap algorithm proposed by Simar and Wilson (2007). Furthermore, we use this method to provide bias-corrected efficiency scores \( \hat{d}(x, y) \) according to

\[
\hat{d}(x, y) = \hat{d}(x, y) - \text{bias}(\hat{d}(x, y)) = 2\hat{d}(x, y) - B_1 \sum_{h=1}^{B_1} \hat{d}_h^*(x, y)
\]

where \( B_1 \) is the number of bootstrap replications and \( \hat{d}_h^*(x, y) \) is the \( b \)th simulated efficiency score based on the assumptions stated above. Hence, in a first step we estimate bias-corrected efficiency scores \( \hat{d}(x, y) \) as described above. In a second step, we use those efficiency scores to draw inferences from the parameter estimates of \( \beta \) by bootstrapping the empirical distribution relying on the assumption stated in Equation (5). Following Simar and Wilson (2007), we simulate 100 draws in the first step and 2,000 draws in the second step.

4.3. Testing Returns to Scale

One goal of our analysis is to discover whether the technology underlying the provision of public services at the local level exhibits increasing, constant, or decreasing returns to scale. We follow the approach suggested by Färe and Grosskopf (1985) and assess local returns to scale by comparing different DEA estimators under alternative assumptions of constant, variable, and non-increasing returns to scale. Then, we adapt the bootstrap procedure suggested by Simar and Wilson (2002) to control for the impact of the environmental variables while statistically testing the resulting test statistics.

An observed municipality or municipal association operates under constant returns to scale if the DEA estimator assuming constant returns to scale, which is defined as

\[
\hat{\theta}_{\text{CRS}}(x, y) = \min \{ \theta > 0 | \theta c \geq C \lambda, y \leq Y \lambda, \lambda \geq 0 \},
\]

is equal to the DEA estimator allowing for variable returns to scale defined in Equation (4). Hence, we can use the following test statistic given in Equation (9) for testing whether or not the underlying

\(^{13}\) This is because \( \hat{d}(x, y) \) is by construction a biased estimator of \( d(x, y) \), the \( \hat{d}(x, y) \) are serially correlated, and the error term \( e \) is correlated with \( z \) via its correlation with the inputs \( x \) and outputs \( y \).
technology exhibits constant returns to scale in our sample. We follow Simar and Wilson (2002) and use the mean of the ratios of the estimators \( \hat{\theta}^{-1}_{\text{CRS}}(x, y) \) and \( \hat{\theta}^{-1}(x, y) \) to calculate the test statistic.

(9)  
\[
\hat{S}_{1}^{\text{CRS}} = n^{-1} \sum_{i=1}^{n} \left[ \hat{\theta}^{-1}_{\text{CRS}}(x_i, y_i) / \hat{\theta}^{-1}(x_i, y_i) \right] \geq 1.
\]

We can test the null hypothesis \( H_0^1 \) of global constant returns scale versus the alternative hypothesis \( H_1^1 \) of no global constant returns to scale by testing \( \hat{S}_{1}^{\text{CRS}} = 1 \) versus \( \hat{S}_{1}^{\text{CRS}} > 1 \).

Since we do not know the distribution of the test statistic defined in Equation (9), we bootstrap \( B=2,000 \) pseudo samples based on the assumption stated in Equation (5) to simulate the empirical distribution of \( \hat{S}_{1}^{\text{CRS}} \) under the null hypothesis \( H_0^1 \) of global constant returns to scale.\(^{14}\) If we reject this null hypothesis \( H_0^1 \), we can test a less restrictive null hypothesis \( H_0^2 \) of non-increasing returns to scale versus the alternative hypothesis \( H_1^2 \) of variable returns to scale by replacing \( \hat{\theta}^{-1}_{\text{CRS}}(x, y) \) in Equation (9) with \( \hat{\theta}^{-1}_{\text{NIRS}}(x, y) \), which is defined by

(10)  
\[
\hat{\theta}_{\text{NIRS}}(x, y) = \min \{ \theta > 0 \mid \theta c \geq C \lambda, \ y \leq Y \lambda, \ i \lambda \leq 1, \ \lambda \geq 0 \}.\]

The empirical distribution of the resulting test statistic is then bootstrapped under the \( H_0^2 \) of non-increasing returns to scale.

