Voting for Inefficiency

Pio Baake* Maria Nieswand† Lilo Wagner‡
DIW Berlin *DIW Berlin* *DIW Berlin*

February 15, 2015

Abstract

In this paper, we propose a probabilistic voter model that rationalizes the inefficient use of labor input in municipality-owned enterprises. Our focus is on the relationship between absolute as well as relative inefficiency measures and the size of the municipality in which public firms operate. We show that larger communities maintain enterprises which are on average more efficient in absolute as well as relative terms and that efficiency shows a lower variance for large enterprises. Using data from a sample of electricity distribution companies in Germany we find our theoretical predictions consistent with the companies' observed behavior.

Keywords: Voting, Cost inefficiency, Asymmetric information, Electricity distribution

JEL Classification Numbers: H83, H72, L32, D24

Acknowledgment: This paper is produced as part of the project KOMIED (Municipal infrastructure companies against the background of energy policy and demographic change), a project financed by Leibniz Association. We thank Astrid Cullmann, Peter Bogetoft and Christian von Hirschhausen for valuable comments and fruitful discussions. Also we thank the participants of the 8th North American Productivity Workshop as well as the participants of the Brown Bag Seminar of the Cluster of Industrial Economics at DIW Berlin.

*DIW Berlin, Mohrenstr. 58, 10117 Berlin, Germany, Email: pbaake@diw.de.
†DIW Berlin, Mohrenstr. 58, 10117 Berlin, Germany, Email: mnieswand@diw.de.
‡DIW Berlin, Mohrenstr. 58, 10117 Berlin, Germany, Email: lwagner@diw.de.

1 Introduction

Public decision-making and the behavior of public enterprises have been of core interest since many years in various fields of both theoretical and empirical research, e.g., public choice, public finance and efficiency analysis. This great attention stems from the fact that, in numerous economies worldwide, the public sector is an important provider of goods and services. However, a central concern being raised in this context is the lack of efficiency in public goods and service provision. One channel through which inefficient production might be fostered is the close relationship between public enterprises and their owners, e.g., local governments. At the cost of efficiency, many public firms claim to pursue objectives of public interest and social welfare. In contrast to managerial failure or inability, inefficiency serves pursuing alternative objectives. For example, local governments can boost votes by artificially generating employment opportunities within the public enterprises.

In this paper, we propose a probabilistic voter model that rationalizes the inefficient use of labor input in municipality-owned enterprises. This is done by competing parties that promise to create inefficiently many public jobs. Our main interest is to investigate the relationship between absolute as well as relative inefficiency measures and the size of the municipality in which public firms operate. Size is thereby represented by the number of voters. Our analysis suggests that larger communities maintain enterprises which are on average more efficient in absolute as well as relative terms and that efficiency shows a lower variance for large enterprises. Thus, we expect larger firms to show smaller relative inefficiency measures. To test this hypothesis empirically, we quantify efficiency for a sample of electricity distribution companies by means of conditional efficiency analysis. We find the predictions consistent with the companies' observed behavior. Our analysis furthermore succeeds in linking absolute and relative efficiency measures of the public enterprises, which has not been done before.

The relationship of governments and public enterprises is particularly taken into account by the public choice literature, which provides alternative approaches to explain the behavior of public enterprise. As summarized by Duncombe et al. (1997) two most common public choice models are the (median) voter and principal-agent models, the latter beginning with the seminal work by Niskanen (1971). While voter models rely on the assumption that political outcomes are based on voter preferences, principal-agent models emphasize the importance of the objectives of public managers to the public choice process.

As implied by Niskanen (1971) and others, inefficiency and government size should be negatively related because bureaucrats maximize the 'bureaucratic slack' (budget minus least cost of production) by increasing monitoring costs which are supposed to be higher in more complex environments. This presumes the existence of a powerful bureaucrat who seeks to maximize his budget while the government itself acts as a benevolent social planner. The government, however, consists of elected politicians whose goal it is to remain into power and who themselves might have an incentive to provide excess employment opportunities. This approach is taken in the present paper.

Consider the following setting. Two parties compete for voters out of which after the elections some are recruited by a public enterprise. This firm is said to be absolutely efficient if its input-output ratio is defined by an exogenously given efficient production function. Inefficiency however creates employment opportunities from which an individual voter potentially benefits. This benefit forms through two channels: first, his probability of being recruited by the public enterprise increases with inefficiency, and second, once employed, inefficiency is attractive because a constant work load is split among several workers. However, the budget must be balanced implying that inefficiency creates a social burden which has to be borne by the whole community. Varying inefficiency measures are explained by the presence of some manager who possesses private information about the costs for efficiency, for instance monitoring costs or his private management qualities. When the manager type distribution is independent of the community size, larger communities maintain more efficient enterprises on average. The reason for this result is that efficiency saves more labor costs in larger communities while manager salaries are unaffected by size.

The relative efficiency measure can be derived by taking the absolute efficiency which is produced under the lowest cost manager type. Although less strong, the relationship between community size and efficiency is maintained when efficiency is defined as a relative measure.

The remainder of the paper proceeds as follows: Section 2 present the voter model while Section 3 discusses the relationship between efficiency and the number of voters the parties are competing for. In Section 4 the empirical analysis of relative efficiency is conducted. After introducing the estimation method, data, variable selection and results are presented. Finally, Section 5 concludes.

