ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF ECONOMICS AND COMMERCIAL LAW
GÖTEBORG UNIVERSITY
116

ESSAYS ON THE PERFORMANCE OF FIRE AND RESCUE SERVICES

Henrik Jaldell

Essays on the Performance of Fire and Rescue Services

Henrik Jaldell

Abstract

This thesis consists of an introduction and five separate papers. All papers deal with measuring performance for the Swedish fire and rescue services. The first paper describes the production process of fire safety, while Papers 2-5 are empirical analyses measuring performance for some aspects of that process.

Paper 1 (*The problems of defining outputs in the public sector's service production - a discussion with an application to the fire service*) discusses how outputs and inputs should be measured in the public sector, and how they could be used in productivity and efficiency studies. There are two different levels where the studies could be performed: the vertical and the horizontal levels. The vertical level is the distinction between performing studies on the macro, the national, and the within-unit level. The horizontal level is the distinction between different outputs, whether determinant variables, direct outputs or consequences are used. The paper also includes an application of these ideas to the fire service.

Paper 2 (*Swedish fire and rescue services' manning levels - a stochastic frontier analysis using panel data*) studies the productivity and efficiency of the Swedish fire and rescue services during 1989-1995 using a stochastic frontier specification for panel data. The manning level is modelled as a function of risk, environment, and number of turn-outs. The results show that the size of population was the main determinant of manning levels. No productivity change was discovered. The efficiency differences found were substantial with a mean input saving potential of 30%.

In Paper 3 (*Measuring the efficiency of Swedish fire services' stand-by level*), the DEA-model is used to find efficiency scores and returns to scale corrected for environmental variables. The paper studies the stand-by level of Swedish fire services. This level has two output dimensions 1) the turn-out time (the faster the better), measured as number of people reached within five and ten minutes, and 2) the suppressing power, measured by the total number of firemen turning out (the more the better). The empirical results show that the long run input saving potential is about 30%.

In Paper 4 (*Productivity change of Swedish Fire Services between 1992 and 1998*), Malmquist productivity indexes are used to find out how productivity has changed among Swedish fire services between 1992 and 1998. The paper studies the stand-by level and the empirical results show that productivity has decreased for full- and mixed-time fire services. Less input used has resulted in less output produced.

Paper 5 (*Measuring performance differences using an ordinal output variable: The case of Swedish fire services*) investigates how to find performance differences in fire services with an ordinal output variable. Performance is measured by adjusting the outputs for inputs using the ordered probit model. No performance differences were found between full-time and part-time firemen for fires in detached houses. The results also indicate that "team spirit" is more important for performance than the actual number of firemen fighting a fire.

Keywords: efficiency, productivity, public sector, intermediate output, panel data, stochastic frontier, data envelopment analysis (DEA), Malmquist, ordinal output, ordered probit

Address: Henrik Jaldell, Karlstad University, SE-651 88 Karlstad, Sweden.
Tel. +46-54-7001369, e-mail: Henrik.Jaldell@kau.se

Acknowledgements

Many friends and colleagues have helped me with the writing of this thesis. Thank you all! Special thanks go to:

Lennart Hjalmarsson, my supervisor, for your valuable comments and advice.

The reference group at the Swedish Rescue Services Agency (Räddningsverket): Sven-Erik Frödin for financial support and for never giving up on me despite the (almost) never ending story of my thesis, but instead encouraging me to continue working. Birgitta Juås for suggesting the problem of the thesis. It was harder than expected! Bengt Martinsson for your stubborn questioning which spurred me to understand the role of the fire and rescue services. Fredric Jonsson for your will to read and understand endless tables of numbers. Bengt Mattsson and Björn Sund for your valuable comments, and for the help with defending me, and economic science, in the group. Dick Sträng for providing data for Paper 3 and 4. Göran Melin, since during a discussion with you I really got the picture of what the fire and rescue services really are producing.

Almas Heshmati and George Battese for valuable comments to Paper 2 and Paper 5 (Almas).

I must also (again) thank Bengt Mattsson and Birgitta Juås for their excellent economic courses in Karlstad that inspired me to carry on to postgraduate studies.

Henrik Jaldell
Karlstad, April 2002

This thesis consists of an introduction and five papers:

1. **The problems of defining outputs in the public sector's service production – a discussion with an application to the fire service.**

2. **Swedish fire and rescue services' manning levels – a stochastic frontier analysis using panel data.**

3. **Measuring the efficiency of Swedish fire services' stand-by level.**

4. **Productivity change of Swedish fire services between 1992 and 1998.**

5. **Measuring performance differences using an ordinal output variable: The case of Swedish fire services.**

Introduction

1 Background

The analyses in this thesis have originated from two reports about the Swedish Fire Service, which evaluated the marginal product of firemen and the marginal product of response time (Juås, 1994 and 1995). The purpose of those reports was to use production functions to examine these issues. However, due to data problems and difficulties in defining the production process within the production theory of economics, this was not done. Instead, costs-benefit analyses were employed.

What, then, are the specific problems with the production of fire safety (Juås, 1994, p. 15, 94-95)? The first is that fire is a continuous process. The task of the fire brigade is to change this process. The output should be the difference between what could have happened, the potential damage, and what actually happened. The second problem is that the outcome of the firemen's work can differ among buildings with the same response time, even though the same number of firemen and trucks are used. Factors such as the weather, building material and the construction are very important to how the fire develops. A third problem is that the fire service mostly produces a stand-by service and the question is how can this "waiting for something to happen" be valued? A fourth problem is that the fire services are engaged in many different kinds of operations except fires, such as drowning accidents, traffic accidents, storms, floods etc. A fifth problem is that there is a lack of data for both inputs and outputs.

At first, my intention was to use production functions to analyse the Swedish fire service while taking the above problems into account. However, I found it more appropriate to use frontier functions, which makes it possible to measure not only productivity, but also efficiency.

Therefore the purpose of this thesis is to develop production frontier models that are fit for the Swedish fire services, and to evaluate efficiency and productivity. An attempt is also made to solve some of the above problems about the production process of fire safety.

Fire safety is a complicated production process, and unfortunately very few economic studies have looked at the fire service before. (A survey is presented in the Appendix to Paper 1.) This thesis can therefore be seen as a first attempt to try to measure efficiency and productivity for the fire service using production frontier techniques.

2 Efficiency and productivity

Production frontier estimation is quite straightforward. Given an estimated frontier, inefficiency for each production unit can be measured as the distance to the frontier, and given observations over time productivity changes can be also be measured. However, there are two main problems: 1) Defining inputs and outputs, and 2) Determining the production frontier technique to be used.

The first question is discussed thoroughly in Paper 1. The second question has many dimensions. A choice must be made between assuming a convex or non-convex technology, between estimating a parametric or non-parametric frontier, between estimating a deterministic or stochastic frontier, and between using econometric techniques or using linear programming. Two additional questions are 1) What assumption should be made about statistical noise when choosing a stochastic frontier? 2) What model should be used when having data over both time and by cross-section (panel data)?

The differences between the specifications and some of their advantages and disadvantages will be described below. However, the examples here do not cover all possible aspects.[1]

2.1 Convexity/Non-convexity

In analysing the efficiency and productivity of private firms and public services an assumption that there exists a best-practice technology must be made. An implicit assumption is that the best-used technology is almost the same as the best-practice technology, and therefore that empirical data can be used to find it (see Grosskopf, 1986). Setting up some reasonable axioms like profit maximisation, convexity and monotonicity, Varian (1984) described how an outer bound of the technology could be found. Given five units with different output/input combinations (*B*, *H*, *C*, *I* and *D*), the outer bound is *YO* in Figure 1. For all bundles within the outer bound, profits are at least as great as the profits at any other choices. Note that since *YO* starts at *A,* it implies a "free lunch".

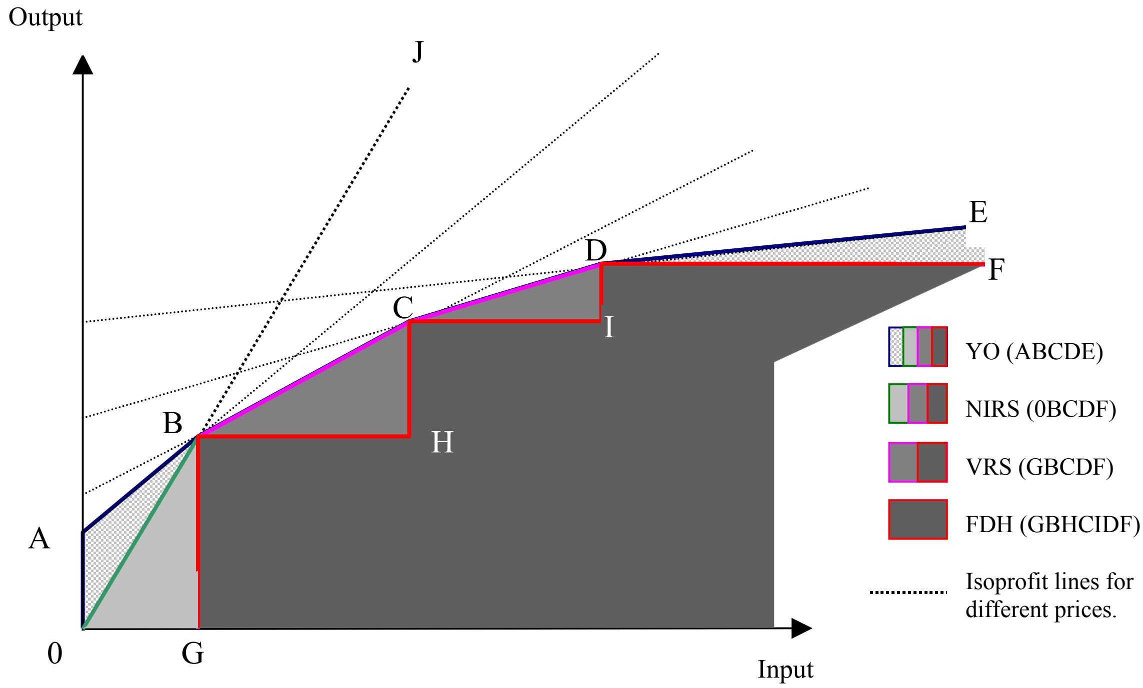
An inner bound of technology can instead be found by assuming non-convexity.[2] Production units are only compared to units with the same output or input level (Deprins et al., 1984). In the figure below this is called *FDH* (free disposal hull). A popular technique used to find the best-used technology is the data envelopment analysis, DEA (Charnes et al., 1978).[3] When using the DEA convexity, monotonicity and some scale property are assumed. By assuming that the returns to scale can vary from increasing to constant and to decreasing, we have variable returns to scale (*VRS)*. Note that the *VRS*-frontier implies set-up costs. By further assuming non-increasing returns to scale (*NIRS*), one moves further. The DEA technologies are thus situated between the outer and inner bound of the technology. The exception is if instead constant returns to scale (*CRS*) is assumed, *0BJ*. Note that the *CRS*-frontier normally is above the outer bound, *YO*.

[1] Recent surveys of production frontier techniques include Fried et al (1993) and Coelli et al (1998) covering most aspects, Färe et al (1994) and Cooper et al (1999) covering linear programming, and Kumbhakar and Lovell (2000) covering stochastic frontiers.

[2] Varian instead defined VRS as the inner bound.

[3] Färe and Li (1998) compared the inner and outer bounds for DEA.

Figure 1. Different frontier technologies



What reference technology should then be chosen? It depends on both the relevant assumptions about the technology and the data available. The *YO*-hull not only requires an assumption about profit maximisation (or cost minimisation) behaviour, but it also requires knowledge about both input and output prices. The rest (*VRS*, *NIRS* and *FDH*) do not require an assumption about profit maximisation behaviour or knowledge about input and output prices. *NIRS* implies non-negative profits, and the choice between *VRS* and *FDH* depends on the assumption of convexity. Non-convexity is more relevant in the public sector, since output is often given for different units by the environment. Thus, convexity may not be a relevant concept (see also Thrall, 1999, and Cherchye, Kuosmanen and Post, 2000).

None of the studies in this thesis has assumed a non-convex technology. The reason is that the input/output combinations for Swedish frontier fire services are not given. It should be possible to compare a real fire service to a convex combination of two or more frontier units. One risk of assuming convexity is that the hypothetical peer unit (the target) on the frontier probably consists of a combination of the largest and smallest units. Thus, there is a pedagogical problem because the manager of the in-between-sized fire service may not find it interesting to compare himself neither to the largest units, nor to the smallest units. The results may thus give no help to practical policy questions.

2.2 Parametric/Non-parametric, Deterministic/Stochastic, and Econometric/Linear programming

Using a parametric frontier means assuming a functional form for the function. Examples include the Cobb-Douglas function, the CES function, the Zellner-Revankar, the quadratic function, and the translog function. The advantage of using a specific form is that it gives a structure for the technology. Questions such as:

What are the marginal elasticities?
The elasticity of substitution?
The price elasticity of the factor demand functions?
The elasticity of scale?
The degree of homogenity?

are easily answered, since they are all defined with specific bounds in advance for the above functions.

For example, for the Cobb-Douglas, the elasticity of substitution is equal to 1, the price elasticity of factor demand is also equal to 1, and the elasticity of scale is constant. For the CES function elasticity of substitution is instead constant, and for the Zellner-Revankar function elasticity of scale is variable, but only a function of output, i.e. homotheticity is implied. The flexible translog function has both variable elasticity of substitution, variable elasticity of scale, and allows for non-homotheticity. The advantage of using a flexible form is that questions about the structure of the technology can be tested statistically. The disadvantage is that there is less degree of freedom when estimating the parameters.

