COST EFFECTIVENESS: THE FORGOTTEN DIMENSION OF PUBLIC SECTOR PERFORMANCE

Hans de Groot (Innovation and Governance Studies, University of Twente)

Bart L. van Hulst (Innovation and Public Sector Efficiency Studies, Delft University of Technology)

1. Introduction

The performance of public organizations is one of the key topics in public administration research (for an overview, see Walker and Boyne, 2009). The main focus has been on the process of performance measurement and management and typically challenges the assumptions and outcomes of 'new public management' and related reforms. A smaller part of the research effort addresses the determinants of public performance; however often without calculating the size of impacts on public management. In particular, the dimension of cost effectiveness is usually overlooked. This paper aims at partly filling this gap. Its relevance is not only derived from academic criteria but also from the practitioner's urgent need for results on the cost effectiveness of public programs. Many governments currently are cutting budgets on an unprecedented scale after the world-wide financial and economic crisis. Reducing budgets and safeguarding the level of essential public services as much as possible, requires thorough knowledge of their cost effectiveness.

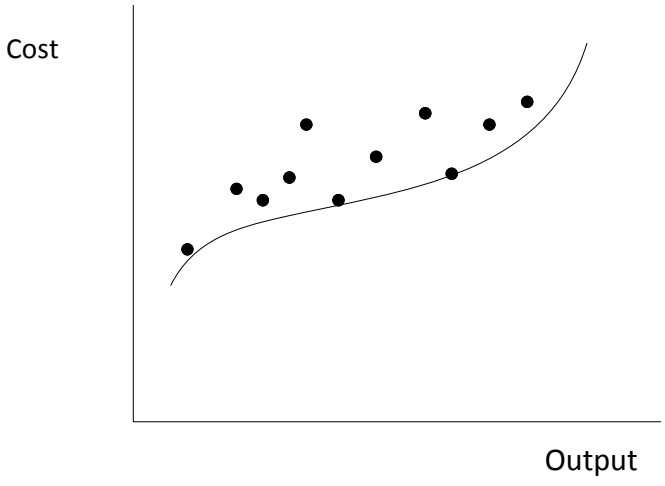
The methodology of the paper closely follows the evidence-based research agenda put forward by, among others, Meier and O'Toole (2009). The paper adds new insight by focusing on measures of cost effectiveness and efficiency, usually not considered in current research. An empirical illustration is given by studying the efficiency of administrative services of local government and its determinants, using a large sample of around 100 Dutch municipalities for three years, employing data on costs, outputs and organizational characteristics. The empirical illustration is of particular interest because it links costs to physical outputs in terms of delivered services. This differs from previous research that proxies physical output by the size of the population (for example, Borger et al (2000) and Kalb (2010).

2. Theoretical considerations

Current approaches to public sector performance in the context of the evidence-based research agenda mentioned before typically try to explain program performance as a function of independent variables that measure product or client characteristics, environmental influences and organizational and managerial variables such as the extent of decentralized decision making, the level of external networking by employees or managers, their intrinsic skills and qualities, available resources etc. Cost-effectiveness or efficiency of the institutions or programs involved is typically not considered. Although resources are