2 The model

Two parties A and B compete for the votes of N voters. Whoever accumulates more votes wins the election and governs a public enterprise (an electricity distribution firm). More precisely, the firm is not governed by the government directly, but the task is delegated to some manager who possesses private information about the costs of efficiency. The parties compete over menus of contracts which they pledge to offer to the enterprise's manager (or bureaucrat) in case of election. A contract consists of the manager's salary ζ and an efficiency level $\alpha \in (0,1]$ to which the manager commits upon contracting. More specifically, α denotes the ratio between the efficient labour input level L^e and the actual labour input L_L, i.e. $\alpha = L^e/L_L$. Thus, $\alpha = 1$ means the firm is efficient, while smaller α imply lower efficiency levels. For simplicity, we restrict our attention to labor input. This is basically identical to presuming non-substitutability between labor and capital inputs.[1] The manager's salary can be made contingent on the efficiency measure, thus z is a mapping from α into the set of real numbers. ζ_i represents the payment that party i offers. Each party thus competes for voters over an election program $\mathcal{A}_i = \zeta_i(\alpha)$.

The production technology is the enterprise's set $\tilde{\Psi}$ of feasible input-output combinations. The set is defined by $\tilde{\Psi} = \{(L_L, \tilde{y}) \in \mathbb{R}^2_+ | L_L \text{ can produce } \tilde{y}\}$. Absolute efficiency is defined by its frontier $\tilde{\Psi}^\delta$. That is, for every output $\tilde{y}$, the minimum L_L in $\tilde{\Psi}$ defines $\tilde{\Psi}^\delta$. It is denoted by L^e and this absolutely efficient input-output relation is assumed to be strictly positive. Further assumptions on the efficient production set are not required. The efficient labour input level can thus be written as a mapping $L^e : y \to \mathbb{R}_+$. For technical convenience, we will allow labour input to take on real values in the following.

The manager has private information $\theta \in [0,1]$ about the costs of efficiency, given by $C(\alpha)H(\theta)$.[2] For simplicity, it is assumed that θ is uniformly distributed. The manager's quasi-linear utility function is given by $U^M(\alpha,\theta) = \zeta(\alpha) - C(\alpha)H(\theta)$. For the costs of efficiency, we impose the following assumption.

Assumption 1 $C' > 0, C'' > 0$ *and* $H' < 0 \leq H''$

Thus, higher manager types θ face lower costs of efficiency. Assumption 1 is standard in models of adverse selection. The first derivatives ensure that the single crossing property for $\zeta^K(\alpha,\theta,\overline{U})$ is satisfied, with $\zeta^K(\cdot)$ being defined such that $U^M(\zeta^K(\alpha,\theta,\overline{U}),\alpha,\theta) := \overline{U}$.[3] Second, $C'' > 0$ guarantees interior solutions and $H'' \geq 0$ makes sure that the efficiency level increases when the costs of efficiency are low (i.e. θ is high). Finally, note that manager costs are assumed to be independent of firm size. This assumption can be stated differently: monitoring more workers can be delegated to intermediate managers who themselves do not dispose of informational advantages over the manager.

Consumption homogeneous voters consume the good produced by the public enterprise. Since the sales price is usually regulated at governmental level and thus beyond the sphere

[1] This assumption is without loss of generality. If capital inputs were substitutable by labor, the voting environment will result in inefficiently high labor input. Also, empirically, we evaluate efficiency in the market for electricitiy distribution. Clearly, capital – cables – and labor are not substitutable here.

[2] The costs of efficiency can alternatively be interpreted as the manager's ability.

[3] The single crossing property states that $\partial^2\zeta^K(\cdot)/\partial\alpha\partial\theta < 0$.

of influence of some municipal administration, we assume it to be exogenously given and normalized to 1. We further impose that regulation is efficient, meaning that regulation itself does not create further distortions.[4] At this given price, each voter demands x units of the good. The total number of units demanded in a community equals the output produced, $\tilde{y}$. For simplicity x (and thus also $\tilde{y}/N$) is normalized to 1. Without loss of generality, the utility derived from consuming this quantity at the given price is normalized to zero.

Similarly, public wages are fixed in tariff contracts (TVöD in Germany) which are the outcome of nationwide bargaining between the representatives of governmental organizations and those of the union. Denote by w this exogenously given wage. Then, the firm's profit can be written as

$$\Pi(\alpha, N) = N - \frac{1}{\alpha} w L^e(N) - \zeta(\alpha). \tag{1}$$

The government's actions are subject to a budget constraint. Thus, potential losses have to be financed by a local tax. Tax payers may be either voters who are employed by the public firm or those who are privately employed.[5] This payment denoted by t is assumed to be flat. It is a mapping from the efficiency level α and the community size N into $\mathbb{R}$. Paying this tax generates an excess burden to individuals of $\mu \geq 1$, such that every payment effectively costs an individual μt.

Privately employed are assumed to earn the same wage as public employees. Public employees however benefit from inefficiency if present, because each worker then produces less output.[6] This additional utility from inefficiency is given by some function $U^I : \alpha \rightarrow \mathbb{R}$.[7] We make the following assumption.

Assumption 2 $U^{I\prime} < 0$, $U^{I\prime\prime} < 0$, $U^{I\prime\prime\prime} \leq 0$

Assumption 2 states that the additional utility from inefficiency is constantly decreasing in efficiency. Also, the function is concave, meaning that higher marginal utility from being inefficient is derived at high efficiency levels. Thus, pausing is more valuable for workers

[4] In Germany, transit prices of all nets are regulated by the Bundesnetzagentur. This fixed price assumption is however not crucial for our results. We discuss this topic further below.

[5] Assuming the absence of unemployed voters is without loss of generality to our model if these unemployed have to contribute the same individual tax payment t.