5. Data

Municipal budget data for 2004 are taken from local government accounts. These and other data are obtained from the Statistical Office of Saxony-Anhalt, as well as from the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). If municipalities belong to a municipal association, their budgets, including the budget of the joint administration office—if there is one—are aggregated. Overall, we analyze data from 46 independent municipalities and 157 administrative collectivities, of which 122 are Type A associations and 35 are Type B associations.

Revenues and expenditures from the local government accounts are organized according to category (Gliederung) and source (Gruppierung). Our input measures are based on municipal expenditures since input quantities and prices are not available.

\(^{14}\) Simar and Wilson (2002) actually use a bootstrap procedure based on a smoothed kernel density estimation combined with the reflection method. That is, the density of the efficiency scores is estimated from a sample with \( 2n \) observations, including reflected observations around the boundary at 1. We adapt the procedure provided in Simar and Wilson (2007) to test for returns to scale and draw from the left-truncated normal distribution describing the impact of environmental variables on the efficiency scores.
Table 2: Structure of current (non-investment) expenditures in the municipal core budgets in Saxony-Anhalt, 2004 in percent (mean)—District-affiliated municipalities

<table>
<thead>
<tr>
<th>Category</th>
<th>Category number</th>
<th>Share of current expenditures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration</td>
<td>0</td>
<td>15.62</td>
</tr>
<tr>
<td>Public safety</td>
<td>1</td>
<td>3.40</td>
</tr>
<tr>
<td>Schools</td>
<td>2</td>
<td>3.51</td>
</tr>
<tr>
<td>Science, research, culture</td>
<td>3</td>
<td>1.28</td>
</tr>
<tr>
<td>Social security</td>
<td>4</td>
<td>17.82</td>
</tr>
<tr>
<td>Healthcare, sports, recreation</td>
<td>5</td>
<td>2.95</td>
</tr>
<tr>
<td>Buildings, housing, traffic</td>
<td>6</td>
<td>6.11</td>
</tr>
<tr>
<td>Public facilities and economic development</td>
<td>7</td>
<td>7.50</td>
</tr>
<tr>
<td>Municipal enterprises, public utilities, public real estate, and special assets</td>
<td>8</td>
<td>2.62</td>
</tr>
<tr>
<td>Financial management</td>
<td>9</td>
<td>39.24</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations.

Regarding the source, the expenditure data are taken only from the current account. We also exclude certain categories. Although financial management is a large part of the expenditures (cf. Table 2), it is excluded except for the subcategory “interest payments.” Including the other items in this category would result in distortions as they are either compensations for expenditures of other members of municipal associations or the joint administration or transfer payments to the district that do not correspond to any output at the municipal level; some expenditures do not even represent factor costs at all, such as the redistributed business tax revenues.\(^{15}\)

To avoid comparison problems between municipalities with different degrees of outsourcing, we have excluded some services that are often allocated to municipal enterprises or municipal special purpose associations for which no output indicators are available (waste and sewage disposal—part of Category 7—as well as the provision of electricity, gas, water, etc.—part of Category 8).\(^{16}\)

The first input, labor, is measured as expenditure for staff in all categories except those mentioned above. Labor expenditures on average account for more than half the total expenditures. Capital expenditures are measured by the sum of interest payments and expenditures for rent and lease. They represent a relatively small fraction of municipal expenditures. The third input is resources and intermediate inputs. As such it is a considerably broad input category and makes up a substantial part

\(^{15}\) Table 2 also illustrates the necessity of analyzing global municipal efficiency because the interest payments in Category 9 as well as the expenditures of Category 0 and some items of Categories 7 and 6 represent overhead costs that otherwise would have to be allocated pro rata to the particular municipal service.

\(^{16}\) This seems to be more of a problem for the independent cities and districts than for the district-affiliated municipalities. We found 298 municipal enterprises and special purpose associations owned by the district-affiliated municipalities. About two-thirds are public utilities and municipal housing companies. Deducting the expenditures for sewage disposal, which consist primarily of transfer payments or compensations to special purpose associations, leaves only 23 municipal enterprises that might cause minor distortions if we neglect the municipal enterprise sector in our analysis.
of the budget.\textsuperscript{17} A further breakdown into several more homogenous inputs, however, is neither possible nor advisable because the number of dimensions of the estimation problem should be kept as small as possible.