The non-parametric frontier on the contrary lets the data points decide how the structure of the frontier looks. This frontier is measured by linear programming techniques. DEA is not only a non-parametric technique, but also a deterministic technique. There is no allowance for measurement errors or statistical noise.[4] Instead the advantage is that a functional form does not have to be decided in advance, since the model envelops the data tightly with linear segments, and a distribution for the inefficiency term does not have to be assumed.[5]

The question of choosing a parametric or non-parametric production frontier is only relevant when assuming a convex technology. Assuming a non-convex technology implies using a non-parametric production frontier.

2.3 Stochastic frontier and panel data

A stochastic frontier or composed error model (Aigner, Lovell and Schmidt, 1977, and Meussen and van den Broeck, 1977) divides the errors into two parts: inefficiency and white noise. A typical model is

$$y=f(x,\beta)+v-u \qquad (1)$$

where y is output, x is input vector, β are parameters to be estimated, u is the inefficiency term and v is the white noise.

To be able to estimate this function the distribution of the inefficiency term must be decided upon in advance. The question is, should it have a half-norm, exponential, truncated normal or gamma distribution? Ex ante, this question is difficult to answer because the technology is unknown before the analysis. Ex post, the distributions could be tested for from a statistical point of view. The functional form, f, must also be decided ex ante, even if some ex post statistical tests can rule out some functional forms (e.g. going from translog to Cobb-Douglas).

[4] A stochastic DEA is under construction, and deterministic parametric functions can be estimated using both linear programming and econometric methods.

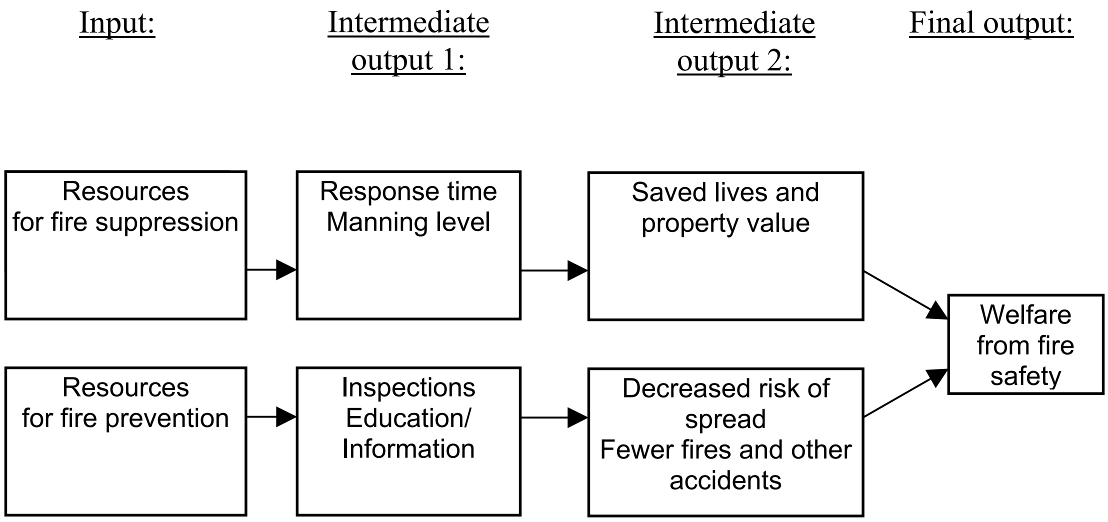
[5] Hjalmarsson, Kumbhakar and Heshmati (1996) have compared deterministic and stochastic frontiers.

The stochastic frontier can be used with panel data. Panel data means data over both individuals, i, and time, t. Not only an inefficiency term, u, but also a firm-specific term, λ, and a time-specific term, μ, can be estimated.[6] The model can then be written as

$$y_{it}=f(x_{it},\beta)+v_{it}-u_{it}+\lambda_i+\mu_t \qquad (2)$$

where the inefficiency term is time varying. The firm-specific and time-specific terms could be estimated as fixed effects (constant), or as random effects (varying). Panel data thus makes it possible to get more information, but the disadvantage is that (for the random effect model) more assumptions about distributions have to be made.

Figure 2. Input and output for the different fire service levels.



3 The papers

The fire services have two important tasks. First to prevent fires from happening, and second, if they happen, to suppress them as quickly as possible. The production process can be divided into three different output levels as illustrated in Figure 2. The resources

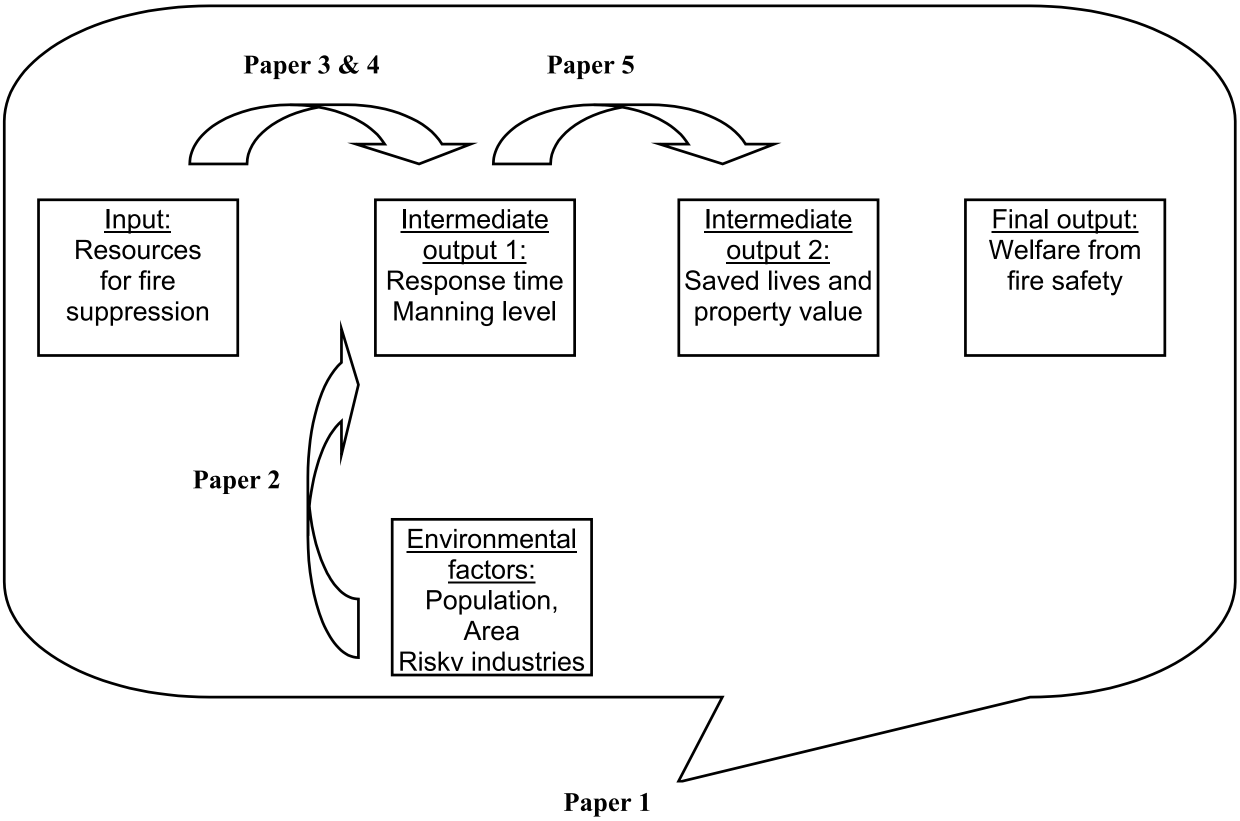
[6] Even more information can be received as shown by Heshmati (1994).

for suppression are used to produce the first intermediate outputs, the stand-by level, which can be measured by the response time (the faster the better), and the number of available firemen (the more the better). At the scene of the fire saved lives and property value are the second intermediate outputs. The input here is the first intermediate output (more firemen make it easier to suppress the fire). The final output is the increased total utility, i.e. the welfare change, due to all fire service activities.

In this thesis only the suppression aspect is studied. The reason is that prevention is much harder to measure. Considering the input, there is no proper definition of what is meant by fire prevention and fire managers have difficulties in describing the nature of resources that are used for fire prevention. With regard to output many other factors beside what the fire service does affect the number and spread of fires. The find statistically significant fire prevention activities data over a very long period is therefore probably needed.

Some of the different connections in the above production process are studied in the papers in this thesis. Figure 3 shows which paper that deals with what problem.
It is the supply of available statistics that has led to the choice of these analyses. For example, statistics about the change of the sizes of the manning level during the 90's, enabled a comparison of these with the change of the environmental factors in Paper 2. In Papers 3 and 4, available information about response times to the population was used as one output measure for the stand-by level. From 1996 centrally collected reports with statistics from all turn-outs in Sweden exist. Conclusions about the performance of the fire services at the scene of the fire were made from these reports. These conclusions are discussed in Paper 5.

Figure 3. Description of the papers.



3.1 Paper 1 – The problems of defining outputs in the public sector's service production – a discussion with an application to the fire service

The purpose of this paper is to identify the problems of defining inputs and outputs when performing efficiency and productivity studies for the public sector. The problems are discussed with an emphasis on fire services. In the paper, the horizontal structure of the problem is discussed, i.e. whether one of the intermediate outputs or the final output should be used in the productivity analysis. Bradford, Malt and Oates (1969) originally discussed this. The paper ends with a framework for how to proceed with an efficiency and productivity analysis. (The same as the one presented in Figure 2 above.) In an appendix to the paper there is also a literature survey of the few earlier economic studies of fire services.

3.2 Paper 2 – Swedish fire and rescue services' manning levels – a stochastic frontier analysis using panel data

The purpose of this paper is to find out what the size of the manning level depends on. Does it depend on environmental factors such as the size of population and the size of the area, risk factors such as the number of fires, other accidents, and risky industries, or is it independent of these factors meaning that size only depends on tradition? An input requirement frontier function is estimated using econometric methods with the Battese and Coelli (1995) frontier specification. This specification makes it possible to separate environmental factors affecting the production technology and those affecting the efficiency term. The main conclusion from the study is that the size of the manning level mainly depends on the size of the population, and that mean efficiency is low, about 0.7. Thus, it seems that tradition is also a relevant factor.

3.3 Paper 3 – Measuring the efficiency of the Swedish fire services' stand-by level

In this paper the efficiency in the production of the first intermediate output, the stand-by level, is investigated. The stand-by output has two dimensions 1) the response time of the fire service and 2) the suppressing power of the fire crew. The first dimension is measured according to the number of people that are reached within five and ten minutes, respectively, and the second dimension is measured by the total manning level of the fire service. Only the operative aspect has been studied, since no statistics about the resources on the administrative side existed. A data envelopment analysis (DEA), assuming a variable returns to scale technology and including environmental factors such as population and area as fixed inputs and outputs, was performed. The reason for choosing DEA was that prices on inputs for each fire service were not available. Mean efficiency is calculated to about 0.7. Another result is that most fire services operate under constant and decreasing returns to scale.

3.4 Paper 4 – Productivity change of Swedish fire services between 1992 and 1998

This study looks at the productivity change of the stand-by level between 1992 and 1998. The same variables as in Paper 3 are used. A Malmquist productivity index is calculated, and the results show that productivity has decreased following the budget cuts that many municipalities have had. In other words: The cuts in input have led to even less output. The productivity indexes are also compared to some public choice variables. The total cost of the municipality, the fire service's share of total cost and the fire service's external income are factors that statistically significantly affect total productivity.

3.5 Paper 5 – Measuring performance differences using an ordinal output variable: The case of Swedish fire services

In this paper, the second intermediate output, i.e. what happens at the fire scene, is studied. The output variable is constructed using statistics about the spread of fires in private houses after the arrival of the fire crew. The statistics about the spread of the fire is not defined as a continuous variable, but instead the output measure is ordinal. This measure has then been compared to the inputs response time and size of the fire crew (the same as the first intermediate outputs). The comparison has been done using an ordinal probit model. The results show that using more firemen has a positive effect on fires in private homes, and that performance is not affected by the use of a full-time or part-time fire crew. Full-time firemen are better trained than part-time firemen, and therefore a positive effect on the performance could be hypothesised.

4 Conclusions

The results from the above analyses show low efficiency among fire services, and they also indicate that performance comparisons are rare in this field. If performed the output aspect has been mostly analysed on its own and the input aspect on its own. Their relationship, i.e. productivity, has seldom been discussed. However, the lack of useful statistics for comparisons between output measures and input measures of fire services has also contributed. Hopefully the ideas put forward here may enable fire managers to ask themselves whether their fire service are more productive than others, and whether there are any improvements that can be done. Let us hope that there will be an interest to develop better and comparable output and input measures for the fire service sector in the future.

References

Aigner, D., C. A. K. Lovell and P. Schmidt, 1977, *Formulation and estimation of stochastic frontier production function models*, Journal of Econometrics, 6, 21-37.

Battese, G. E. and T. J. Coelli, 1995, *A model for technical inefficiency effects in a stochastic frontier production function for panel data*, Empirical Economics, 20, 325-332.

Bouckaert, G., 1992, *Productivity analysis in the public sector: the case of the fire service*, International Review of Administrative Sciences, 58, 175-200.

Bradford, D., R. Malt and W. Oates, 1969, *The rising cost of local public services*, National Tax Journal, 22, 185-202.