[6] Thus, wages are not the competitive wages. The assumption that publicly employed workers derive a greater utility than an average privately employed worker is supported for instance by Shleifer and Vishny (1994), who similarly claim that inefficiency arises due to the fact that beneficiaries of it produce less and earn more. They also present evidence for this claim.

[7] This utility is actually independent of firm size. To see this, consider two firms of different size, one with efficient labour input L_1^e and one with efficient labour input L_2^e. When efficient, the firms produce y_1 or y_2 units of output for which each worker has to effectively spend h^e working hours. For a given efficiency level $\overline{\alpha}$, each worker then effectively spends $\overline{\alpha} h^e$ working hours, independent of firm size.

facing tight work plans. The assumption concerning the third derivative is made in order to obtain unambiguous results. It is not very strict though, because results hold also for positive third derivatives if not too large.

The quasi-linear utilities of the two voter groups can then be written as

$$\begin{aligned} U^E(\alpha, N) &= w - \mu t(\alpha, N) \\ U^F(\alpha, N) &= w - \mu t(\alpha, N) + U^I(\alpha). \end{aligned} \tag{2}$$

As mentioned above, taxes are used to counterbalance firm losses. Thus, an individual pays tax t, given by

$$\begin{aligned} t(\alpha, N) &= \frac{1}{N}\Big(\frac{1}{\alpha} w L^e(N) - \Pi(\alpha, N)\Big) \\ &= \frac{2}{\alpha N} w L^e(N) + \frac{\zeta(\alpha)}{N} - 1. \end{aligned} \tag{3}$$

Voting is probabilistic, meaning that voting is subject to random effects, mirroring common as well as individual preference shocks. More precisely, let $d \in [-1/2\delta, +1/2\delta]$ denote the common preference shock in favor of party B. Similarly let $s_j \in [-1/2, 1/2]$ denote the private preference shock of voter j for the same party. Both common and individual preference shocks are uniformly distributed. A voter j then elects party A if and only if

$$\mathcal{U}(\mathcal{A}_i, N) \geq \mathcal{U}_B(N) + d + s_j,$$

where $\mathcal{U}(\cdot)$ denotes a voter's expected utility from a win of party i, given this party's election program $\mathcal{A}_i$. In particular, prior to elections, voters are not aware which role they play afterwards. They may be either privately or publicly employed. The respective probability to be selected to belong to one of the groups is determined by the election program. A contract which induces a high inefficiency level for some cost type implies that a high share of workers will be employed publicly for the same type. Note that we refrain from considering unemployment for the moment. Thus, all voters which will not be employed by the public enterprise will have an employment in the private sector. Further below, it is shown that this abstraction is without of generality.

Finally, the game structure is the following:

(1) Parties A and B announce their election programs. Both parties seek to maximize their probabilities of being elected.

(2) Private and common preference shocks realize, voters observe the programs and elect

one of the two parties.

(3) The party that gets more votes wins the election and governs the public enterprise.

3 Efficiency and population size

It is standard from models of adverse selection to show how the cost minimizing incentive compatible incentive scheme is characterized. The manager is offered by each party, a menu of contracts $\zeta_i(\alpha)$ from which he chooses one according to his cost type θ. Following to the revelation principle, the compensation scheme can be expressed as a function of the cost type. The efficiency when party i wins the election is $\hat{\alpha}_i : \theta \to (0,1]$ which by standard arguments should be an increasing function. The compensation scheme for party i is then given by

$$\hat{\zeta}_i(\theta) = C(\hat{\alpha}_i(\theta))H(\theta) - \int_0^\theta C(\hat{\alpha}_i(\tau))H'(\tau)d\tau.$$

Accordingly, the utilities of each voter group, $U^E(\cdot)$ and $U^F(\cdot)$ (see equation (2)) can be expressed as functions of θ, $\tilde{U}_i^E(\theta, N)$ and $\tilde{U}_i^F(\theta, N)$ respectively.[8] From this, the expected voter's utility from a win of party i is given by

$$\mathcal{U}(\mathcal{A}_i, N) = \int_0^1 \Big(\frac{L^e(N)}{\hat{\alpha}_i(\theta)N}\tilde{U}_i^F(\theta, N) + \frac{\hat{\alpha}_i(\theta)N - L^e(N)}{\hat{\alpha}_i(\theta)N}\tilde{U}_i^E(\theta, N)\Big)d\theta.$$

The parties then maximize their probabilities of winning the election. In equilibrium, efficiency measures are symmetric and they are functions of the population size N. We denote them by $\alpha^*(\theta, N)$. The following can then be shown.

Lemma 1 *$\alpha^*(\theta, N)$ exists and is unique. It holds that $\partial\alpha^*(\theta, N)/\partial\theta > 0$.*

Therefore, the lower a manager's costs for efficiency, the higher the contractual efficiency levels. This result is in line with standard analysis of adverse selection problems and not innate to the voter environment. Comparative statics further reveal the following.

Proposition 1 *The equilibrium efficiency level $\alpha^*(\theta, N)$ is strictly increasing in the population size N.*

Proposition 1 says that efficiency levels increase with population – or firm – size. The reason is that the firm's marginal costs for efficiency are lower at larger enterprises. Stated

[8]Correspondingly, $t(\alpha, N)$ is then expressed as $\tilde{t}(\theta, N)$ by replacing $\zeta(\alpha)$ through $\hat{\zeta}_i(\theta)$.

differently, increasing efficiency has two effects: first, it lowers labour costs and second, it raises manager salaries. In larger communities, increasing efficiency (which is a relative measure) saves more labour costs, since the total number of workers is higher than in small communities. On the other hand, community size has no effect on the manager's costs for efficiency. The result implies that expected absolute efficiency measures should also increase in population size.