We use only the economically relevant expenditure sources for input calculation. Imputed costs and internal offsets, for example, are reported only for book-keeping purposes. Redistributed tax revenues, general apportionments paid, and allocations to reserves cannot be matched to measurable municipal output either. Furthermore, expenditures for covering previous-year deficits are excluded since the analysis is restricted to one year and should not be confounded by previous years’ financial management. Allocations to the capital budget are left out so as to be consistent in restricting the analysis to the current account.

To correct the data for double cost counting, rent income from other municipalities is subtracted from the capital input, and all three inputs are proportionately reduced by reimbursements of expenditures and grants received from member municipalities and counties.\textsuperscript{18} The aggregation and adjustment of the inputs is closely related to the calculation formula of the statistical office of Saxony-Anhalt (2010).\textsuperscript{19}

The outputs are geared to the municipality’s functions\textsuperscript{20} (based on the expenditure structure in Table 2) and resemble those used in other studies (e.g., Geys and Moesen 2009; Geys et al. 2007; Kalb 2010a). Nevertheless, they are often only a rough proxy. On average, 87\% of the expenditures in Category 4, social security, are made for childcare. Moreover, almost all children (from three to six years) visit a childcare center, while less than 1\% is in family daycare.\textsuperscript{21} Therefore, the number of approved places in childcare centers in the municipality is an appropriate output.\textsuperscript{22} Students in elementary school are the measure for Category 2, which comprises mainly elementary schools.

Recreational area is used as a proxy for local public health, sport, and recreation facilities; traffic area serves as a measure of municipal street-related outputs. As the recreational area is relatively small and measured in the same terms as traffic area, it is not used as a separate output but combined with the traffic area.

\textsuperscript{17} The detailed calculation scheme is available from the authors on request.

\textsuperscript{18} However, grants or cost refunds to private or public enterprises or to private nonprofit organizations (e.g., childcare facilities) are included in our input figures, except for the explicitly excluded expenditure categories. The underlying assumption is that the receiving units provide local public services (or at least intermediate inputs for the municipality’s administration) that the municipality otherwise would have to provide itself. Details of the calculation are available from the authors on request.

\textsuperscript{19} Statistical Office Saxony-Anhalt (2009).

\textsuperscript{20} The municipalities are obliged to provide many of those services by higher levels of government.

\textsuperscript{21} Statistical Offices of the Federation and the States (2008).

\textsuperscript{22} These numbers are available only for 2006. The statistics show that the number of available/approved places has increased somewhat over the period 2002 to 2006 and also afterward. No information on the actual number of children in childcare centers is available before 2006. Thus, our output measure might be slightly biased upward (Statistical Office Saxony-Anhalt 2010).
Certain municipal services that are either public consumption goods for private households or public inputs for the private enterprise sector (or both) cannot be measured properly (or adequate data has not been published). This problem arises for services involving public safety, but also for many other services, such as economic development or business-related infrastructure. Thus, we assume that these particular public outputs are correlated with the size of the population (public consumption goods) and the number of employees subject to social security contribution (public inputs). An overview of the outputs is given in Table 3.

**Table 3: Outputs**

<table>
<thead>
<tr>
<th>Output</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Number of inhabitants</td>
</tr>
<tr>
<td>Childcare places</td>
<td>Number of approved places in the childcare centers within the municipality</td>
</tr>
<tr>
<td>Children in elementary school</td>
<td>Number of children attending elementary schools in the municipality</td>
</tr>
<tr>
<td>Traffic and recreational area</td>
<td>Traffic and recreational area in hectare</td>
</tr>
<tr>
<td>Employees subject to social security</td>
<td>Number of employees working in the municipality who are subject to social</td>
</tr>
<tr>
<td>contribution</td>
<td>security contribution</td>
</tr>
</tbody>
</table>

Source: Authors’ compilation.

At the second stage, the DEA score is regressed on environmental variables, which are supposed to explain differences in the efficiency level, as outlined in Section 3. They are described in Table 4.

The institutional variables are dummy variables for the type of municipality and the number of member municipalities forming one municipal association. The base group of the dummy variables is the independent municipalities. Municipal associations comprise up to 22 municipalities, although having more than 10 is rare.