Charnes, A., W. W. Cooper and E. Rhodes, 1978, *Measuring the efficiency of decision making units*, European Journal of Operational Research, 2, 429-444.

Cherchye, L., T. Kuosmanen and T. Post, 2000, *What is the economic meaning of FDH? A reply to Thrall*, Journal of Productivity Analysis, 13, 263-267.

Coelli, T., D.S.P. Rao, and G.E. Battese, 1998, *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers.

Cooper W., L.M. Seiford, and K. Tone, 1999*, Data Envelopment Analysis - A Comprehensive Text with Models, Applications, References*, Kluwer Academic Publishers.

Deprins D., L. Simar and H. Tulkens, 1984, *Measuring labor-efficiency in post offices*, in M Marchand, P. Pestieau, and H. Tulkens, eds., The performance of public enterprises: concepts and measurement, North-Holland.

Färe, R., S. Grosskopf, and C.A.K. Lovell, 1994, *Production Frontiers*, Cambridge University Press.

Färe, R., and S. K. Li, 1998, *Inner and outer approximations of technology: a data envelopment analysis*, European Journal of Operational Research, 105, 622-625

Fried, H. O., C.A.K. Lovell, and S. S. Schmidt (eds.), 1993, *The Measurement of productive efficiency*, Oxford University Press.

Grosskopf, S., 1986, *The role of the reference technology in measuring productive efficiency*, Economic Journal, 96, 499-513.

Heshmati, A., 1994, *Estimating Technical Efficiency, Productivity Growth and Selectivity Bias Using Rotating Panel Data: An Application to Swedish Agriculture*, Ekonomiska studier no. 50 (PhD dissertation), Göteborg University, Sweden.

Hjalmarsson L., S. Kumbhakar and A. Heshmati, 1996, *DEA, DFA and SFA: A comparison*, Journal of Productivity Analysis, 7(2/3), 303-328.

Juås, B., 1994, *Räddningstjänst vid byggnadsbränder*, Högskolan i Karlstad, Forskningsrapport 94:7. (Summarised in Optimal fire safety, Räddningsverket, P21-098/95,1995.)

Juås, B., 1995, *Tidsfaktorns betydelse vid räddningstjänstens insatser*, Högskolan i Karlstad, Forskningsrapport 95:15. (Summarised in Optimal fire safety 2, Räddningsverket, P21-157/96,1996, and in Mattsson & Juås, 1997, The importance of the time factor in fire and rescue services operations in Sweden, Accident Analysis and Prevention, 29(6), 849-857.)

Kumbhakar, S., and C.A.K. Lovell, 2000, *Stochastic frontier analysis*, Cambridge university press.

Meussen, W., and J. van den Broeck, 1977, *Efficiency estimation from Cobb-Douglas production functions with composed error*, International Economic Review, 18(2), 435-444.

Thrall, R.M., 1999, *What is the economic meaning of FDH?*, Journal of Productivity Analysis, 3, 45-65.

Varian, Hal, 1984, *The Nonparametric Approach to Production Analysis*, Econometrica, 52(3), 579-597.

1

Paper 1

The problems of defining outputs in the public sector's service production - a discussion with an application to the fire service

by
Henrik Jaldell

ABSTRACT

This paper discusses how outputs (and inputs) should be measured in the public sector, and how they could be used in productivity and efficiency studies. There are two different levels where the studies could be performed: the vertical and the horizontal levels. The vertical level is the distinction between performing studies on the macro, the national, and the within-unit level. The horizontal level is the distinction between different outputs, whether determinant variables, direct outputs or consequences are used. The paper also includes an application of these ideas to the fire and rescue service.

1 Introduction

This paper will investigate the problem of measuring outputs (and inputs) in the public sector with an emphasis on fire and rescue services (hereafter referred to "fire services").[1] Why write papers on how to define outputs and inputs? Well, perhaps because real life is not as simple as models assume that it is. If we are to compare productivity and efficiency, which involves comparing outputs to inputs across various fire service units, we must know the answers to questions such as: What is the fire services really doing? What are their objectives? How can we measure them? [2]

Section 2 in this paper introduces the theory of productivity and efficiency measurement using frontier techniques. Various definitions of efficiency for a given production frontier are presented. The problem of transferring efficiency measures from the private to the public sector is then discussed, in two dimensions. First is the appropriate "vertical" level of measurement; macro, national or within-unit level. These all correspond to similar levels for the private sector (i.e. macro, industrial and micro), but for the public sector, there is also a question of political allocation. The second dimension is the appropriate "horizontal" level. Should the direct outputs or the final consequences (effects) be used? Using final consequences it is important to handle equity concerns properly, while using direct outputs it is important to handle quality differences properly. It will be argued that direct outputs correspond better to private goods, and are therefore best suited in a production theory framework.

Section 3 summarises how inputs and, especially, outputs have been defined and measured in earlier studies on fire services. A complete description is given in the

---

[1] The official name in Sweden is "fire and rescue services". However, most of the discussion here is about fires and therefore only the word fire is used.

[2] The question of how to implement productivity and efficiency improvements is not the subject of this paper, though it is very important. Two recent discussions of the problems involved are Bouckaert (1993) for Belgian civic registry offices, and Nolander and Roos (1998) for Swedish pharmacies.

Appendix. Finally, in section 4 a synthesis is presented as a framework for future research on the economics of fire services.

2 Productivity and efficiency measurement using frontier techniques

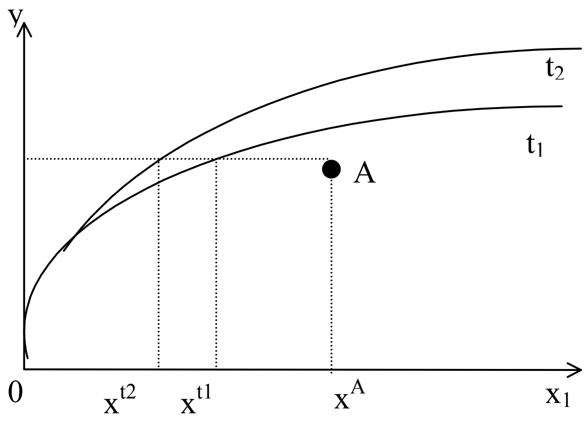
The theory of productivity and efficiency measurement using frontier techniques has mainly been developed for the private sector. In this section, the problem of transferring it to the public sector will be discussed.[3]

2.1 Measuring efficiency and productivity

The production units using the "best" technique define the production frontier. Technical change is defined as a shift of the frontier, shown in Figure 1 where there has been a positive shift from t_1 to t_2. An increase in average productivity for the whole sector is not necessarily the same as a shift of the frontier, however, because an increase in average productivity can be decomposed into three parts (Lovell, 1993): (i) the frontier has shifted due to better best-practise technology, (ii) the ”non-frontier" units have moved closer to the frontier, becoming more efficient by catching up with the best technology, and (iii) the environment has changed.

[3] Unfortunately, there is no general agreement on definitions of the terms productivity and efficiency (nor for the occasionally used effectiveness), and the distinction between them is not always clear. Sjöblom (1990) used the word "idea-soup" to describe the situation. In Swedish, there is even more confusion, since both efficiency and effectiveness are translated as "effektivitet".

Figure 1. Productivity change with one input and one output.



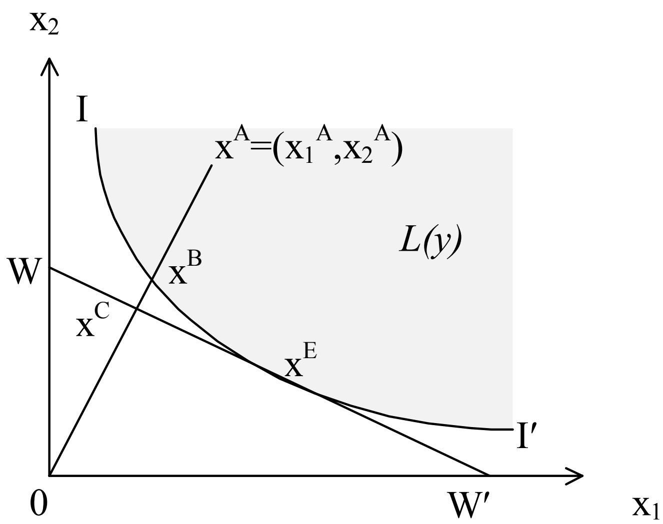
Firms not located on the frontier, and thus less productive than those on the frontier (e.g. unit *A* in figure 1) are called inefficient. Farrell (1957) defined three efficiency measures assuming constant returns to scale relating to technical efficiency, price efficiency and overall efficiency. Technical efficiency is measured as the relative reduction in input possible while still producing the same amount of output, or as the relative increase in output possible using the same amount of input. These are called input-saving and output-increasing efficiency respectively. The measure of technical efficiency belongs to the range (0,1], and is 1 if the observation is fully efficient. In Figure 1, $0x^{t1}/0x^{A}$ and $0x^{t2}/0x^{A}$ measure input-saving technical efficiency. To measure price efficiency we compare observed average costs to those of the least-cost producers. The measure of price efficiency also belongs to (0,1], and is 1 if the observed unit uses the cost-minimising input-mix. Overall efficiency is then measured as the product of technical and price efficiency.[4]

All these measures can be seen in Figure 2 where the efficient isoquant, labelled *II′*, bounds the input requirement set, *L(y),* while the actual input choice is labelled x^A. The

[4] Often price efficiency is called allocative efficiency, and overall efficiency is called cost efficiency or revenue efficiency. Allocative efficiency here is, however, equals not allocative efficiency on a macro level, se below.

technically efficient input vector of x^A is labelled x^B; i.e. it is possible to scale down both inputs and still be in *L(y)*. Then the Farrell measure of the technical efficiency of x^A is $0x^B/0x^A$. The line *WW'* shows the observed input price-ratio. A unit using inputs x^E is both technically efficient and price efficient. The cost of producing at point x^C is the same as at x^E, so the measure of price efficiency is $0x^C/0x^B$ for unit *A*.[5] Overall efficiency for unit *A* can then be calculated as $0x^B/0x^A * 0x^C/0x^B = 0x^C/0x^A$. How to calculate these measures with real data is discussed in section 2.2.2.5 below.

Figure 2. Efficiency measures with one output and two inputs.



A structural efficiency measure for the whole industry can also be calculated using the average of efficiencies of the individual firms, or using an average firm. With non-homogenous production functions, it is also possible to calculate measures of scale efficiency, showing how close a firm is to optimal scale. Førsund and Hjalmarsson (1987) has a thorough discussion of all these efficiency measures, within a neo-classical production theory framework, while Färe, Lovell, and Grosskopf (1994) presents the efficiency measures within an axiomatic framework.

[5] Cost-minimising proportions are only independent of the scale of production, as drawn here, in the case of a homothetic technology.

Førsund and Hjalmarsson (1987) defined three different levels of the economy where it is interesting to study efficiency: the macro level, the industry level, and the within-firm level. At the macro level, allocative efficiency measures are used, for example to measure the efficiency loss due to monopoly or monopsony powers. Both the supply and demand sides of the economy are taken into account.

At the industry level, demand is given and only supply is of interest. The object is to compare units within the industry to best-practise technology, which makes frontier production theory a natural framework in which to work.

Within-firm efficiency, concerned with how a single firm uses its resources, is a natural interest for managerial and engineering sciences. For a multi-plant firm, however, frontier production theory can be used, since it is interesting to compare the different plants to each other.

The shift of the production frontier, i.e. productivity change, is most often interesting to study over time, while how efficiency differs between firms is studied both over time and at a certain point in time.

2.2 The public sector

In producing the service to the public, the public services will probably not be fully efficient in a competitive market sense. There are three main reasons for this: the public sector has other objectives, public choice and property rights reasons, and monopolisation.

Public firms may have other objectives than profit maximisation assumed for a private firm. Examples of other objectives are that the distribution may be important to the public firm (i.e. they do not just care about efficiency but also about equity), that the public firm have a longer planning cycle than a private firm, or that it may be run by ideological reasons. The stated objectives may also be poorly defined and complex, which tend to increase cost levels.[6]

The public choice approach, following Borcherding, Pommerehne and Schneider (1982), explains inefficiency with failures in monitoring and controlling the public firms. These problems are called principal-agent problems. For example, a municipality council have difficulties in specifying incentives to public firms to behave like the council wants. The managers (the agent) may have their own goal they want to achieve (power, luxury offices etc).

The property rights approach concentrates on the impossibility of transferring ownership rights among individuals in the public sector, as compared to the relative ease within the private sector. Therefore there is less pressure on the public firm, than on a private firm, that has a competitive market environment that pressures it to increase efficiency.

Since the fire service has a monopoly in each municipality, there is no alternative company to call, and thus there may exist inefficiency because of no competition. In the private sector, a firm supplying poor quality or service automatically loses profits and will eventually exit the market, but in the public sector, this process does not work.

When measuring efficiency in the public sector it is important not only to know both what vertical level one is interested in (corresponding to the private sector's macro, industrial and within-firm levels), but also what horizontal level (directly produced output or final output) one is interested in.

---

[6] Models with other objectives than efficiency are discussed, for example, by Steinberg (1986) and Carroll (1993).