often included as a control variable (for example, the level of expenditures per student in educational programs), cost effectiveness is not addressed separately. This is an important omission, as policy makers usually are not interested in effectiveness per se – does the program work – but often in the question: how much effect is a program generating for each dollar spent. Or alternatively: is this program the least costly way of achieving the public objectives, or are their less costly alternatives? A typical question in education could be: is it more effective - in terms of student achievement - to invest in higher quality teachers or in smaller classes? Both measures can be shown to have some effect, but their cost-effectiveness is different. To answer the cost-effectiveness question, a more specific approach is needed. We argue that existing methods from econometrics and operations research can be used to measure and investigate the cost-effectiveness of public programs or institutions (for an overview, see Fried et al, 2008). Methods such as stochastic frontier analysis (SFA) and data envelopment analysis (DEA) determine a virtual frontier of relatively efficient decision making units (compare the solid line in figure 1). The parametric SFA frontier allows for random measurement error in the data as well as systematic deviations from the efficiency frontier arising from intrinsic inefficiencies. The non-parametric DEA frontier is completely determined by the existing data, but does not allow for measurement error. Therefore we prefer the more realistic approach of the parametric SFA frontier, although one has to assume a particular functional form. A widely used functional form is the translog function (Christensen et al, 1973). For our purpose, cost efficiency, i.e. relating total costs and outputs (or outcomes, in which case we speak of cost-effectiveness) is the most relevant efficiency indicator. Pure technical efficiency would only consider the relation between physical inputs, such as the amount of labor or capital, and outputs. When considering cost-efficiency, the usual assumption is that decision making units also take into account input prices in order to minimize costs. These input prices therefore are part of the frontier estimation exercise, unless prices do not vary substantially for different decision making units, in which case they can be neglected.

Figure 1 Example of an efficiency frontier



The distance to the efficiency frontier is a measure of the (relative) inefficiency of each decision making unit. In practice, an efficiency score is calculated for each unit from the proportional decrease in all inputs needed to reach the frontier at a given output. Depending on the scope of the analysis and available data, assumptions have to made regarding a possible intertemporal shift of the efficiency frontier, for instance as a result of technological progress. Typically, as a first step, cost-effectiveness or efficiency scores are determined as described, taking into account exogenous factors that also determine efficiency, such as client characteristics. In a second step, efficiency scores can be explained by regressing them on variables that measure management quality or other controllable organizational characteristics. In this stage, different assumptions from organizational or political theories of public sector behavior can be tested.

3. Empirical illustration

Our approach to measuring cost-effectiveness or efficiency in the public sector is empirically demonstrated using data on Dutch municipalities. Their large number, more than 400, and their freedom in choosing internal organizational structure and management instruments, make them well suited to analyze determinants of cost-effectiveness. The dataset is created in the context of a developing research program on evidence-based public management in local government of three Dutch universities. The dataset used for this paper, however, is limited to local administrative public services such as the provision of passports, documents from the registry and small building permits. Data are available on costs, organizational data and outputs, including survey results on citizen satisfaction with administrative municipal services. The extensive dataset used here is only available for a subset of around 100 municipalities for 3 years(a total of 229 usable observations)and is largely based on an annual benchmark exercise (BMC, 2010). Table 1 gives some descriptive statistics of key variables in the dataset for 2009.

Table 1. Descriptive statistics of key variables in 2009 (N=85)

Variable	Minimum	Maximum	Mean
Personnel costs	184	30,507	3,638
Population size	5,429	747,093	81,924
Passports issued	740	112,443	9,669
Identity cards issued	258	54,000	5,631
Driving licenses issued	309	77,067	8,979
Certificates of death, birth & marriage	119	57,807	3,566
Certificates of residence	348	239,965	8,762
Building permits	-	2,411	274