**Table 4: Environmental variables**

<table>
<thead>
<tr>
<th>Environmental variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population density</td>
<td>Population divided by the total area in square kilometers</td>
</tr>
<tr>
<td>Share of senior citizens</td>
<td>Population aged 65 years and older as share of total population</td>
</tr>
<tr>
<td>Relative population change</td>
<td>Absolute value of the relative population change between 2000 and 2004</td>
</tr>
<tr>
<td>Dummy variables for type of municipality</td>
<td>Type A associations, Type B associations, base group is independent</td>
</tr>
<tr>
<td></td>
<td>municipalities</td>
</tr>
<tr>
<td>Number of municipalities</td>
<td>Number of municipalities within the administrative collectivity, equal to one</td>
</tr>
<tr>
<td></td>
<td>for independent municipalities</td>
</tr>
<tr>
<td>Debt per capita</td>
<td>Total debt divided by population</td>
</tr>
<tr>
<td>Relative equalization transfers</td>
<td>Equalization transfers as a percentage of total adjusted current income</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>Number of unemployed divided by population aged between 15 and 65 years</td>
</tr>
</tbody>
</table>

Source: Authors’ compilation.

Furthermore, we include a set of control variables to account for additional variables affecting efficiency at the local level. We control for equalization transfers (grants) as a percentage of total adjusted current income, which could be interpreted as an indicator for the flypaper effect. Additionally, debt per capita is included as a fiscal control variable. All municipalities have debts. The
average is about 950€ (per inhabitant) but there is a great deal of variation.

As demographic variables we include population density, the share of senior citizens (which captures the aging aspect), and relative population change to account for overall population decline. Population density varies enormously between approximately 20 and 1,200 inhabitants per square kilometer. To account for possible nonlinear effects of population density, a square term is also included. Senior citizens are inhabitants aged 65 years and older; they are included as fraction of the total population. The mean share is 20%, with a maximum of about 27%. Population change is measured by the absolute value of the relative population change between 2000 and 2004.

We also control for the unemployment rate, measured as the number of unemployed divided by population aged 15 to 65. On average, 10% of the population was unemployed in 2004. Although this is a relatively high number, it should be kept in mind that it would be even higher if measured as percentage of the labor force. Descriptive statistics of all variables are presented in Table 5.

Table 5: Descriptive statistics (N=203)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inputs:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor (in €)</td>
<td>2,894,448</td>
<td>2,894,205</td>
<td>460,538</td>
<td>17,700,000</td>
</tr>
<tr>
<td>Capital (in €)</td>
<td>428,926</td>
<td>502,415</td>
<td>48</td>
<td>3,222,858</td>
</tr>
<tr>
<td>Resources and intermediate inputs (in €)</td>
<td>2,212,260</td>
<td>2,430,012</td>
<td>347,249</td>
<td>17,800,000</td>
</tr>
<tr>
<td><strong>Outputs:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>9,615.13</td>
<td>7,833.50</td>
<td>2,229.00</td>
<td>45,737.00</td>
</tr>
<tr>
<td>Childcare places</td>
<td>443.08</td>
<td>340.29</td>
<td>102.00</td>
<td>2,046.00</td>
</tr>
<tr>
<td>Children in elementary school</td>
<td>235.81</td>
<td>194.76</td>
<td>0.00</td>
<td>1,179.00</td>
</tr>
<tr>
<td>Traffic and recreational area</td>
<td>465.15</td>
<td>219.86</td>
<td>67.00</td>
<td>1,191.00</td>
</tr>
<tr>
<td>Employees s.t. social security contribution</td>
<td>2,508.83</td>
<td>3,169.39</td>
<td>213.00</td>
<td>17,918.00</td>
</tr>
<tr>
<td><strong>Environmental variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td>141.90</td>
<td>169.63</td>
<td>21.16</td>
<td>1,216.41</td>
</tr>
<tr>
<td>Share of senior citizens</td>
<td>0.20</td>
<td>0.02</td>
<td>0.13</td>
<td>0.27</td>
</tr>
<tr>
<td>Relative population change</td>
<td>0.04</td>
<td>0.02</td>
<td>0.00</td>
<td>0.17</td>
</tr>
<tr>
<td>Number of municipalities</td>
<td>5.49</td>
<td>4.00</td>
<td>1.00</td>
<td>22.00</td>
</tr>
<tr>
<td>Equalization transfers</td>
<td>0.31</td>
<td>0.09</td>
<td>0.00</td>
<td>0.48</td>
</tr>
<tr>
<td>Debt per capita</td>
<td>942.90</td>
<td>657.11</td>
<td>24.14</td>
<td>4,041.48</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>0.10</td>
<td>0.02</td>
<td>0.01</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations.
6. Results and Discussion