2.2.1 The vertical levels

How to measure efficiency at the public sector's various levels has mostly been discussed by political and managerial scientists, not by economists. The political scientists Dalton and Dalton (1988) proposed six measures of public sector efficiency (which they called productivity). [7]

At the macro level, Dalton and Dalton (1988) defined social effectiveness, i.e. how capable the public sector is in approximating the private market's supply. This is the topic in the public economics literature, where the question is allocative efficiency. If one dials the alarm number in case of a fire, for example, can one be sure that the fire service will turn-out with sufficiently well-educated and motivated firemen? Is supply equal to demanded quantity and quality?

At the national (industry) level two of Dalton and Dalton's (1988) proposed measures are the same as the two technical efficiency measures defined above, output increasing and input saving efficiency, and they also include price and overall efficiency. In most public sector studies, the input saving measure is most appropriate since most often inputs can be varied but output is given. For example, given the risk of the environment, an input saving efficiency measure would compare actual resources used to the least possible resources. Output increasing measures would compare value and lives saved to the most possible with the same crews, or would compare the number of fires to the least number achieved with the same resources on fire prevention. For fire services in Sweden, this is a natural level to study since they are independent units.

For the within unit level Dalton and Dalton (1988) called efficiency "organisational effectiveness". Examples of high organisational effectiveness could be that the goal of 90 seconds to a turn-out is always met, and that a sufficient number of healthy firemen

[7] As mentioned earlier there is no agreement on the definitions of these measures. Gary, Flynn, Jenkins, and Rutherford (1988, according to Sjöblom, 1990), for example, define six measures: economy (cost in

are always available. Studying services with many branches such as police, social insurance or pharmacies, Farrell efficiency measures could also be used.

An additional level for the public sector is what Dalton and Dalton (1988) call political allocation, which depends upon the perspective, circumstances, goals and interests of the political parties. For the public sector in a democratic state, the principles of liberty, merit, equality, and human rights should also be taken into account. An example of the political allocation would be if there were differences between the municipalities' spending on fire service, depending on which party had the political majority. In economics, these questions are analysed within the public choice literature.[8] Incorporating equity into the analysis is important, since one of the main reasons for having the good publicly supplied in the first place is that free market distribution gives an "unfair" allocation.

2.2.2 Relating outputs to inputs: The horizontal levels

The main difficulty with some public sector activities is how to define the output variable, since there is no market where quantities of outputs are sold. Following Ross and Burkhead (1974), there are for the public sector five different methods of relating outputs to inputs, i.e. of doing efficiency and productivity studies:[9] (1) using work measures; (2) measuring outputs by inputs; (3) the determinants approach; (4) using changes in consequences or effects related to inputs; and (5) using changes in the quantity of direct outputs related to inputs (the production function approach)

relation to output), efficiency, effectiveness (fulfilment of goal), efficacy (fulfilment of goal), equity and electability.

[8] Two examples of combining technical efficiency measurement and public choice variables for explaining inefficiency are Duncombe, Miner, and Ruggiero (1997) for US schools, and Grossman, Mavros, and Wassmer (1996), analysing the spending of large US cities.

[9] Mellander (1993) suggested a sixth way; to measure productivity and efficiency when data on physical output cannot be found. However, the method relies on restrictive assumptions of homotheticity and constant returns to scale.

2.2.2.1 Work measures

The distinction between work measures and productivity measures is important. For example, consider an automobile factory. A work measure is how many hubcaps are installed per hour. If a worker runs instead of walks around a car, the work measure will increase, but this will only result in increased productivity if the number of cars produced is increased. Work measures are measures of intermediate activities. Productivity and efficiency measures are concerned with the linkage between inputs and final products. The problem with work measures is that they are not based upon a general theory of production.

2.2.2.2 Measuring outputs by measuring inputs

The most popular way to measure output in the public sector has been to use the value of the inputs that goes into the service. This is the method used, for example in the GNP accountings, and is the same as just using total costs in the private sector without regard to how much is actually produced. With this method, productivity (by definition) cannot change, which has lead to a decrease in the popularity of this method.[10]

2.2.2.3 The determinants approach

The determinants approach uses expenditures as the dependent variable in multiple regression, where the independent variables are all factors which may influence the level of expenditure, including proxies for quality changes. The purpose is to find the factors that influence expenditure for a certain service, and thus to explain differences in the expenditure levels of different services. Ross and Burkhead (1974) find two problems with this method. First it does not separate demand and supply factors, whereas productivity analysis is only about the production of goods and services, i.e. the supply side of economics. The second problem is that there is no behavioural theory underlying the method, so the chosen factors have no theoretical justification. However,

[10] This standpoint is expressed for example in a Swedish government report (Ds 1994: 24) on productivity in the public sector.

determinants studies are useful for explaining differences in levels of expenditure over time and among units.

2.2.2.4 Using consequences as outputs

Since there are ambiguities about what is meant by public output Bradford, Malt, and Oates (1969) distinguished between the services directly produced, called direct outputs, and the things of interest to the citizen-consumer, called consequences. For example, the direct outputs resulting from police inputs used (police officers, cars, communications equipment) might include the number of blocks provided with a specified degree of surveillance, the number of blocks provided with readily available police-officer reserves, the number of intersections provided with traffic control, and so on. A citizen however, is primarily interested in the actual consequences, effects, or outcomes such as the degree of safety from criminal activity, and the smoothness and rapidity of the flow of traffic. Consequences depends both on the direct outputs, and on environmental variables. The distinction between the outputs is important because the trend of the cost of providing them may be quite different. Studies comparing consequences to inputs are often called effectiveness studies, to distinguish them from efficiency studies (comparing direct outputs to inputs).

The consequences are not just supplied by the public sector, but are also often public goods in an economic sense. Pure public goods have two properties: non-excludability and non-rivalry. Non-excludability means that no household can be excluded from consuming the good once provided, while non-rivalry means that consumption of the public good by one household does not reduce the quantity available for consumption by any other.[11] Problems with non-excludability and non-rivalry are one reason why environmental variables are so important in studying efficiency in the public sector.

[11] Poole (1980) argued that fire protections has close private substitutes, such as home alarms and own fire extinguishers, and therefore it does not have to be provided by the public sector. A more competitive environment for fire and rescue brigades has also been discussed in Sweden (SOU 1994:67).

It is difficult for the fire service to exclude anyone from fire suppression activity: Once a fire has been suppressed in one house, all houses nearby (including possible free-riders, who have not paid for fire protection) have also benefited from the service. For fire prevention, it is (in theory) possible to just help and inform these individuals who have paid fees, but again all people nearby will benefit from the fire prevention activities, and thus free riding is possible. Fire prevention does seem non-rivalrous, but for fire suppression there is some rivalry: It may be difficult for the fire service to fight two, or three, fires at the same time.

Since one of the reasons for supplying the service publicly is equity, it is important to incorporate this into the measures of public sector productivity. Using consequences as outputs also equity concerns should be controlled. Brudney and Morgan (1988) proposed that equity could be incorporated into the productivity measures by different weighting schemes. For example: "Assume that library service to 'low-education' residents is three times as important to other". The problems with this approach are the specification of target groups and how to choose weights. Therefore, this approach seems arbitrary and subjective in the choice of weights. A better approach is to use equity as a restriction, not as an objective, and to present how the result affects different groups. [12]

In any case, according to Ross and Burkhead (1974), using consequences as outputs is useful for program evaluation and for public expenditure justification, but it does not provide a useful framework for measuring the quantity of public sector output in the terms necessary for productivity studies, because the products of the production process are confused with the consequences of the products. The public agencies' control of the final outputs is too small to be of any interest in productivity studies. Ross and Burkhead say (1974, p. 48): "--- production theory allows one to make hypotheses regarding the relationship between the quantities of labor and capital used in the

---

[12] Equity should only be *controlled* for in productivity and efficiency studies, and thus equity is a restriction, not an objective.

production process and the quantity of guns of a given quality which come out of that production process. The theory does not question how those guns are used.” [13]

In the private sector, the distinction between direct and final outputs is not a major problem. To maximise profits, the firm must produce products demanded by the consumers, thus fulfilling the effectiveness criterion, and in doing so, they must produce at low costs, thus fulfilling the efficiency criterion. Effectiveness measures are therefore almost never measured in the private sector.

2.2.2.5 Using direct outputs: the production function approach

The public sector output measure that comes closest to output in the private sector is the direct output, as defined by Bradford, Malt, and Oates, not the consequences. Since public sector direct outputs are produced in a similar fashion as in the private sector, it is natural to use production theory when measuring those outputs.[14]

A frequent objection to productivity and efficiency analyses of public services using intermediate outputs, in which some units are found to be less efficient than others, is that quality has not been considered. The quality dimension could not be included, since the goods and services have no market prices, and thus there is no feedback on quality. The main problem, according to Hjalmarsson (1991) is that quality is a demand problem, while productivity measurement is a supply problem. Hjalmarsson lists various cases of the quality problem: In the simplest case, homogeneity makes it easy to divide the products into different quality classes. At the other extreme, every product can be unique, and including quality is very difficult. Thus in some cases a common sense of what constitutes ”good” quality does not exist, while in other cases it is possible to measure quality but perhaps at a high cost. If quality is interesting the best

[13] A guide to the problems of measuring consequences or outcomes is found in Smith (1996).
[14] For an example of this reasoning, see Grosskopf, Hayes and Hirschberg (1995), studying police efficiency.

approach is to study quality ex post, i.e. after the efficiency scores have been obtained, and ex ante, i.e. trying to adjust the output variables.[15]

The method of using direct outputs has also been criticised because direct outputs are of little concern to consumers (e.g. Vedung 1995). Ross and Burkhead (1974) pointed out two basic difficulties with this argument: First, one really has the same problem in the private sector; why do we not measure consequences in the private sector? The second problem is that it confuses efficiency and effectiveness. Efficiency measures are concerned with the direct products, while effectiveness is a measure of the consequences. Again, consider the automobile industry, where efficiency measures how many cars of a certain quality are produced using labour and capital, measured either in costs or in physical units. An effectiveness measure would look at the consequences, i.e. how the car is used, measured in passenger-kilometres, for example.

Another objection to using efficiency measures instead of effectiveness measures is that the relation between direct and final outputs is not known. Bouckaert (1992) emphasised that *a priori* it is not always appropriate to assume a positive relation between efficiency and effectiveness. The relation may be negative: increased efficiency may lead to decreased effectiveness. This objection seems to be correlated with the quality dimension, especially in the service sector. Thus, producing direct outputs more efficiently may lead to a loss in quality of the final output.

Production theory has been widened by frontier techniques, making it possible to estimate efficiency. Production frontier literature can be divided depending on which estimation method is used: the econometric approach, or the programming approach. Both of these can be used with either a parametric or a non-parametric specification of the frontier, and with either a deterministic or a stochastic frontier. In a deterministic

[15] A practical example of how to incorporate quality into efficiency analysis is proposed by Bjurek, Gustafsson, and Kjulin (1992), who regressed technical efficiency scores obtained from a DEA study on child care against various quality variables.

specification, all deviation from the frontier is due to inefficiency, while in a stochastic specification a white-noise term is also added.[16] In the parametric case, a parametric functional form (e.g. Cobb-Douglas or translog) is assumed for the production frontier before estimation. It is now most popular to use the econometric approach for parametric stochastic frontiers, and the programming approach for non-parametric deterministic frontiers. The public service often produces several outputs using several inputs. Therefore, when estimating a parametric frontier, a cost frontier is most often used for the public sector, which makes it possible to estimate both technical and price efficiency.[17] However, some economists argue that, since public providers have other objectives and constraints, comparisons between public providers should only be made on the basis of their technical efficiency (Lovell, 1993).

In the public sector, the most popular technique for efficiency is the non-parametric deterministic programming approach called data envelopment analysis, DEA, in which one neither specifies a functional form nor assumes a specific behaviour. However, a basic assumption is that the technology is convex.[18] The main problem with DEA is that it is deterministic, and thus does not consider statistical noise. It has mainly been motivated in economics within an axiomatic framework (Färe, Grosskopf, and Lovell, 1994), but also within a neo-classical framework (e.g. Førsund, 1996). Two recent examples of specialised data envelopment models developed with the public sector in mind are the indirect approach by Färe, Grosskopf and Lovell (1988) which adjusts

---

[16] Examples of surveys for the parametric frontier include Schmidt (1985-86) and Førsund and Hjalmarsson (1987); for the non-parametric frontier, Ganley and Cubbins (1992), Färe, Grosskopf, Lovell (1994) and Charnes, Cooper, Lewin, and Seiford (1994), and for both; Fried, Lovell, and Schmidt (1993) and Coelli, Rao, and Battese (1998). A comparison of the results from these approaches is made in Hjalmarsson, Kumbhakar, and Heshmati (1996). Greene (1997) discusses deterministic versus stochastic frontiers using econometric techniques.

[17] This is, however, not that easy as demonstrated for example by Kumbhakar (1996).

[18] The ordinary DEA-model can be adjusted for convex input sets, with non-convex output sets, assuming a piecewise Cobb-Douglas technology, or both non-convex input and output sets, the FDH-model (free disposal hull) (Färe, Grosskopf, Lovell, 1994). The FDH-model has higher efficiency numbers compared to the ordinary DEA-model, since it bounds the data points more closely. (The frontier looks like a flight of stairs.) The reason for assuming non-convexity is mainly pedagogical; production units should not be compared to artificial units, but only to real units. The FDH-model has been criticised for assuming non-economical behaviour (Thrall, 1999).

efficiency for the problem of a fixed budget, and the approach by Ruggiero (1996) which adjusts for environmental variables. However, no suggestions of how to incorporate equity concerns have been found.