The problem with empirically measuring efficiency is that the absolutely efficient technology set $\tilde{\Psi}$ cannot be observed. Efficiency evaluation is usually done using nonparametric methods. These methods have the advantage that they do not need to rely on functional form assumptions. On the other hand however, absolute efficiency cannot be measured. Instead, nonparametric methods estimate the relative efficiency.

More precisely, a new production set is defined given the observed input-output combinations. Each observation is then compared to the efficient frontier of this new production set. The productive efficiency of a municipality's enterprise relative to this new technology frontier is what we refer to as the relative efficiency. Formally, it is given by

$$\beta^*(\theta, N) = \frac{\alpha^*(\theta, N)}{\overline{\alpha}(N)} \qquad \text{with } \overline{\alpha}(N) := \alpha^*(1, N). \tag{4}$$

Before being able to derive a statement on the relationship between community size N and the relative efficiency measure, it is necessary to know how the incentive scheme varies with community size. The following Lemma explores this relationship.

Lemma 2

$$\frac{\partial}{\partial N}\left(\frac{\partial \alpha^*(\theta, N)}{\partial \theta}\right) < 0$$

if N is large.

Lemma 2 states that incentive schemes flatten with community size at least for large communities. The intuition for this result is again based on the fact that the size of the community does not affect the manager's cost for increasing efficiency. On the one hand, flattening incentive schemes increases informational rents paid to high manager types. On the other hand however, it implies higher efficiency levels for lower manager types. As before, efficiency is more important in large communities as savings on labor costs are especially large here. Thus, higher efficiency levels in large communities result in similar contracts for managers. An immediate consequence of lemma 2 is the following.

Proposition 2 *The equilibrium relative efficiency measure $\beta^*(\theta, N)$ is strictly increasing in the community size if N is large.*

Proposition 2 states that, like the absolute efficiency, relative efficiency should equally increase in population size. Thus, concerning the variable N, absolute and relative efficiency measures hint into the same direction. This correlation arises unambiguously because incentive schemes flatten with population size, see Lemma 2. Hence, while absolute efficiency increases in N for all manager types, the effect is stronger for higher cost types. This implies that the increase of numerator in equation (4) is higher than that of the denominator. Note however that the effect on absolute efficiency is stronger. Thus, when observing increasing relative efficiency measures, the true effect of population size on efficiency will be underestimated.

Finally, how would we expect efficiency variance to vary with firm size? The following proposition gives an answer to this question.

Proposition 3

$$\frac{\partial Var(\alpha^*(\theta, N))}{\partial N} < 0 \quad \textit{and} \quad \frac{\partial Var(\beta^*(\theta, N))}{\partial N} < 0.$$

4 Relative inefficiency in the German electricity distribution sector

4.1 Methodology

Consider M municipalities $j = 1, ..., M$, where each municipality governs exactly one public enterprise. These public enterprises transform p inputs $L^j \in \mathbb{R}_+$ into q outputs $y^j \in \mathbb{R}_+$ using a common production technology Ψ, which is defined as $\Psi = \left\{(L, y) \in \mathbb{R}_+^{p+q} \mid L \text{ can produce } y\right\}$. The boundary of Ψ, denoted by Ψ^δ, is the production frontier against which each observation is compared to. In practice, however, this efficient technology cannot be observed, and thus, needs to be estimated from an observed sample of input-output-combinations.

For the estimation of the technology set and the company-specific efficiency scores, we apply the conditional free disposal hull (FDH) proposed by Daraio and Simar (2005). This approach allows accounting for external factors, i.e. factors that do not belong to the set of inputs and outputs. Considering external factors is important, because they potentially influence the performance measure of the public enterprises. In general, there are two ways in which these factors, also referred to as z-variables, may impact the estimates. They can affect i) the attainable set and/or ii) the distribution of the relative efficiency estimates (Bădin et al., 2012).

Z-variables induce a frontier shift if they influence the attainable production set. For some of the public enterprises this could yield in input-output-combinations that are unfeasible for them to reach. Thus, the estimated relative efficiency measures that lack an economic meaning when z-variables are neglected in the determination of efficiency. However, if the attainable set is not influenced by the external factors, the so called *separability condition* holds (Bădin et al., 2012). The potential impact, i.e. the impact of z-variables on the distribution of inefficiency, is of particular interest in this paper as it allows actually explaining differences of the public enterprises' performance.[9]

To present the conditional FDH estimator formally, we rely on the probabilistic formulation of the production process introduced by Cazals et al. (2002). It examines the probability of an observation (L, y) to be dominated by other observations. We focus on input-oriented efficiency estimation. In this context, a public enterprise is considered dominated if another unit produces at least as much output while using no more of any input (Simar and Wilson, 2008). Due to conditioning the production process, thereby, only companies with similar environments are compared to each other. Opposite to various approaches that incorporate heterogeneity, this approach takes z-variables into account without making assumptions about their influence on the attainable input-output set, i.e. no *a priori* assumptions regarding the *separability* are required (De Witte and Kortelainen, 2009). The conditional joint probability function (representing the production process) and its decompositions are given by

$$\begin{aligned} H_{\mathcal{L}Y|Z}(L, y|z) &= Prob\left(\mathcal{L} \leq L, Y \geq y | Z = z\right) \\ &= Prob\left(\mathcal{L} \leq L | Y \geq y, Z = z\right) Prob\left(Y \geq y | Z = z\right) \\ &= G_{\mathcal{L}|Y,Z}(L|y, z) S_{Y|Z}(y|z) \end{aligned} \tag{5}$$