3.1 Frontier estimation results

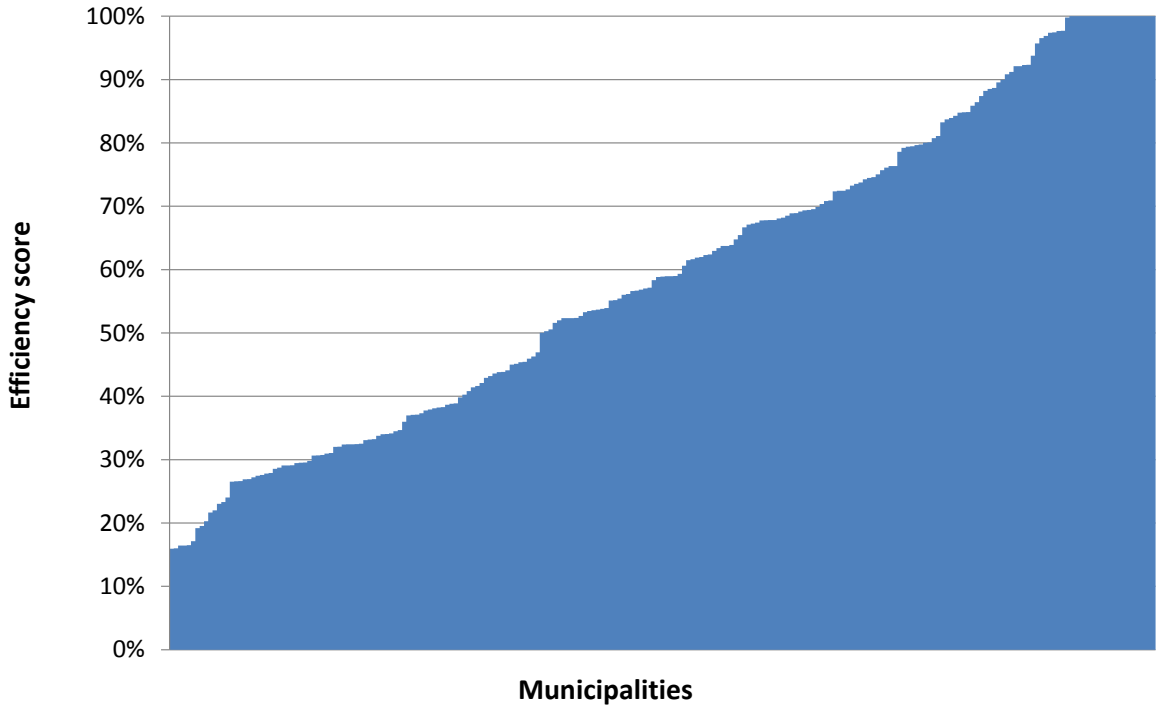
We have estimated a translog stochastic cost frontier, employing as the three main output variables (1) the sum of the number of passports, driving licenses and national identity cards,

(2) the sum of certificates of death, birth & marriage and (3) small building permits. The last output is considered separately, as the provision of building permits is a relatively costly service. It is worthwhile to consider additional exogenous variables that capture qualitative output characteristics that also affect service costs - and therefore efficiency - given the level of quantitative output. For example, a lower educated or immigrant population could require more time to be serviced and therefore generate higher costs. In this analysis we proxy these factors by the percentage of the population living in urbanized areas. Input price differences have been neglected, given nationally regulated local public sector wages and small regional price variation in the Netherlands. Estimated parameters for the frontier are presented in table 2, while efficiency scores are depicted in figure 2.

Table 2. Estimation results

Variable	**Parameter estimate**	**T-value**
Constant	**-0.94**	**-9.3**
2009	**0.19**	**2.3**
2010	**0.34**	**3.6**
Documents	**0.54**	**3.3**
Certificates	0.13	1.1
Permits	**0.29**	**2.4**
Documents x Documents	0.08	0.2
Documents x Certificates	-0.09	-0.3
Documents x Permits	0.11	0.4
Certificates x Certificates	0.16	0.7
Certificates x Permits	-0.10	-0.5
Permits x Permits	-0.01	0.0
% Urban	**0.17**	**2.3**
% Urban x %Urban	**0.06**	**3.2**
% Urban x Documents	0.004	0.1
% Urban x Certificates	-0.01	-0.5
% Urban x Permits	-0.04	-1.1
SIGMA	**1.17**	**11.1**
LAMBDA	**3.39**	**2.9**

Figure 2. Distribution of efficiency scores (N=229)



R-squared of the estimation is 0.72. We also see that most linear terms are significant and have the expected signs. Interaction and quadratic terms are in most of the cases not significant. This means that the translog specification is somewhat overdone and a loglinear Cobb Douglas specification might suffice. Sigma and Lambda are indicative parameters of the frontier analysis and are in an acceptable range.

From table 2 we can conclude that production of documents and permits are the most important cost drivers for the administrative services examined here. Production of certificates has a small impact on the costs. We also see that the percentage of the population living in urban areas has the expected positive and significant effect. Figure 2 shows the distribution of the efficiency scores, the efficiency scores ranges from 16% (lowest score) to 100% (fully efficient). About 10 percent of the observations are fully efficient. On average the efficiency scores are 60%, so there are possibilities to increase efficiency. In the second stage of our analysis we try to determine the factors that explain differences in efficiency.