Empirical examples of public sector efficiency analyses using frontier techniques made in Sweden concern hospitals, theatres, courts, and district attorney offices (Ds 1994:24); social insurance offices (Hjalmarsson and Kumbhakar, 1991 Bjurek and Hjalmarsson, 1995); child care (Bjurek et al, 1992); hospitals (Färe et al, 1994); schools (Heshmati, 1997) and pharmacies (Althin, 1995). International studies include post offices, local governments, hospitals[19], schools, and highway maintenance patrols (see Lovell, 1993, for a list). No study has been found using frontier techniques fire service, other than Bouckaert (1992), which only uses graphs.

2.2.2.6 Conclusions

In reality, most often there is no real choice between the above methods: You have data on determinants, or consequences, or direct outputs, and you have to stick with what you have. If there is a choice, direct outputs should clearly be used when using production theory. However, the distinction between types of outputs is not always clear, and the major advantage of using consequences is that they are more acceptable to the decision-maker than are direct outputs. Therefore, a direct output, as close to a consequence as possible, is the best choice.

3 Summary of earlier studies of fire services

As discussed at some length above the main problem is output: What is the fire service really producing? With respect to what should different fire brigades be compared? To be able to compare them, clearly a standardised measure must be used.

We will now look at how the outputs (and inputs) of the fire service have been defined in earlier studies. The papers are presented in the Appendix. Sections A1–A10 lists papers and articles with empirical work, while sections A11–A13 discuss some earlier theoretical work.

Table 1 summarises the various output variables from the literature survey presented in the Appendix, including a classification of whether the determinant approach (DA), the production function approach (PF), or some other method (O) is used. The outputs are for both fire prevention and fire suppression activity. In the last column, there is also a description of how the various researchers classified their output variables as direct output (D), final output (C), output as an environmental (E) or output as a quality (Q) variable.

As indicated in the table, there are differences of opinion on how to classify some of the same variables. Population, number of fires, size of area and casualties are sometimes direct outputs, other times final outputs. However, the researchers have agreed on property value to be a final output.

[19] A special issue of Journal of Health Economics was dedicated to efficiency analyses (1994, vol. 13).

Table 1. Examples of proxies for output measures.

Section in app.	Work	Approach#	Variables	Output class¤
A1	Hirsch (1959,1973)	DA	night-time population	D
			area	D
			density of dwelling units	D
			index of scope and quality of fire protection	C
			value of real property	D
A2	Ahlbrandt (1973)	DA	population	C
			size of area	C
			assessed value	C
			number of stations	Q
			number of personnel	Q
			fire insurance rating	Q
			% of houses lacking plumbing facilities	E
A3	Wallace (1977)	O	quality index incl. average time and number of firemen	Q
A4	Coulter (1979)	DA	value of property losses	C
			number of deaths	C
			number of injuries	C
			number of fires	C
A5	Kristensen (1983)	O	costs	
A6	Southwick and Butler (1985)	DA	number of civilian deaths	D
			value of property losses	C
			number of alarms	C
			number of building fires	C
			number of total fires	C
A7	Bouckaert (1992)	PF	number of fires	D
			number of fire prevention activities	D
			population	C
			size of area	C
			value of property value	C
			number of casualties	D,C
A8	Johansson (1992)	O	population	D
			number of fires	D
A9	Duncombe (1991, 1992), Duncombe and Yinger (1993), Duncombe and Brudney (1995)	PF	ration of property value to property value lost	C
			service for emergence	Q
			% of fire fighters paid	Q
			% of houses built before 1940	E
			% of people below poverty line	E
			% of property value in industrial property	E
			% of houses higher than 2 floors	E
A10	Juås (1994, 1995)	O	saved property value	D
A11	Schaenman and Swartz (1974)	-	see Table A1 in Appendix	

Section	Work	Approach[#]	Variables	Output class[¤]
A12	Morley (1986)	-	number of fires and other emergencies	D
			fire inspections and investigations	D
			property loss	C
			response time	C
			casualties	C
			number of people saved	C

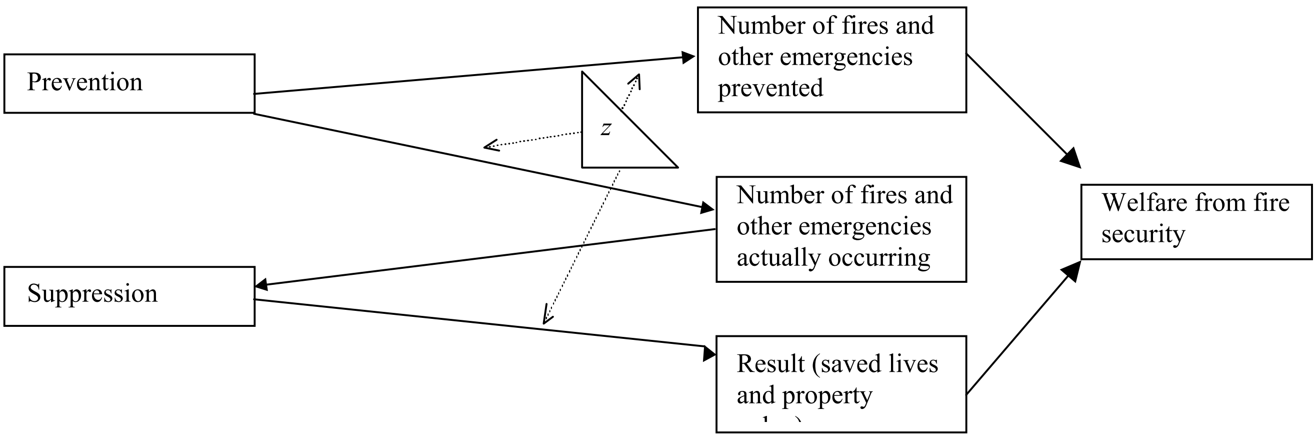
[#] DA = determinant approach
PF = production function approach
O = other approach
\- = no empirical analysis

[¤] D=direct output
C=consequence (final output)
E=environmental variable
Q=quality variable

4 A framework for empirical studies

The role of the fire service is to supply people with the good they demand, the feeling of security. It does this by two main activities, fire prevention and fire suppression, as outlined in figure 3.

Figure 3. The fire service system.



z= environmental factors (e.g. weather, risky industries, risk of spread etc.)

The most important activity is to prevent fires and other emergencies, such as traffic accidents, flooding etc., from happening. Such fires and other emergencies may then be due to failure of the fire service, but they may also depend on things not in their control. No matter how much resource the fire service put on prevention activities, such as inspections and information, fires and other accidents will happen.

Uncontrollable factors also influence fire suppression. If there is a fire the fire service turns out, and the result of the turn-out depends not only on their own activity (the turn-out time, the number and skill of the firemen), but also on environmental factors (the weather, building conditions, etc.). The question now is how we can measure the outputs of the fire service.

4.1 Intermediate outputs

4.1.1 Fire prevention

Fire prevention is executed through building codes, education, information and inspections. The direct outputs are then the number of inspections done, and the number of people educated and informed. (Fire services are normally not responsible for building codes.) There seems to be one intermediate step missing here, describing how many fires are prevented.

The outputs of fire prevention include not only the occurrence of fewer fires, but also, if a fire occurs, they include reduced loss of life and property value. One ideal measure would then be the number of fires actually prevented from happening, but in reality, the best one can do is to use the number of fires (or the inverse thereof), while keeping influence of the surroundings constant. Another ideal output of fire prevention would of course be the number of lives and the property value actually saved due solely to those activities, but in practice this is indistinguishable from those saved due to fire suppression activities.

4.1.2 Fire suppression

The fire service must be prepared to turn out at any time, without knowing whether a fire will actually occur or not. This intermediate output can be measured both by response time, the faster the better, and by how many firemen that will turn out, the more the better. Therefore, the manager, using resources received from the municipality, must decide: how many firemen to have, if they shall be full-time or part-time, how many stations to have, and how to divide the firemen between the stations.

Once a fire occurs and the fire crew turns out, another output is produced: how well the fire brigade succeeds in suppressing the fire, which can be measured by the lives and property value saved. The number of lives and the property value saved are not measurable, but one can instead use the inverse of the property value lost and the number of dead and injured.

The property value lost and the number of dead and injured depend first on the condition of the fire when the fire brigade arrives, then on the number and quality of the firemen available, but also on the environmental variables discussed earlier. In addition, from the suppression activity there is also a feeling of security, which reflects the final output.

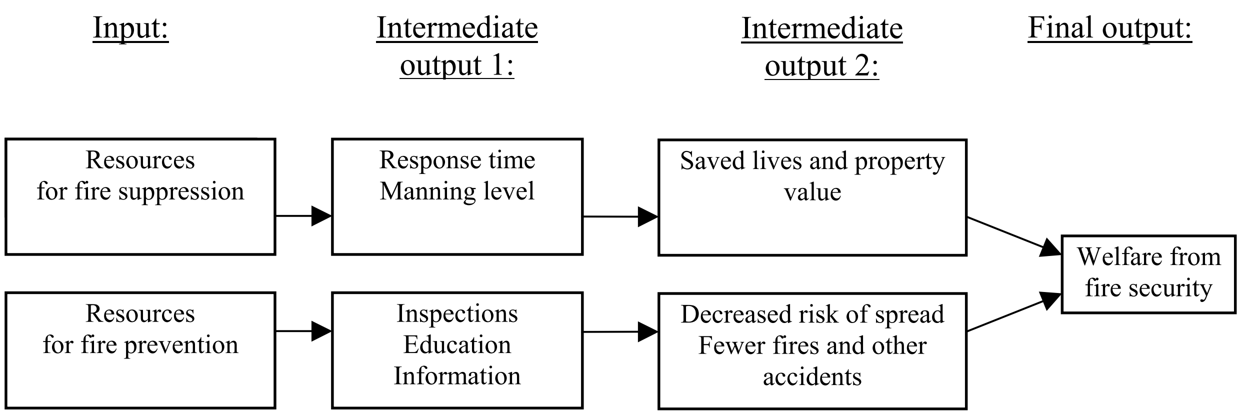
4.2 Final output

The feeling of security, or in other words welfare, that people get from fire prevention and fire suppression activities is the final output or consequence by consumers. It is the final output or consequence. The final output is very hard to measure, because it is difficult for people to distinguish between the security they feel from prevention as opposed to suppression activities if asked for example in a contingent valuation study. As with all demand functions it is a function of the price (or cost) of the good provided, the price of substitutes and complements, income, and preferences.

5 Conclusions

Figure 4 extends the two-step model for public sector output suggested by Bradford, Malt and Oates (1969) to incorporate the two levels of intermediate outputs of both fire prevention and fire suppression.

Figure 4. Output from fire service in more detail.



The structure of fire prevention and suppression activities with tree different levels of outputs may be an explanation for the diversity in output measures as described in Table 1.[20] This structure will be used in empirical studies of the Swedish fire and rescue services' productivity and efficiency.

[20] Another explanation is that data does not exist for all the interesting measures, and therefore proxy variables have been used instead.

References

Ahlbrandt, R., 1973, Efficiency in the provision of fire and rescue brigades, *Public Choice*, 16, 1-15.

Althin, R., 1995, *Essays on the measurement of producer performance*, Lund Economic Studies 60, Lund University.

Bjurek, H., 1994, Essays on efficiency and productivity change with applications to public service production, Ekonomiska studier 52, Göteborg University.

Bjurek, H., B. Gustafsson, and U. Kjulin, 1992, *Efficiency, productivity and determinants of inefficiency at public day care centres in Sweden,* Memorandum 160, Department of Economics, Göteborg University.

Bjurek, H. and L. Hjalmarsson, 1995, Productivity in multiple output public service: a quadratic frontier function and Malmquist index approach, *Journal of Public Economics*, 56, 447-460.

Boercherding, Pommerehne, and Schneider, 1982, Comparing efficiency of private and public production: the evidence from five countries, *Zeitschrift für Nationalökonomie*, suppl.2, 127-156.

Bouckaert, G., 1992, Productivity analysis in the public sector: The case of the fire and rescue service, *International Review of Administrative Sciences*, 58, 175-200.

Bouckaert, G., 1993, Efficiency measurement from a management perspective: A case of the civil registry office in Flanders, *International Review of Administrative Sciences*, 59, 11-27.

Bradford, D., R. Malt, and W. Oat, 1969, The rising cost of local public services: Some evidence and reflections, *National Tax Journal*, 22(2), 185-202.

Brudney, J. and D. Morgan, 1988, *Local government productivity: Efficiency and equity*, ch. 9, 161-176, in Kelly (1988).

Carroll, K., 1993, The effects of multiple objectives in the theory of public sector supply, *Public Choice*, 75, 1-20.

Charnes, A., W. Cooper, A. Lewin, and L. Seiford, 1994, *Data envelopment analysis: Theory, methodology and applications*, Kluwer Academic Publishers, Boston.

Coelli, T., D.S.P. Rao, and G. Battese, 1998, *An introduction to efficiency and productivity analysis*, Kluwer Academic Press, Netherlands.

Coulter, P., 1979, Organizational effectiveness in the public sector: the example of municipal fire protection, *Administrative Science Quarterly*, 24, 65-81.

Dalton, T., and L. Dalton, 1988, *The politics of measuring public sector performance: Productivity and the public organization*, in Kelly (1988), ch. 2, 19-68.

Ds 1994:24, *Den offentliga sektorns produktivitetsutveckling 1980-1992*, ESO, Fritze, Stockholm.

Duncombe, W., 1991, Demand for local public services revisited: The case of fire protection, *Public Finance Quarterly*, 19, 412-436.