where Z denotes the vector of z-variables, $G_{\mathcal{L}|Y,Z}$ the conditional distribution of X on Y and Z, and $S_{Y|Z}$ the conditional survivor function of Y (Daraio and Simar, 2007a). The support of $G_{\mathcal{L}|Y,Z}(\cdot|\, y, z)$ can be defined as the attainable set taking z-variables into account (Daraio and Simar, 2007b). This can be used to estimate the minimum amount of inputs for output level y given $Z = z$. Under the assumption of free disposability of the conditional production set, denoted by Ψ^z, the lower boundary of $G_{\mathcal{L}|Y,Z}$ can be considered as the input-oriented Debreu-Farrell measure of efficiency β.

[9] Simar and Wilson (2007) and Bădin et al. (2012) point out that it is important to take the implications of *separability* (and in particular its absence) into account when analyzing the effects of z-variables on estimated relative efficiency scores. If *separability* does not hold many standard regression techniques can not be applied.

For a unit producing output level y with operating conditions decribed by z, the according conditional FDH estimate of relative efficiency, denoted by $\hat{\beta}(L, y|z)$, can be expressed as

$$\hat{\beta}(L, y|z) = inf\left\{\beta \,|\, \hat{G}_{\mathcal{L}|Y,Z,n}\,(\beta L \,|\, y, z) > 0\right\} \tag{6}$$

where the subscript n in $\hat{G}_{\mathcal{L}|Y,Z,n}$ indicates that its obtained empirically. Since FDH is a full-frontier approach, i.e. $Prob((x, y|z) \in \Psi^z) = 1$, efficiency is bound to unity (Simar and Wilson, 2008). A production unit is considered efficient if $\hat{\beta}(L, y|z) = 1$. Smaller values indicate inefficiency and the amount by which the inputs can be reduced proportionally.[10]

Conditioning the production process to a given set of z-variables can be done by Kernel estimation. This approach assigns probabilities to the companies that reflect the similarity of their operational environments. The more similar the values of a particular z-variable between two companies are, the higher are their respective kernel probabilities and the more likely both companies will be selected in each others reference sets.

Since we only consider a continuous z-variable in our empirical application we use the Epanechnikov kernel to estimate the kernel probabilities as proposed for example by Daraio and Simar (2005). Following Bădin et al. (2010) we further rely on a company-specific and data-driven bandwidth-selection methods based on Hall et al. (2004) and Li and Racine (2007, 2008), which is a crucial component of the Kernel smoothing procedure.[11]

Following Bădin et al. (2012), we apply nonparametric regression techniques to identify the average effect of the z-variable on the distribution of efficiency and the corresponding variance. The regression model can be written as

$$\hat{\beta}(\mathcal{L}, y|z) = \mu(z) + \sigma(z)\epsilon \tag{7}$$

where $\mu(z)$ and $\sigma(z)$ are the average and variance functions; ϵ is the error term. While the first function reveals insights on the average effect of z on the conditional efficiencies, the second involves information on the dispersion of the efficiency distribution as a function of z (Bădin et al., 2012). For the regression we apply the local constant regression type.

4.2 Data and Variable Selection

For our empirical analysis we combine data from two publicly available sources for the year 2005. The first source, i.e. VWEW Energieverlag (publisher of the German energy

[10]Note that in our application public enterprises only optimize over labor as we treat the second input, i.e. capital, as quasi-fixed.

[11]An advantage of this method is that it does not rely on the separability condition. Hence, assuming separability is avoided in both the conditional efficiency estimation and the bandwidths selection.

industry VWEW (2006)), provides information on inputs and outputs necessary to model the production process of the electricity distribution companies. Given the main purpose of this paper, we thereby only include companies that are fully publicly owned. The second source, i.e. ...., contains regional data on working population. We use the municipal code number to merge the data on firms and population. This is justified since public energy companies mainly operate in the area of their registered office. In total our sample consists of 108 observations.

Regarding the selection of variables representing the underlying production process of the electricity distribution companies we follow Cullmann (2010). The characteristics of our data we use are shown by Table I. We use two input variables, i.e. labor (L_L) and capital L. Labor input is defined by the number of employees. Unfortunately the data on employees is not given for the different layers of the value-added chain separately. We account for the fact that some of the firms are active in electricity distribution and generation by subtracting one employee per 20 GWh of produced electricity from the total number of employees. This approximation is proposed by Auer (2002) and overcomes the mentioned issue.

Capital input is measured by the length of electricity lines and cables in kilometers. In order to capture different costs associated with the individual voltage levels (high, median and low voltage), each of these types are weighted using factors by the German network association (VDN). High voltage lines and cables are multiplied by the factor 5; medium and low voltage lines and cables by the factors 1.6 and 1, respectively. Even though, our empirical production model contains a capital input, it is still coherent with the presented theoretical model since, in our estimation, we consider capital as quasi-fixed. Therefore, public enterprises are only allowed to optimize over labor.

The plurality of outputs in network industries is well acknowledged. Therefore, two output variables are chosen. The first one represents the annual amount of electricity delivered to all final customers (y_D) measured in MWh.[12] The second output variable is the total number of customers (y_C) including households, industrial and other customers.

Furthermore, we use the number of inhabitants aged between 15 and 65 to capture the employable part of the population (z_P). By this variable we identify the operating environment of each company in the sample and compare their performance only among municipalities with a similar number of employable inhabitants. In addition, this variable is used to explain differences in the in the estimated relative efficiency among the companies.