3.2 Explaining efficiency scores

Before proceeding to the explanation of efficiency scores in terms of managerial and organizational characteristics, we present some statistics on efficiency scores for subgroups of municipalities (table 3).The quality of staff is measured by the average pay scale of front

office workers. In the table we distinguish three levels of staff quality, implying that 1/6, 2/3 and 1/6 of the municipalities have low, average and high quality staff respectively. The same applies to the level of absenteeism. We also have information on how the services are organized. Three types are distinguished, following recent developments in the provision of local administrative services. Some municipalities offer only traditional services through their central front office. Traditional services are typically limited to passports, identity cards, driving licenses and different certificates from the registry. Other municipalities provide extended services through their front office, either with or without integrating their back offices for those services. Extended services include more specialized services such as providing building permits and welfare benefits.

Table 3. Average efficiency scores for different groups of municipalities

Group description	Average efficiency score (* = p > 0.05)
Total average	**60%**
Small municipalities (< 50,000 inhabitants)	62%
Medium size municipalities (50-100.000 inh)	58%
Large municipalities (> 100,000 inhabitants)	55%
Low quality staff	70%*
Average quality staff	60%
High quality staff	44%*
Low level absenteeism	60%
Medium level absenteeism	55%
High level absenteeism	66%*
Extended services, integrated back offices	48%*
Extended services, separate back offices	65%*
Traditional services	57%
Low client satisfaction scores	57%
Average client satisfaction scores	57%
High client satisfaction scores	68%*

Table 3 shows the following differences in efficiency scores when comparing subgroups of municipalities. Population size does not seem to matter. Higher quality of staff, however, leads to lower efficiency levels. Apparently, possible efficiency gains by employing higher quality (= higher paid) staff do not outweigh their higher costs. Note that higher absenteeism is associated with higher efficiency. This is a perverse result: usually higher absenteeism leads to higher costs per output, as absent workers do not provide services and still have to be paid. We have no details on possible insurance contributions that compensate for those costs, but a net gain in efficiency seems unlikely. Municipalities providing extented services using separate backoffices show a more than average efficiency, while the reverse is true for those with integrated back offices. Interestingly, municipalities with high client satisfaction scores also show high efficiency scores. Apparently, there is not necessary a trade off

between quality of services and efficiency, as sometimes claimed. Meier and O'Toole (2009) in their public education studies also frequently find positive spillovers from one goal to another due to effective management.

The efficiency scores determined in the first step of the analysis have been tentatively explained by managerial and organizational characteristics using multiple regression. In this exploratory study, we do not formulate explicit hypotheses on public sector managerial behavior. Instead, we include the potential explanatory variables as described above. Because our dependent variable, the efficiency score, theoretically ranges from 0% to 100%, we transform the dependent variable and then apply aTobit regression (Tobin, 1958). The transformation implies taking the reciprocal of the efficiency scores. Instead of the groups defined in table 3 we use the continuous equivalent of those variables when possible. Estimation results are presented in table 4. For comparison, and to facilitate the interpretation of results, ordinary least squares estimation results for the untransformed efficiency scores are included as well. Note that as a result of the transformation the interpretation of parameter estimates in the second stage is as follows: a negative sign means a positive effect on efficiency scores.

Table 4. Estimation results second stage Tobit regression

Variable	**Parameter estimate**	**T-value**
Size of municipality	0.03	0.49
Quality of staff	**9.32**	**6.05**
Level of absenteeism	-2.75	-1.12
Extented services with separate back offices as compared with integrated back offices	**-0.69**	**-3.95**
Traditional services as compared with extended services and integrated back offices	-0.33	-1.31
Client satisfaction	-0.38	-2.07
Sigma	**1.07**	**19.85**

Table 4 confirms the results already obtained from analyzing subgroups. Higher quality (= higher paid) staff does not payoff in higher efficiency. The perverse result for absenteeism also shows up here. The effects of organizational structure are again significant: separate back offices turn out to be more efficient than integrated back offices. Client satisfaction no longer has a significant, positive impact on efficiency (although it is close to being significant), but also no negative impact. That still implies that there is no tradeoff between appreciation of services and efficiency.