Duncombe, W., 1992, Costs and factor substitution in the provision of local fire and rescue brigades, *Review of Economics and Statistics*, LXXIV, 180-185.

Duncombe, W., and J. Brudney, 1995, The optimal mix of volunteer and paid staff in local governments: an application to municipal fire departments, *Public Finance Quarterly*, 23(3), 356-384.

Duncombe, W., J. Miner, and J. Ruggiero, 1997, Empirical evaluation of bureaucratic models of inefficiency, *Public Choice*, 93, 1-18.

Duncombe, W., and J. Yinger, 1993, An analysis of returns to scale in public production, with an application to fire protection, *Journal of Public Economics*, 52, 49-72.

Farrell, M., 1957, The measurement of productive efficiency, *Journal of the Royal Statistical Society*, Series A, III, 253-280.

Fried, H. O., C.A.K. Lovell, and S. S. Schmidt (eds.), 1993, *The Measurement of productive efficiency*, Oxford University Press.

Färe, R., S. Grosskopf, B. Lindgren and P. Roos, 1994, *Productivity developments in Swedish Hospitals: A Malmqvist output index approach*, ch. 13 in Charnes A., W. Cooper, A. Lewin, and L. Seiford (1994).

Färe, R., S. Grosskopf, and C.A.K. Lovell, 1988, An indirect approach to the evaluation of producer performance, *Journal of Public Economics*, 37, 71-89.

Färe, R., S. Grosskopf, and C.A.K. Lovell, 1994, *Production Frontiers*, Cambridge University Press.

Førsund, F., and L. Hjalmarsson, 1987, *Analysis of industrial structure: A putty-clay approach*, Industriens Utredningsinstitut, Stockholm.

Førsund, F., 1996, On the calculation of scale elasticity in DEA models, *Journal of Productivity Analysis*, 7 (2/3), 283-302.

Ganley, J. A., and J. Cubbins, 1992*, Public sector efficiency measurement: applications of data envelopment analysis*, North-Holland, London.

Gary, A., A. Flynn, W. Jenkins, and B. Rutherford, 1988, *Productivity measurement in government: The experience of accountable management*, Central Paper for the EGPA-conference.

Greene, W.H., 1997, Frontier production functions, ch. 3 in M.H. Pesaran and P. Schmidt (eds.) *Handbook of Applied Econometrics, Vol. 2: Microeconomics*, Blackwell Publishers.

Grosskopf, S., K. Hayes, and J. Hirschberg, 1995, Fiscal stress and the public production of public safety: a distance function approach, *Journal of Public Economics*, 57, 277-296.

Grossman P., P. Mavros, and R. Wassmer (1996), *Public sector inefficiency in large U.S. cities*, Working Paper 1996:6*,* Department of Economics, University of Århus, Denmark.

Heshmati, A., 1997, *A comparison of parametric and non-parametric approaches to the adjustment of output for service quality differences: with applications to Swedish schools*, Memorandum 234, Department of Economics, Göteborg University.

Hirsch, W.Z., 1959, Expenditure implications of metropolitan growth and consolidation, *Review of Economics and Statistics*, 41, 232-241.

Hirsch, W.Z., 1973, *Urban Economic Analysis,* McGraw-Hill, New York.

Hjalmarsson, L., 1991, *Metoder för forskning om produktivitet och effektivitet med tillämpningar på offentlig sektor*, Ds 1991:20.

Hjalmarsson L., S. Kumbhakar and A. Heshmati, 1996, DEA, DFA and SFA: A comparison, *Journal of Productivity Analysis*, 7(2/3), 303-328.

Johansson, C., 1992, *Kostnadsvariationer inom det kommunala brand- och räddningsväsendet*, Meddelanden från ekonomisk-statsvetenskapliga fakulteten vid Åbo Akademi, Series A:366.

Juås, B., 1994, *Räddningstjänst vid byggnadsbränder*, Högskolan i Karlstad, Forskningsrapport 94:7.

Juås, B., 1995, *Tidsfaktorns betydelse vid räddningstjänstens insatser*, Högskolan i Karlstad, Forskningsrapport 95:15.

Kelly, R. M., 1988, *Promoting Productivity in the Public Sector*, Macmillan Press, Hampshire

Kristensen, O., 1983, Public versus private provision of governmental services: The case of Danish fire protection services, *Urban Studies*, 20, 1-9.

Kumbhakar, S., 1996, Efficiency measurement with multiple outputs and multiple inputs, Journal of Productivity Analysis, 7(2/3), 225-256.

Lovell, C. A. K., 1993, *Production frontiers and productive efficiency*, ch. 1 in Fried, Lovell and Schmidt (1993).

McMillan, M., 1987, On measuring congestion of local public goods, *Journal of Urban Economics*, 26, 131-137.

Mellander, E., 1993, *Measuring productivity and inefficiency without quantitative output data*, Economic Studies 14, Department of Economics, Uppsala University-

Morley, E., 1986, *A Practitioners Guide to Public Sector Productivity Improvement*, van Nostrand Reinhold Company, New York.

Norlander, N-O. and P. Roos, 1998, Implementing the Malmquist productivity index: the case of the national corporation of Swedish pharmacies, in R. Färe, S. Grosskopf, and R.R. Russell (eds.), *Index numbers: Essays in honour of Sten Malmquist*, Kluwer Academic Publishers, Boston.

Poole, R, 1988, Fire protection, ch. 15 in T. Cowen (ed.), *The theory of market failure: A critical examination,* George Mason University Press, Fairfax, Virginia.

Ross, J., and J. Burkhead, 1974, *Productivity in the local government sector*, Lexington Books, Toronto.

Ruggiero, J., 1996, On the measurement of technical efficiency in the public sector, *European Journal of Operational Research,* 90, 553-565.

Schaenman, P., and J. Swartz, 1974, *Measuring fire protection productivity in local government*, National Fire Protection Association, Boston.

Schmidt, P., 1985-86, Frontier production functions, *Econometric Reviews*, 4(2), 289-328

Sjöblom. S., 1990, *Produktivitet och effektivitet i offentlig sektor – perspektiv och problem*, Meddelanden från ekonomisk-statsvetenskapliga fakulteten vid Åbo Akademi, Series A: 308.

Smith, P. (ed.), 1996, *Measuring Outcome in the Public Sector*, Taylor and Francis, London.

SOU 1994:67, *Räddningstjänst i samverkan och på entreprenad*, Fritze, Stockholm.

Southwick, L., and Butler, R., 1985, Fire department demand and supply in large cities, *Applied Economics*, 17, 1043-1064.

Steinberg, R., 1986, The revealed objective functions for nonprofit firms, *Rand Journal of Economics*, 17(4), 508-526.

Thrall, R. M., 1999, What is the economic meaning of FDH? *Journal of Productivity Analysis*, 11, 243-250.

Vedung, E., 1995, *Produktivitet och effektivitet i kommunal tjänsteproduktion*, Meddelanden från Ekonomisk-statsvetenskapliga fakulteten vid Åbo Akademi, ser A: 426.

Wallace, R., 1977, Productivity measurement in the fire and rescue service*, Public Productivity Review*, 2(3), 12-36.

Appendix: Earlier studies on fire and rescue services

A1. Hirsch (1959) [21]

This is the oldest study of fire and rescue services from an economic perspective. Hirsch lists four activities for the fire department: fire prevention, fire detection, extinguishment and reduction of life hazards. The first two activities are carried out by fire engineers planning and checking buildings for risks and also involve others such as the municipality's building code administration, architects, building contractors, etc. In these activities the fire department act as an authority. The second class of activities, crisis response is the main function of the fire department: to turn-out and arrive at a fire and suppress it as quickly as possible, and afterwards to put the property in order, plus control that the fire does not start again.

According to Hirsch demand is represented by the incidences, a function of location, time, method of reporting, and weather conditions. On the supply-side, Hirsch estimated a cost function for fire and rescue brigades in the St. Louis city-county area in 1957. Total expenditure was estimated as a linear and quadratic function of quantity (night-time population), quality (index of scope and quality of fire protection used), and service conditions affecting input requirements (area, density of dwelling, increase in night-time population, receipts for wholesale, retail, and service establishments, and value of real property). All variables were statistically significant except density of dwelling.

A2. Ahlbrandt (1973)

Ahlbrandt estimated an average expenditure function for Seattle-King County, where expenditure per capita was estimated as a Cobb-Douglas function of the output level, factor prices, technology, quality, and the environmental factors that influence production decisions. Proxies for output were population, area, and assessed value of

---

[21] Reproduced in Hirsch (1973)

real estate. Capital was assumed to be invariant, so only labour expenditures were considered. The quality variables should have been response time, number of firemen available, level of training, and water pressure, but since response time and level of training were not available, the number of stations and the number of permanent personnel were used as proxies. A fire insurance rating was used to capture other dimensions. In addition, structural conditions within the community were incorporated by a variable showing the percentage of houses lacking some or all plumbing facilities.

The main results were that although fire protection is a public good, private supply of fire protection is cheaper than public supply. Sacks (1986) criticised Ahlbrandt and showed that the same model for Connecticut did not imply lower expenditure for private fire protection.

A3. Wallace (1977)

Wallace divided fire prevention and suppression into various subfunctions or stages. For prevention, they were inspections, code enforcement, and educational campaigns. For suppression, the stages were dispatching and firefighting. Dispatching had three steps: discovery of fire, notification of authorities, and dispatch of fire companies. Firefighting had four steps: travel to the fire, actual firefighting and extinguishment, overhaul (checking to assure the fire is out), and picking up equipment and return to service. In discussing output, Wallace pointed out that quantity measures such as the number of inspections and the number of fires do not give enough information, since the important thing is that the inspections minimise loss of life and/or property. Since there are no measures of the number of fires prevented through inspections or the dollar loss of property saved from fire, proxies must be developed.

The rest of the paper was about choosing quality adjustments (weights) for the output. For measuring the productivity of inspections, the output, number of inspections, was weighted with time spent fighting fires in inspected properties for different years. For dispatching, the output, number of dispatched units, was weighted with a complicated quotient of number of alarms with no turn-outs and number of alarms with no actual fires. For the firefighting stage, the quality index is the inverse ratio of the average time out of service in the study year to the average time out of service in the year before the study year. Average time is defined as $AT = \frac{\sum_{j}^{A}\sum_{i}^{N} M_{ij}T_{ij}}{A}$, where M is number of firemen from one company responding to an alarm, T is time the company was out, A is total alarms, and N is number of companies at an incident. Using this index, Wallace gave an empirical example from Syracuse, New York, where he found a remarkable increase in fire-fighting productivity of 12% and 35% between 1972 and 1973.

A4. Coulter (1979)

Coulter studied what he called the organisational effectiveness of public fire and rescue brigades. Prevention effectiveness referred to minimising the incidence of fires (e.g. the number of fires per capita), while suppression effectiveness referred to minimising loss per incident (e.g. dollar property loss per capita). Productivity was measured as the total cost of fire, i.e., expenditures plus property losses, per capita. Coulter hypothesised that variations in the effectiveness of municipally-provided fire and rescue services were a function of variations in their urban environment. Environmental characteristics largely determine the nature of fire hazards, but they are also influenced by the organisational and delivery characteristics of the fire and rescue service.

Using discriminant analysis fourteen variables were found significant for expenditures: unionisation, contract alarms, precipitation, constant manning, land area, very cold days, status of fire chief, density, social class, local alarms, institutionalised population, number of paid firemen, emergency fire/medical services and crowding.

Eleven variables were found significant for fire prevention: very cold climate, social class, building inspector's training, number of paid firemen, land area, minimum requirement for inspectors' education and training, inspection program comprehensiveness, fireman inspections, fire safety planning, manager-type government, and adjusted tax.

Ten variables were found significant for fire suppression: further education incentive, constant manning, land area, thunderstorms, maximum response time, fire and rescue service planning, population, number of full-time firemen, housing deterioration, and water supply deficiency.

Thirteen variables were found significant for fire-cost per capita: emergency response versatility, fire and rescue service planning, social class, maximum response time, status of fire chief, constant manning, inspection program comprehensiveness, land area, administrative staff size, institutionalised population, further education incentive for firemen, mutual aid responses, and emergency fire/medical and rescue brigades.

A5. Kristensen (1983)

Almost half of Denmark's municipalities have a contract with a private company, Falck, which is responsible for fire and rescue service. Leaving out those with voluntary or with a mixture of private and public fire service Kristensen compared 241 of Denmark's 275 municipalities with respect to costs. He found that private fire services had lower costs than public. Three reasons were given: economies of scale (Falck operates nation-wide); competitive pressure is harder on a private company, since a public service may believe that the city council will never change to a private supplier; and separation between demand and supply.

A6. Southwick and Butler (1985)

Southwick and Butler simultaneously estimated both demand and supply functions for the fire and rescue services for 65 major cities in the United States. Quantity demanded was hours of fire protection per capita, which was a function of fire losses, relative income, and population density. Quantity supplied was fire losses, which was a function of hours of protection per capita, population density, poverty, and population size. As proxies for fire losses five different measures were used (so five different demand and supply systems were estimated): number of deaths, total fires, building fires, alarms, and property losses.

Demand for firemen hours was found to be inelastic with respect to firemen wage, but to have a higher elasticity with respect to fire losses. There was no substantial difference of marginal elasticity with respect to population between large and small municipalities and thus no case for a public good.