[12]Final customers include households, industrial customers and others.

Table I: Descriptive Statistics for German Electricity Distribution Companies

Variable	Name	Min	Mean	Median	Max	Std.dev.
Input						
Employees [number]	$\hat{L}_L$	1	64	31	570	92.8
Cables and lines [kilometers]	$\hat{L}_C$	44.1	776.8	454.1	7526.4	1118.3
Outputs						
Electricity delivered [MWh]	y_D	10126	215533	97128	2920310	357608.3
Final customers [number]	y_C	1780	27770	13639	395312	46124.7
Continuous external variables						
Employable population [inhabitants$^{(1)}$]	z_P	887	27686	17281	258640	35543.6

Sources: *Statistisches Bundesamt, VWEW Energieverlag.*
Notes: observations=108, year=2005. $^{(1)}$Population between the ages 15 and 65.

4.3 Empirical Results

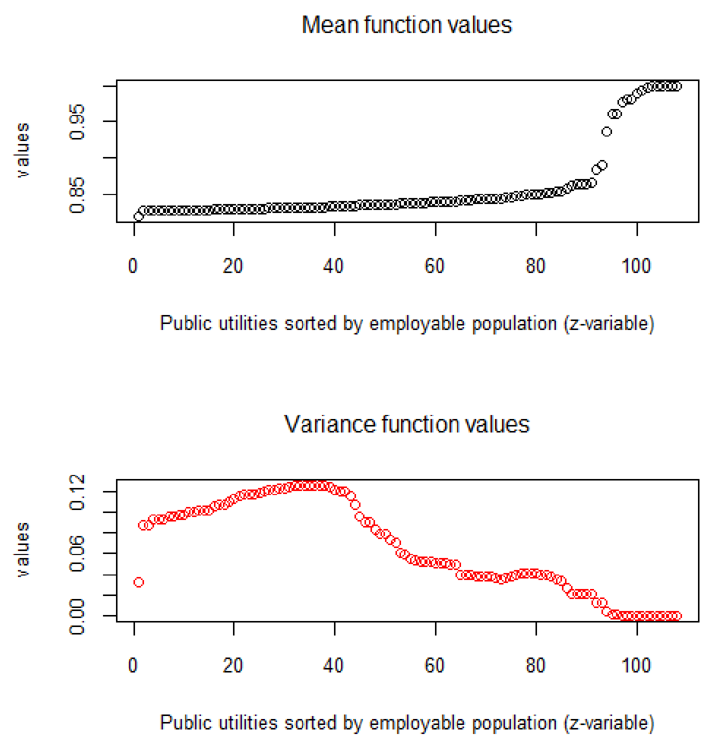
Our estimation results are shown in Figure 1.[13] The upper panel represents the mean function obtained by regressing the individual conditional efficiency estimates for the public utilities on the z-variable *unemployable population*. The public enterprises are sorted with respect to the z-variable in increasing order, thus, the firms of municipalities with larger values of z are located on the right side of the x-axis. Therefore, we can observe that *unemployable population* has on average a positive and increasing impact on efficiency, i.e. public enterprises are more efficient in municipalities with larger employable population.

Similarly, the variance function plotted in the lower panel of Figure 1 can be interpreted. Clearly, the variance of efficiency first slight increases for smaller values of z but then notably decreases for larger values. In general, the variance between firms in municipalities with smaller z is much bigger than for firms in municipalities with larger z.

From these findings we can draw the conclusions that i) the employable population positively effects the performance of public electricity distribution companies where ii) the companies with higher employable population perform more efficient. Therefore, we find empirical evidence for our presented voter model.

[13]In order to interpret the results properly, we investigated a potential frontier shift caused by *employable population*. This is done by computing the unconditional FDH estimates and compare them with their conditional counterparts (Bădin et al., 2012). Our results do not suggest that *employable population* causes a frontier shift. Therefore, we can interpret the impact of *employable population* on the performance of the public enterprises as an impact on the distribution of inefficiency.

Figure 1: Second-stage regression



5 Conclusion

In this article, we formalize the idea of 'rational inefficiency' in public firms, meaning that productive inefficiency is actively initiated by local governments which seek to win elections. This analysis reveals that large communities tend to maintain rather efficient enterprises on average than do smaller communities. Further, we expect efficiency variance to decrease in community size.

Nonparametric efficiency methods estimate the relative efficiency of a firm, i.e. the efficiency relative to the most efficient firm in a group. This means that if the most efficient firm in a group is not efficient itself – and this is what our model predicts – relative and absolute efficiency are not identical. The relative efficiency measure can be derived from the absolute. The voting model predicts that correlation between communitiy size and absolute or relative efficiency measures respectively should show the same sign (positive), although the effect is stronger for absolute efficiency measures. The same relation holds for the variance measures (which show however a negative correlation with community size). We have empirically tested our hypothesis using data from the German electricity distribution sector and this investigation supports our predictions.

Clearly, this analysis is not conclusive. In particular, more work should be devoted to analyzing the implications of unemployment. The abstraction from unemployment in our model is without loss of generality meaning that unemployment rates should not have any effect on the party's voting programs here. This hypothesis however remains to be tested, especially since it appears very unintuitive on first sight.

References

Auer, *Benchmarking und Regulation elektrischer Netze in liberalisierten Strommärkten: Grundlagen, internationale Erfahrungen und Anwendungen auf Österreich*, Technische Universität Wien, Institut für Elektirsche Anlagen und Energiewirtschaft, 2002.