Interpretation of the impact of the variables on efficiency is a bit hard due to the mentioned transformation. However the continuous explanatory variables are standardized, so the

differences in magnitude of the parameter estimates give an indication of the impact. For example, a municipality with an average efficiency score and average staff quality is predicted to have a 6% (3.6 percent point) higher efficiency score if the quality of staff is reduced by 1%. If client satisfaction increases with 1%, an increase of only 0.2% (0.1 percent point) for the efficiency score is predicted.

A separate discussion is warranted for possible economies of scale. Does increase of outputs lead to less than proportional costs? We have tested the sum of estimated parameters for the output variables against the hypothesis that the sum is not significantly different from one – that is: no economies of scale. That hypothesis could not be rejected. Note, however, that our cost measure only involves personnel cost. We cannot exclude the possibility that using a broader definition of inputs, including material costs, would lead to a different picture.

4. Summary and conclusion

Public sector performance is one of the key topics in public administration. This paper discusses an often forgotten dimension: cost-effectiveness or efficiency. In particular in the aftermath of the financial crisis and worldwide cutbacks on public expenditures, policy makers urgently need more of that type of performance information. We have shown that existing econometric methods can be used to model cost-effectiveness. As an empirical example, we studied the cost-efficiency of administrative services provided by Dutch municipalities. Both quantitative and qualitative output measures have been employed, instead of the usual population proxy. The results show a large variation of cost-efficiency scores, with an average of 60% compared with the most efficient municipalities. A tentative explanation of the cost-efficiency scores by organizational characteristics suggests a number of preliminary conclusions. Interestingly, higher client satisfaction scores do not imply lower efficiency scores, suggesting that there is not necessary a tradeoff between quality and efficiency. Higher quality of staff, on the other hand, does not pay off in terms of higher efficiency. Apparently, possible productivity gains are not enough to compensate for higher personnel costs.

In future research the complete package of Dutch municipal services will be analyzed with respect to its efficiency or cost-effectiveness, as has been done before on a limited scale for other countries (compare Borger et al (2000) for Belgium and Kalb (2010) for Germany), however employing simple output proxies. Additional data on managerial quality and managerial instruments, such as outsourcing service delivery to third parties, networking efforts, etc., will be collected to gain more insight in the determinants of cost-effectiveness of local public services. The analysis of the complete public service package will also be related to political preferences for spending and taxation levels.

References:

Borger de, B., &Kerstens, K. (2000). What is known about municipal efficiency? The Belgian case and beyond. In J. L. T. Blank (Ed.), Public provision and performance: contributions from efficiency and productivity measurement (pp. 299-330). Amsterdam: Elsevier.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental Logarithmic Production Frontiers, *The Review of Economics and Statistics*, 55(1), 28-45.

Fried, H. O., Lovell, C. A. K., & Schmidt, S. S. (2008).The measurement of productive efficiency and productivity growth. New York: Oxford University Press.

Kalb, A. (2010). Public Sector Efficiency. Applications to local governments in Germany. Heidelberg: Universität Heidelberg.

Meier, K.J. and L. J. O'Toole Jr. (2009), The proverbs of New Public Management: lessons from an evidence-based research agenda, *The American Review of Public Administration*, Vol. 39, No. 1, 4-22.

BMC (2010). Benchmarking Publiekszaken 2010. Den Haag: SGBO/BMC.

Tobin , J. (1958). Estimation of relationships for limited dependent variables, *Econometrica* 31 (1958), 24-36.

Walker, R.M. and G.A. Boyne (2009), Introduction: determinants of performance in public organizations, *Public Administration* vol. 87, No. 3, 433-439.