A7. Bouckaert (1992)

Bouckaert first defined the differences between inputs (expenses), outputs (fire prevention, fire fighting, other emergencies, and ambulances), and consequences (positive: population, area, property value, and negative: damage, casualties). Since labour was a mix of volunteer and paid firemen the expenses of 13 fire departments in Belgium were used instead of labour as input. (The reason for using expenses instead of cost was that there was no cost accounting.) The number of fires and the number of fire prevention activities were chosen as outputs. Fires have different sizes, but when a similar group of cities is analysed, a similar distribution of fires can be assumed. For a measure of prevention the inverse of number of fires was used, and for a measure of suppression the inverse of number of casualties was used. For consequences, there was only reliable data for the population, size of area, property value, and number of casualties. The main purpose of the paper was to study efficiency, defined as differences in the relation between input and output, and effectiveness, defined as input vs. consequence. The efficiency analyses were made graphically and partially analysing only one input, and one output and consequence at a time. Professional services were found to be less efficient, but more effective, since they put more effort into fire prevention.

A8. Johansson (1992)

Johansson compared cost per capita and cost per turn-out for the various municipalities of two provinces of Finland (Åbo and Björneborg). She found no patterns whatsoever for productivity during the 1980's.

A9. Duncombe and Yinger (1993) and Duncombe and Brudney (1995)

Duncombe and Yinger estimated a translog cost function for fire protection in New York State with the division between direct and final (consequences) outputs in mind (see). In order to test the Bradford, Malt and Oates's (1969) hypothesis that environmental variables affect consequences, but not direct outputs the cost function was estimated with proxies for consequences. The single proxy for consequence used in Duncombe (1991, 1992), the inverse of property loss, *L,* relative to the property value, *V,* was extended to two: one measuring fire suppression, *L/F* (F=number of fires)*,* and one measuring fire prevention, *F/V*. The environmental variables were percent houses built before 1940, percent of population below poverty line (proxy for building condition), percent property value in industrial and utility use, percent houses higher than 2 floors (proxy for high-risk property), and population density (proxy for risk of spread). There were also two control variables: service for emergency and mutual aid, and percentage of paid firemen. Capital and labour costs were also included. Since certain fire prevention activities are not recorded in the fire departments budget, half of each city's spending on "other" public safety (outside police and fire) was added to total costs.

The main purpose of the paper was to suggest measures of marginal elasticities (or what Duncombe and Yinger called returns to scale) for the public sector. They especially pointed out that marginal elasticities with respect to output and population could be divided into three parts: the ordinary marginal elasticities as calculated in the industrial sector, a correction for service quality for the output, and a correction for congestion for the population measure.[22] The main findings were economies of scope between fire suppression and fire prevention; marginal elasticity with respect to population was unity; and returns to scale with respect to service were slightly increasing. Since environmental variables affected both direct and final outputs the hypothesis that environmental variables only affect the second stage of output could be rejected.

[22] For a public service, it is interesting to study the degree of publicness of the good to see the closeness to a pure public good. Duncombe and Yinger decomposed the marginal elasticity of population with respect to cost into two parts: one for publicness (the congestion effect) and one for returns to scale in production of the direct output. They found, as did Brueckner (1980), a high degree of publicness for fire services, while returns to scale in production were increasing, thus resulting in a total marginal elasticity of unity. In an earlier study, Borcherding and Deacon (1972) also found a marginal elasticity of unity.

The Duncombe and Yinger paper only considered fire departments with full-time firemen, while Duncombe and Brudney (1995) also included an analysis of the problem with having volunteer firemen. They showed that the substitution elasticity between paid and volunteer firemen was surprisingly inelastic.

A10. Juås (1994[23], 1995[24])

Using cost-benefit analysis Juås measured the "time value" of a 5- or 10-minute delay for the fire brigade. A 5-minute delay is the difference between a full- and a part-time crew, while a 10-minute delay could result if a remote part-time station is closed down and the central full-time brigade takes over instead. The time factor can often be fixed at zero or a very low value, especially when it comes to false alarms. For non-zero-delay turn-outs Juås computed different values for each of ten reported turn-out objects.

Juås started with a discussion about the special kind of business that fire and rescue service is. She described it as having three specific attributes: it is an organic process, the time dimension is important, and the product is stochastic. These three attributes are described below (my translation from Juås, 1994).

Organic process:

> "If you produce a dead thing like a car, the car stands waiting for you until you do something, such as weld or paint it. A fire however, is an organic process, which changes over time. The fireman's production is to change the process. So output can be defined as the difference between the potential course of events and what actually happened (using the fire brigade)." (p. 5)

Time:

> "A child can put out a fire just after it has started, but 10 minutes later it is impossible for ten adults. So how large the damage will be depends not just on how much resources are used, but also on when the resources are put in." (p. 8)

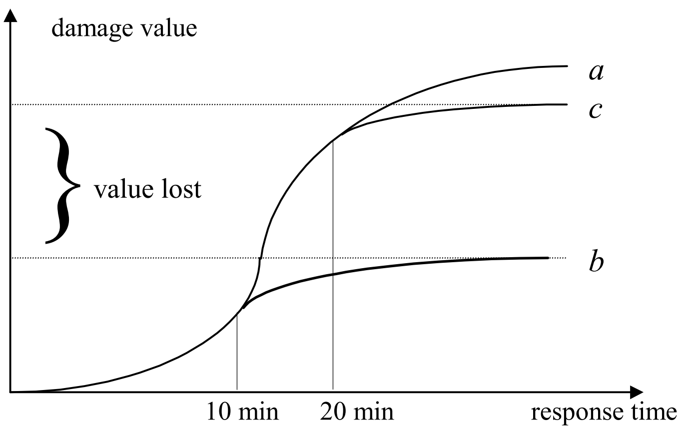
Stochastic:

[23] English summary in Räddningsverket (1995)
[24] English summary in Räddningsverket (1996c) and in Mattsson and Juås (1997)

Fires and fire extinguishing are stochastic processes because:
1. The time from the start of a fire to when an alarm is called in can vary greatly depending on chance.
2. The characteristics of the fire objects. You can never describe an object totally; e.g. someone may have left a window open, which affects the fire.
3. The weather.

The fundamental correlation between response time and value lost due to delay can be seen in Figure A1. What happens with a fire that is not dealt with is described by curve *a*. Curves *b* and *c* shows what happens if the fire and rescue service arrives in 10 and 20 minutes respectively. The value lost of a 10-minute delay is shown in the figure.

Figure A1. Fundamental correlation between response time and value lost due to delay



Source: Räddningsverket (1996c) p 55

In order to find the marginal product of a fire station with respect to building fires, Juås (1994) calculated the saved value for each fire object of a 5- or 10-minute delay. These values were then weighted with the distribution of objects for the fire service in question, and then multiplied with the number of alarms. Juås (1995) reported similar calculations for other types of fires and accidents.

A11. Schaenman and Swartz (1974)

Looking at the problem of how to measure fire protection productivity Schaenman and Swartz focused on the two principal sub-functions of the fire service: fire prevention and fire suppression. They proposed measures for output, input, and community and fire department characteristics, which are presented in Table A1.

There is no measure indicating how many fires a fire service has prevented; the best one can do is to use the number of fires actually occurring, which can be related to the population, or perhaps better, to the number of daytime population. It can also be broken down to different types of fires, and can be related to the number of fire inspections. However, many other factors determine the number of fires, so these measures should perhaps be used with caution in inter-city studies.

Table A1. Output measures proposed by Schaenman and Swartz (1974).

Activity	Interesting variables	Proposed quantity measures
Fire prevention	• Reported fire incident rates	- Number of fires per 1000 population - Number of residential building fires per dwelling units - Number of commercial/ industrial/institutional fires per 1000 occupancies - Number of fires in inspected properties per 1000 occupancies
	• Number of unreported fires per 1000 population	
Fire suppression	• Response time after fire alarm	- Average response time - Percentage of responses taking less than 'x' minutes
	• Percentage of fires with spread of damage after arrival of first fire unit limited to 'x' or less damage steps	
	• Value of property loss per building fire	
Overall fire protection	• Fire related deaths	- Number of civilian deaths per 100 000 population - Number of civilian deaths per 1000 fires Number of fireman deaths per 1000 fire employees - Number of fireman deaths per 1000 fires
	• Fire related injuries	- Civilian, Firemen
	• Property loss per $1000 property value protected	
	• Insurance rating of fire departments	

For a measure of fire suppression, the value of the damage may seem most natural, but a problem with this measure is that it is not in control of the fire service. To fight a fire of $100,000 dollars may require the same effort as to fight a fire of $1,000,000 dollars. How many firemen and trucks will be sent to a fire is most often fixed in advance and thus not dependent on the specific fire in question. (Very large fires, which require firemen from several departments, are of course an exception.) The property loss is also dependent on fire prevention.

A better way to measure fire suppression, according to Schaenman and Swartz, is to concentrate on the degree of spread after arrival. For example it is better that the fire has only spread from one room to two rooms rather than three rooms, before it is finally extinguished. Of course, the measure is discrete and non-linear in nature. A continuous measure could be "percentage of fires in which damage was limited to x steps after arrival".

Another measure of fire suppression is the response time to a fire alarm. The faster response, the less the loss. Response time is of course only one factor influencing fire loss, but it is one over which the fire department has quite good control. It depends on the location of the fire station, the efficiency of the firemen to get going, and the ability to find the fastest way to the fire. Traffic and road conditions and weather are of course uncontrollable factors.

The overall fire protection measures are those that are dependent both on fire prevention and fire suppression: number of deaths and casualties, insurance rating, and property loss.

A12. Morley (1986)

Morley provided four examples of measures to use for fire protection direct outputs: fire prevention inspections, responses to non-fire emergencies, fires extinguished (by type of fire), and fire investigation. Morley also provided four examples of final outputs (consequences): total time for fire suppression, casualties (injuries and deaths), number of people saved and property. Final output could also be measured using survey methods, e.g. contingent valuation, i.e. ask people how much they value the fire and rescue service.

A13. Räddningsverket (1996a, 1996b)

In two papers (1996a is a theoretical discussion, and 1996b is a practical example from one municipality), The Swedish Rescue Services Agency (Räddningsverket) presented a suggestion of how to measure the output of a fire or rescue brigade. Output was divided into three categories: result, effect, and need. Examples of *result* measures are the number and total time of turnouts, the number of firemen employed, and the number of inspections. Examples of *effect* measures are saved value, damaged value, number of dead, number of injured, and higher awareness of risks. *Wants* is the safety people feel, exemplified by the number of satisfied consumers, and actual cost relative to budgeted cost.

In terminology used here: result is an intermediate output, and wants a final output, while it is difficult to classify effect. It seems to be something in between.

References

Ahlbrandt, R., 1973, Efficiency in the provision of fire and rescue brigades, *Public Choice*, 16, 1-15.

Borcherding, T.E. and R.T. Deacon, 1972, The demand for the services of non-federal governments, *American Economic Review*, 891-901.

Bouckaert, G., 1992, Productivity analysis in the public sector: The case of the fire and rescue service, *International Review of Administrative Sciences*, 58, 175-200.

Bradford, D., R. Malt, and W. Oat, 1969, The rising cost of local public services: Some evidence and reflections, *National Tax Journal*, 22(2), 185-202.

Brueckner, J., 1980, Congested public goods: The case of fire protection, *Journal of Public Economics*, 15, 45-58.

Coulter, P., 1979, Organizational effectiveness in the public sector: the example of municipal fire protection, *Administrative Science Quarterly*, 24, 65-81.

Duncombe, W., 1991, Demand for local public services revisited: The case of fire protection, *Public Finance Quarterly*, 19, 412-436.

Duncombe, W., 1992, Costs and factor substitution in the provision of local fire and rescue brigades, *Review of Economics and Statistics*, LXXIV, 180-185.

Duncombe, W., and J. Brudney, 1995, The optimal mix of volunteer and paid staff in local governments: an application to municipal fire departments, *Public Finance Quarterly*, 23(3), 356-384.

Duncombe, W., J. Miner, and J. Ruggiero, 1997, Empirical evaluation of bureaucratic models of inefficiency, *Public Choice*, 93, 1-18.

Duncombe, W., and J. Yinger, 1993, An analysis of returns to scale in public production, with an application to fire protection, *Journal of Public Economics*, 52, 49-72.

Hirsch, W.Z., 1959, Expenditure implications of metropolitan growth and consolidation, *Review of Economics and Statistics*, 41, 232-241.

Hirsch, W.Z., 1973, *Urban Economic Analysis,* McGraw-Hill, New York.

Johansson, C., 1992, *Kostnadsvariationer inom det kommunala brand- och räddningsväsendet*, Meddelanden från ekonomisk-statsvetenskapliga fakulteten vid Åbo Akademi, Series A:366.

Juås, B., 1994, *Räddningstjänst vid byggnadsbränder*, Högskolan i Karlstad, Forskningsrapport 94:7.

Juås, B., 1995, *Tidsfaktorns betydelse vid räddningstjänstens insatser*, Högskolan i Karlstad, Forskningsrapport 95:15.

Kristensen, O., 1983, Public versus private provision of governmental services: The case of Danish fire protection services, *Urban Studies*, 20, 1-9.

Mattsson B, and B. Juås, 1997, The importance of the time factor in fire and rescue service operations in Sweden, *Accident Analysis and Prevention*, 29 (6), 849-857.

Morley, E., 1986, *A Practitioners Guide to Public Sector Productivity Improvement*, van Nostrand Reinhold Company, New York.

Räddningsverket (Swedish Rescue Services Agency), 1995, *Optimal Fire Safety*, Karlstad.

Räddningsverket (Swedish Rescue Services Agency), 1996a, *Mått och indikatorer - tillämpning på Räddningstjänsten i Jönköping*, Karlstad.