Bădin, L., C. Daraio, and L. Simar, "How to measure the impact of environmental factors in a nonparametric production model?," *European Journal of Operational Research*, 2012, *223* (3), 818–833.

Bădin, Luiza, Cinzia Daraio, and Léopold Simar, "Optimal bandwidth selection for conditional efficiency measures: A data-driven approach," *European Journal of Operational Research*, March 2010, *201* (2), 633–640.

Cazals, Catherine, Jean-Pierre Florens, and Léopold Simar, "Nonparametric frontier estimation: A robust approach," *Journal of Econometrics*, 2002, *106* (1), 1 – 25.

Cullmann, A., "Benchmarking and firm heterogeneity: a latent class analysis for German electricity distribution companies," *Empirical Economics*, 2010, *42* (1), 147–169.

Daraio, Cinzia and Léopold Simar, "Introducing Environmental Variables in Nonparametric Frontier Models: A Probabilistic Approach," *Journal of Productivity Analysis*, 09 2005, *24* (1), 93–121.

_ **and** _ , *Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications* Studies in productivity and efficiency, New York: Springer, 2007.

_ **and** _ , "Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach," *Journal of Productivity Analysis*, October 2007, *28* (1), 13–32.

Duncombe, Walliam, Jerry Miner, and John Ruggiero, "Empirical evalutaion of bureaucratic models of inefficiency," *Public Choice*, 1997, *93* (1-2), 1–18.

Hall, Peter, Jeff Racine, and Qi Li, "Cross-validation and the Estimation of Conditional Probabilitiy Densities," *Journal of the American Statistical Association*, 2004, *99* (468), 1015–1026.

Li, Qi and Jeffrey S. Racine, *Nonparametric econometrics: Theory and practice*, Princeton: Princeton University Press, 2007.

_ **and** _ , “Nonparametric Estimation of Conditional CDF and Quantile Functions With Mixed Categorical and Continuous Data,” *Journal of Business and Economic Statistics*, 2008, *26* (4), 423–434.

Niskanen, W.A., *Bureaucracy and Representative Governments*, Aldine, Chicage, 1971.

Shleifer, A. and R. W. Vishny, “Politicians and Firms,” *Quarterly Journal of Economics*, 1994, *109* (4), 995–1025.

Simar, Léopold and Paul W. Wilson, “Estimation and inference in two-stage, semi-parametric models of production processes,” *Journal of Econometrics*, January 2007, *136* (1), 31–64.

_ **and** _ , “Statistical Inference in Nonparametric Frontier Models: Recent Developments and Perspectives,” in Harold O. Fried, C.A. Knox Lovell, and Shelton S. Schmidt, eds., *The measurement of productive efficiency and productivity growth*, Oxford: Oxford University Press, 2008, pp. 421–521.

VWEW, *Jahresdaten der Stromversorgung*, Frankfurt am Main: VWEW Energieverlag, Verlags- und Kommunikationsgesellschaft der Energiewirtschaft, 2006.

Witte, Kristof De and Mika Kortelainen, “Blaming the exogenous environment? Conditional efficiency estimation with continuous and discrete exogenous variables,” MPRA Paper No. 14034, University Library of Munich, Munich March 2009.

6 Appendix

6.1 Proofs

Proof of Lemma 1. The probability that party i wins the election is given by

$$P_i(\cdot) = \frac{1}{2} + \delta \int_0^1 \left[\frac{L^e(N)}{N} \left(\begin{array}{c} \frac{1}{\hat{\alpha}_i(\theta)} \left(U^I\left(\hat{\alpha}_i(\theta)\right) + w \right) \\ -\frac{1}{\hat{\alpha}_j(\theta)} \left(U^I\left(\hat{\alpha}_j(\theta)\right) + w \right) \end{array} \right) - \mu(\tilde{t}_i(\theta, N) - \tilde{t}_j(\theta, N)) \right] d\theta \quad (8)$$

for $j \neq i$. Parties maximize this probability, where

$$\max_{\hat{\alpha}_i(\theta)} P_i(\cdot) \Leftrightarrow \max \Gamma_i(\cdot) \quad (9)$$

$$\text{with } \Gamma_i(\cdot) := \max \int_0^1 \left[\frac{L^e(N)}{\hat{\alpha}_i(\theta) N} \left[U^I(\hat{\alpha}_i(\theta)) - w(\mu - 1) \right] - \mu \hat{\zeta}_i(\hat{\alpha}_i(\theta), \theta) \right] d\theta.$$

The first order condition is then attained by pointwise maximization of equation (8). An optimal incentive scheme $\alpha^*(\theta, N)$ for both parties is hence implicitly defined by

$$\begin{aligned}
&\Omega(\hat{\alpha}_i(\theta)) = 0 \quad \text{with} \\
&\Omega(\alpha^*(\theta,N)) := \frac{L^e(N)}{\alpha^*(\theta,N)} \quad \Big(U^{I\prime}\left(\alpha^*(\theta,N)\right) - \frac{1}{\alpha^*(\theta,N)}\left(U^I\left(\alpha^*(\theta,N)\right) - (\mu-1)w\right)\Big) \\
&\qquad\qquad\qquad -\mu C'(\alpha^*(\theta,N))\Big(H(\theta) - H'(\theta)(1-\theta)\Big).
\end{aligned} \tag{10}$$

Further, for the limits of the first order condition holds the following:

$$\Omega(\cdot)|_{\hat{\alpha}_i=1} = L^e(N)\left(U^{I\prime}(1) - (\mu-1)w\right) - \mu C'(1)\left(H(\theta) - H'(\theta)(1-\theta)\right) < 0, \tag{11}$$

$$\lim_{\hat{\alpha}_i \to 0} \Omega(\cdot) = \lim_{\hat{\alpha}_i \to 0}\left[\frac{L^e(N)}{\hat{\alpha}_i^2}\left(-U^I\left(\hat{\alpha}_i\right) + \hat{\alpha}_i U^{I\prime}\left(\hat{\alpha}_i\right) + (\mu-1)w\right)\right] > 0, \tag{12}$$

where the sign of $\Omega(\cdot)|_{\hat{\alpha}_i=1}$ is negative by Assumption 1. By continuity of $\Omega(\cdot)$ in $\hat{\alpha}_i$, $\alpha^*(\theta, N) \in (0, 1]$ exists.