Before we present the efficiency scores and test the hypotheses derived in Section 3, we analyze the technology underlying the provision of public services at the local level.

6.1. Scale Efficiency

First, we test whether or not the production frontier estimated exhibits global constant returns to scale. Using the method described in Section 4.1, we reject the null hypothesis of global constant returns to scale with a p-value of 0.0000. Next, we test against the null hypothesis of global non-increasing returns to scale, which is also rejected with a p-value of 0.0235.23 Hence, we conclude that the production technology under consideration exhibits areas of decreasing as well as increasing returns to scale. In other words, the size of municipal entities affects theoretically possible productivity.

Only about 20% of the observations in our sample operate under local constant returns to scale, whereas about 43% experience local increasing returns to scale and about 37% operate under local decreasing returns to scale. We bootstrapped p-values under the null hypothesis of local constant returns to scale for each observation. Table 6 summarizes these findings. The third column presents the fraction of rejected tests in the area of increasing, as well as decreasing, returns to scale.

Table 6: Summary statistic of local scale efficiency

<table>
<thead>
<tr>
<th>Local returns to scale</th>
<th>Number of obs.</th>
<th>Significant at the 5 % level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing returns</td>
<td>88 (43%)</td>
<td>18%</td>
</tr>
<tr>
<td>Constant returns</td>
<td>40 (20%)</td>
<td>-</td>
</tr>
<tr>
<td>Decreasing returns</td>
<td>75 (37%)</td>
<td>28%</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations.

Figure 2 illustrates the relationship between scale inefficiency and population size.24 At about 8,000 inhabitants, scale inefficiency is zero. The mean size of all scale-efficient municipal associations is 7,892 inhabitants. At the left- as well as the right-hand side of this population size, scale inefficiency increases and we find a strong negative and positive correlation between both variables, respectively.

23 To prove the robustness of our results, we employ another test statistic for global constant and non-increasing returns to scale, respectively. In addition to the mean of the ratios of the alternative DEA estimators as discussed in Section 4.3, we calculate the ratio of the means. We find p-values of 0.0000 and 0.0230 for the tests against global constant and global non-increasing returns to scale. Furthermore, we bootstrapped all tests discussed above using the smoothed kernel density estimator for the density function of the efficiency scores originally suggested by Simar and Wilson (2002) ignoring the influence of environmental variables. We find similar results, albeit the p-values are lower than the ones presented in the text.

24 This figure can be interpreted as the elasticity of scale efficiency with respect to the empirically observed output mix.
6.2. Bias-Corrected Efficiency Scores

As described in the methodology section, we perform the bootstrap procedure to correct the bias of the estimated efficiency scores. The resulting scores are given in Table 7. By definition, the bias-corrected convex hull is farther away from the consistent but biased DEA frontier. This is why we observe no efficient observation according to the bias-corrected efficiency scores. The histogram of those efficiency scores is shown in Figure 3. Our calculation shows that the mean square error is lower for the bias-corrected frontier estimation (Simar and Wilson 2000). Hence, the bias-correction procedure is justified.

Figure 3: Empirical distribution of bias-corrected efficiency scores

Source: Authors’ calculations.
Table 7: Bias-corrected technical efficiency

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Median</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>203</td>
<td>1.18</td>
<td>1.23</td>
<td>1.05</td>
<td>1.85</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations.

We find that the average municipality could reduce its inputs by about 19% to produce the observed output mix if it produced efficiently.

6.3. The Impact of Environmental Variables on Municipal Efficiency

Finally, we use the bias-corrected efficiency scores to evaluate the impact of environmental variables on municipal efficiency.

The parameter estimates and the $p$-values resulting from the second-stage bootstrap algorithm described in Section 4.2 are given in Table 8. We find a significant negative dummy for Type B associations. Generally, this finding underscores the importance of municipal institutions for efficient local public service production. Furthermore, since the base category is the independent municipalities, the result supports Hypothesis 1. However, the insignificance of the dummy indicating Type A associations may be caused by compensating institutional effects of decentralized administration in this kind of municipal associations, as discussed in Section 3.

Table 8: Results of the second-stage truncated regression

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficients</th>
<th>$p$-values</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.7682</td>
<td>0.0000</td>
<td>0.4597</td>
</tr>
<tr>
<td>Dummy Type A association</td>
<td>-0.0564</td>
<td>0.1975</td>
<td>-0.1446</td>
</tr>
<tr>
<td>Dummy Type B association</td>
<td>-0.1690</td>
<td>0.0015</td>
<td>-0.2699</td>
</tr>
<tr>
<td>Number of municipalities</td>
<td>0.0081</td>
<td>0.0740</td>
<td>-0.0006</td>
</tr>
<tr>
<td>Population density</td>
<td>-0.0007</td>
<td>0.0045</td>
<td>-0.0013</td>
</tr>
<tr>
<td>Population density squared</td>
<td>0.0000</td>
<td>0.2470</td>
<td>0.0000</td>
</tr>
<tr>
<td>Share of senior citizens</td>
<td>3.2486</td>
<td>0.0000</td>
<td>1.6775</td>
</tr>
<tr>
<td>Debt per capita</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.4236</td>
<td>0.4570</td>
<td>-1.5073</td>
</tr>
<tr>
<td>Absolute value population change</td>
<td>0.7162</td>
<td>0.3760</td>
<td>-0.9446</td>
</tr>
<tr>
<td>Equalization transfers</td>
<td>-0.0057</td>
<td>0.0005</td>
<td>-0.0085</td>
</tr>
</tbody>
</table>

Number of observations 203

Additionally, at the 10% significance level, we find empirical support for Hypothesis 2. Hence, ceteris paribus, an increasing number of municipalities in a municipal association decreases its efficiency, which suggests an upper limit for the benefits from political decentralization.

Regarding the control variables employed in our analysis, it is worth noting that we find no empirical
support for the flypaper effect or potential negative efficiency effects of soft budget constraints. Finally, the negative efficiency effect of a higher share of senior citizens might be explained by the fact that older people do not benefit from the main municipal expenditure categories (childcare, primary schools) and hence are not interested in efficiently producing these services; however, this is mere speculation.

7. Conclusion

Bigger and more centralized municipal entities have been the trend in Europe for the past few decades. Expenditure reductions and efficiency gains are often expected to follow from municipal mergers, but evidence from the economic literature in this regard is far from unambiguous. Therefore, we analyze the effect of municipal size and institutional structure at the local level on efficiency in public goods production.

We find the technology underlying the production of public goods at the local level to exhibit variable returns to scale, with about 80% of all observations operating under either increasing or decreasing returns to scale. We adapt the bootstrap procedure suggested by Simar and Wilson (2007) to test this finding. For about 20% of all observations, we reject the constant returns to scale hypothesis at the 5% significance level. Hence, there is a natural limit for scale efficiency gains from enlarging municipal entities. Based on the estimated production frontier, we calculate bias-corrected efficiency scores. According to our results, the inputs, on average, could be reduced by about 19% to produce the observed output mix if production was efficient.

In the second stage of our empirical analysis, we show that the institutional structure has a substantial impact on efficient public service production. Despite the higher transaction cost due to a more complex institutional setting, decentralized municipal structures are at least as and sometimes more efficient than centralized ones. Hence, we conclude that the positive effects of lower information costs and the less intense information asymmetries, as well as the “exit and voice” aspect, outweigh the effect of transaction costs in the production of public services at the local level.

Consequently, it is doubtful whether the standard components of most municipal reforms in Germany and other countries—increasing the size of local governments and transforming municipal associations into centralized independent municipalities—have the intended effect on municipal efficiency. Our results suggest, instead, that higher levels of government should avoid forcing municipal amalgamations and should carefully analyze the particular causes and cures of every single case of municipal inefficiency.
References