Furthermore, solving the $\Omega(\cdot) = 0$ for w and substituting into the second order condition leads to

$$\begin{aligned}
\left.\frac{\partial\Omega(\cdot)}{\partial\hat{\alpha}_i(\cdot)}\right|_{\hat{\alpha}_i(\cdot)=\alpha^*(\theta,N)} &= \frac{1}{\alpha^*(\theta,N)}\left(L^e(N)U^{I\prime\prime}\left(\alpha^*(\theta,N)\right) - \Theta\right) \\
\text{with } \Theta := \mu&\Big(H(\theta) - (1-\theta)H'(\theta)\Big)\Big(2C'(\alpha^*(\theta,N)) + \alpha^*(\theta,N)C''(\alpha^*(\theta,N))\Big).
\end{aligned} \tag{13}$$

By Assumption 1, this expression is strictly negative. Hence, $\alpha^*(\theta, N)$ is also unique.

Finally, solving $\Omega(\cdot) = 0$ for w and substituting into the partial derivatives $\partial\Omega(\cdot)/\partial\theta$ we obtain

$$sign\ \frac{\partial\alpha^*(\theta,N)}{\partial\theta} = sign\left[-\mu C'(\alpha^*(\theta,N))(2H'(\theta) - (1-\theta)H''(\theta))\right] \tag{14}$$

and thus $\partial\alpha^*(\theta, N)/\partial\theta > 0$ by Assumption 1.

Proof of Proposition 1. Solving $\Omega(\cdot) = 0$ for w and substituting it into the partial derivatives $\partial\Omega(\cdot)/\partial N$ leads to

$$sign\ \frac{\partial\alpha^*(\theta,N)}{\partial N} = sign\left[\mu C'(\alpha^*(\theta,N))L^{e\prime}(N)(H(\theta) - (1-\theta)H'(\theta))\right]$$

and thus to $\partial\alpha^*(\theta, N)/\partial N > 0$.

Proof of Lemma 2 $\frac{\partial}{\partial N}\left(\partial\alpha^*(\theta,N)/\,\partial\theta\right)$ is obtained by setting $\Omega(\alpha^*(\theta,N))=0$. It is given by

$$\begin{aligned}\frac{\partial}{\partial N}\left(\frac{\partial\alpha^*(\theta,N)}{\partial\theta}\right)=&\frac{\mu C'(\alpha^*)L^{e\prime}(N)(2H'(\theta)-(1-\theta)H''(\theta))}{\alpha^*(\theta,N)^2L^e(N)\Omega_\alpha(\alpha^*)^3}\\&\left[\Delta_0+L^e(N)\left(\Delta_1U^{I\prime\prime}(\alpha^*)+\Delta_2U^{I\prime\prime\prime}(\alpha^*)\right)\right]\\ \text{with}:\quad\Delta_0=&\mu^2(H(\theta)-(1-\theta)H'(\theta)^2)\\&\left[2C'(\alpha^*)^2+\alpha^*(\theta,N)^2\left(C''^2(\alpha^*)-C'(\alpha^*)C'''(\alpha^*)\right)\right]\\ \Delta_1=&\mu C'(\alpha^*)(H(\theta)-(1-\theta)H'(\theta))-L^eU^{I\prime\prime}(\alpha^*)\\ \Delta_2=&\alpha^*(\theta,N)\mu C'(\alpha^*)(H(\theta)-(1-\theta)H'(\theta))\end{aligned}$$

and thus $\frac{\partial}{\partial N}\left(\partial\alpha(\theta,N)/\,\partial\theta\right)<0$ as long as $\Delta_0+L^e\left(\Delta_1U^{I\prime\prime}+\Delta_2U^{I\prime\prime\prime}\right)<0$. Since $\Delta_1,\Delta_2>0$ the result follows from Assumption 2.

Proof of Proposition 3.

$$\begin{aligned}Var(\alpha^*(\theta,N))=&\int_0^1\alpha^*(\theta,N)^2d\theta-E[\alpha^*(\theta,N)]^2\\ \text{with }E[\alpha^*(\theta,N)]=&\int_0^1\alpha^*(\theta,N)d\theta,\end{aligned}$$

and thus

$$\frac{\partial Var(\alpha^*(\theta,N))}{\partial N}=2\int_0^1\left(\alpha^*(\theta,N)-E[\alpha^*(\theta,N)]\right)\frac{\partial\alpha^*(\theta,N)}{\partial N}d\theta<0.$$

This expression is negative since $\int_0^1\left(\alpha^*(\theta,N)-E[\alpha,N]\right)d\theta=0,\ \ \partial\alpha^*(\theta,N)/\,\partial N>0$ and $\frac{\partial}{\partial N}\left(\partial\alpha^*(\theta,N)/\,\partial\theta\right)<0$.