Räddningsverket (Swedish Rescue Services Agency), 1996b, *Mått och indikatorer - Förslag till mätning av räddningsverksamheten*, Karlstad.

Räddningsverket (Swedish Rescue Services Agency), 1996c, *Optimal Fire Safety 2*, Karlstad.

Sacks, S. R., 1986, Public vs. private production of fire protection services, *Northeast Journal of Business and Economics*, 13(1), 60-69.

Schaenman, P., and J. Swartz, 1974, *Measuring fire protection productivity in local government*, National Fire Protection Association, Boston.

Southwick, L., and Butler, R., 1985, Fire department demand and supply in large cities, *Applied Economics*, 17, 1043-1064.

Wallace, R., 1977, Productivity measurement in the fire and rescue service, *Public Productivity Review*, 2(3), 12-36.

2

Paper 2

Swedish fire and rescue Services' manning levels - a stochastic frontier analysis using panel data

by
Henrik Jaldell

ABSTRACT

This paper studies the productivity and efficiency of the Swedish fire and rescue services during 1989-1995 using a stochastic frontier specification for panel data. The manning level is modelled as a function of risk, environment, and number of turn-outs. The results show that the size of population was the main determinant of manning levels. No productivity change was discovered. Inefficiency was explained by municipality category variables: very large, suburban, and large municipalities were more efficient, while rural and thinly populated municipalities were less efficient. The efficiency differences found were substantial with a mean input saving potential of 30%.

Keywords: fire and rescue service, panel data, productivity, efficiency, public sector

1 Introduction

Efficiency and productivity analyses of various branches of the public sector have increased in popularity in recent years, not only because of an interest in rendering the public sector more effective due to the harsher fiscal environment, but also because of better economic tools with which to perform the studies, such as frontier techniques.[1] Frontier techniques have been used to analyse efficiency and productivity in many public sector branches such as social insurance, courts, post offices, pharmacies etc. However, the fire service has been neglected, probably due to the problem of choosing an appropriate output measure.

The only known studies of fire service efficiency are Ahlbrandt (1973), Kristensen (1983), and Bouckaert (1992), who studied productivity differences between private and public fire services in the state of Washington, in Denmark, and in Belgium respectively. There have also been a few cost and production functions estimated for fire services, including Hirsch (1959), Ahlbrandt (1973), Southwick and Butler (1985), Duncombe and Yinger (1993), and Duncombe and Brudney (1995).

The problem of output is not only that fire protection is a service (Griliches 1992), but also that the activities are quite divergent. Fire prevention, for example, is not totally under the control of the fire service, while fire suppression includes two connected but differently measured activities: waiting for a fire to happen, and then quickly fighting it. Output has been measured differently in the studies mentioned above. Hirsch (1959) and Ahlbrandt (1973) both estimated cost functions where costs were a function of population, a quality index of the fire service, and environmental variables such as area and property values. This is the determinant approach, where costs depend on environment. In Duncombe and Yinger (1993) and Duncombe and Brudney (1995), costs were instead functions of wage, capital, outputs, and environmental variables such as property values and population density. The two outputs (final outputs) were fire suppression, measured as the inverse of the loss per fire, and fire prevention measured

[1] Recent descriptions of the different variants are Fried, Lovell, and Schmidt (1993), and Charnes, Cooper, Lewin and Seiford (1994).

as the inverse of the number of fires. Bouckaert (1992) also used the inverse number of fires as an output for fire prevention, but used the inverse of the number of casualties as an output for fire suppression. Bouckaert defined these as intermediate outputs, and used the number of inhabitants, area and property value as final output. The approach by Duncombe and Yinger, Duncombe and Brudney and Bouckaert is the production function approach, where costs or inputs depend on output.

This study uses an approach similar to those of Hirsch and Ahlbrandt, i.e. the determinant approach. It is assumed that the size of the manning level at the stations, i.e. the firemen on duty prepared to turn out, is a function of the environment, or more specifically how risky the environment is, as discussed in Coombes and Charlton (1994). For example risky industries such as chemical industries requires a larger fire crew. This is also true for environmental factors where higher population, wider area, and higher population density require a larger manning level. The manning level corresponds to labour use and is therefore an input variable.[2] It is also the only input variable, since labour account for about 70% of the total cost of the fire and rescue services, and it is reasonable to assume complementarity between labour and the other inputs.[3] Because there is thus one input and several output variables, an input requirement function is estimated. The reason for using risk and environmental variables rather than cost variables as determinants of the manning level, is that those are the only ones with available data over time in Sweden.

The main objective of this paper is to estimate what actually determines the manning level, and to find patterns of these differences. Since the size of the manning levels in Sweden mainly follow tradition, and not the risk level of the surroundings, there is reason to believe that there exist substantial differences in the manning level among

[2] However, whether the manning level is an input or output variable depends on the situation. Fire suppression is produced in several intermediate steps. One output is what is saved by the fire service by suppressing the fire. One of the inputs used in this step is the number of firemen, i.e. the manning level. However, this is the second output produced; the first output is the attendance level, which is produced whether there is a fire or nor. This in turn determines how fast the fire service can turn out and how many firemen turn out. The faster the better and the more firemen the better.

[3] Examples of input requirement frontiers for the public sector are Kumbhakar and Hjalmarsson (1995a, 1995b), Bjurek, Hjalmarsson and Førsund (1990), and Bjurek and Hjalmarsson (1995); all studied Swedish social security offices.

Swedish municipalities with similar risk environments. The interesting question is to find out the size of the potential efficiency gains. Other interesting questions are whether the harsher fiscal surrounding in Swedish public finances since the recession in 1992 have been reflected in the fire services, which of the above variables in fact are most influential when deciding upon the size of the manning level, and the distribution of inefficiency among different municipality categories.

The availability of panel data for the period 1989-1995 makes it possible to use the stochastic frontier specified by Battese and Coelli (1995).[4] In this specification both the basic stochastic frontier function and the determinants of inefficiency are estimated simultaneously. This method permits estimation of both technical change in the stochastic frontier and time-varying technical inefficiencies, given appropriate parametric specification and distribution assumptions.

Section 2 in this paper describes the fire and rescue services and the data, while the model is introduced and discussed in section 3. Empirical findings are reported in section 4, while section 5 summarises and draws conclusions.

2 The data

According to the Swedish Rescue Service Act of 1987, each municipality in Sweden is directly responsible for fire and other rescue services within its territorial boundaries. Higher responsibility for fire and rescue service, but not direct supervision, lies with the Swedish Rescue Services Agency (Statens Räddningsverk), of the national government. The law states that the municipality is responsible for suppressing the fire or doing a rescue only if a fast response is important or the threatened interest is large. For full-time firemen it shall not take more than 90 seconds to respond to an alarm, while for part-time firemen it shall not take more than 5-6 minutes to respond. Part-time firemen usually have other full-time jobs, but do a 24-hour on-call service every third or fourth week.

[4] Aigner, Lovell, and Schmidt (1977) and Meussen and Van den Broeck (1977) originally proposed the stochastic frontier. Recent surveys of frontier models estimated by econometric methods can be found in Greene (1993, 1997) and, especially for panel data, in Cornwall and Schmidt (1996).

The 288 municipalities in Sweden differ considerable both in population, from 2,800 to 700,000 inhabitants, and in area, from 9 km^2 to 30,000 km^2. Some municipalities have several fire and rescue stations with many full-time paid firemen, while others may only have one station and only part-time firemen, and still other have a mix of full-time and part-time firemen. Altogether there are over 800 fire and rescue stations. The organisation of the fire and rescue services differs somewhat across municipalities. Most municipalities have their own fire and rescue management, but some buy this function from a neighbouring municipality, while some have a collective management for several municipalities.

In order to have one fireman per day in attendance 24 hours per day 5.6 full-time employed men are needed on average (Sträng and Öström, 1994). The cost for having a force of full-time workers is about six times higher than having part-time workers.

Costs per capita thus also vary a lot among the municipalities, from about 200 to 1,100 SEK, with an average of 470 SEK in 1993. Labour costs account on average for about 70% of the total budget; the rest includes costs for material and buildings.

The size of manning level depends on the risk of accidents and fires in the municipality. Risk depends both on probability and on the expected consequences. The independent variables used in this paper are mainly proxies for risk. The number of traffic accidents and the total number of turn-outs to fires (not false automatic alarms) serve as proxies for probability of a turn-out. Fires include those in public buildings, private homes, industries, other buildings, and not in buildings (e.g. in forests). A problem with this variable is that it is quite heterogeneous. A burnt building can be a small shed of little value, or a large building worth millions.

Because the presence of risky industries increases the demand for more firemen, population and the total number of employees in risky industries (pulp/timber, steel/metal, and chemical) are used as proxies for expected consequences.[5] Area is also included as an environmental variable, which takes into account that vast municipalities need more fire stations to be able to turn-out in a reasonable time.

Manning level and turn-out data were obtained from the national Swedish Rescue Services Agency in Karlstad. The data is collected on a fire and rescue service level. Since some of the 288 municipalities co-operate there are 248 services observed during 1989-1995. Three services have missing data in 1989, so the total data set consists of 1733 observations.

Descriptive statistics of the variables are presented in Table 1. The manning level, i.e. the number of firemen in attendance 24 hours a day, has declined over the years, as can be said about traffic accidents, while the total number of fires has fluctuated. Population has increased over time. The area is of course the same, and so is the number of employees in risky industries since only data from one year, 1993, was available.

3 The model

The input requirement function for the manning level in attendance can be written as

$$L = h(X)e^{U+V} \qquad (1)$$

where L is the dependent variable, X are the independent variables specified above, and U is associated with technical inefficiency.[6] V is a random term describing white noise.

As discussed by Lovell (1993) and Lee and Schmidt (1993), there are two approaches to the problem of how to explain inefficiency. The first is to include the explanatory variables in the function and then estimate efficiency. This hinges on the assumption that the explanatory variables do not affect technical efficiency, but only the frontier function.

[5] Data source: SNI 33, 35 and 37 in Statistiska Meddelanden I209501, Table 4.

[6] Inefficiency in the input requirement frontier corresponds to technical inefficiency, and not to allocative and overall efficiency as defined by Farrell (1957). The reason for it not being overall efficiency is that the dependent variable, L, is not weighted by the respective cost of part- and full-time firemen.

Another approach is to estimate the frontier function without explanatory variables, thereby getting the estimates of efficiency in a first step, and then regress the efficiency estimates on the explanatory variables in a second step. The problem with the two-stage approach is that it is inconsistent in its assumptions regarding the independence of the inefficiency effects in the two estimation stages.

In this paper the single-stage procedure of Battese and Coelli (1995) is used instead, in which the parameters of the stochastic frontier and the inefficiency model are estimated simultaneously. This method permits estimation of both technical changes in the stochastic frontier and time-varying technical inefficiencies.

Table 1. Mean, standard deviation, minimum and maximum of the variables, totally and per year.

	Mean	St dev.	Min.	Max.
Manning level	17.94	12.93	5	94
Traffic accidents	33.33	39.38	0	364
Total fires	152.0	302.3	9	3620
Employees in risky industries	546.70	749.30	0	4593
Population	34,880	63,766	3,392	711,119
Area	1,645.7	2,646.9	19	1,9446
Part-time (=1) / full-time (=0)	0.834	0.232	0	1

	Mean per year 1989	1990	1991	1992	1993	1994	1995
n	245	248	248	248	248	248	248
Manning level	18.43	18.24	18.11	17.83	17.80	17.59	17.52
Traffic accidents	35.69	36.39	33.73	33.20	31.52	30.58	32.21
Total fires	163.4	153.6	145.4	156.6	146.6	157.0	141.8
Employees in risky indust.	550.46	546.08	546.08	546.08	546.08	546.08	546.08
Population	34,470	34,434	34,648	34,840	35,054	35,307	35,394
Area	1,659.8	1,646.3	1,646.3	1,646.3	1,646.3	1,646.3	1,646.3
Part-time (=1)/ full-time (=0)	0.835	0.838	0.837	0.834	0.832	0.832	0.831

Assuming a translog function, the model is:

$$lnL_{it} = \beta_0 + \sum_{j=1}^{6} \beta_j x_{jit} + \sum_{j=1}^{6} \sum_{k=1}^{6} \beta_{jk} x_{jit} x_{kit} + \sum_{l=1}^{2} \theta_l D_{lit} + \zeta P_{it} + U_{it} + V_{it} \qquad (2)$$

where i=1,..., 248 are the individual fire and rescue services; t=1,..., 7 are the time periods; and j=1,...,6 are the explanatory variables;

L is the manning level in attendance (mean on duty per hour);

x_1 is the logarithm of the number of traffic accidents; [7]

x_2 is the logarithm of the total number of fires (not false alarms);

x_3 is the logarithm of the total number of employees in risky industries (pulp/timber, steel/metal, and chemicals industries); [7]

x_4 is the logarithm of the population;

x_5 is the logarithm of the area in square kilometres;

x_6 is the year of observation, where x_6=1,..., 6;

D_l are dummy variables for traffic accidents and risky industries: D_1 equals one if number of traffic accidents is zero, and zero otherwise, while D_2 equals one if number of employees in risky industries is zero, and zero, otherwise; [7]

P is a variable indicating if the fire and rescue service consists of part-time (=1) or full-time (=0) firemen, or a mix;

V_{it} is a random error term distributed iid $N(0, \sigma_v^2)$ and independent of the U_{it}'s;

U_{it} represent non-negative technical inefficiency, assumed to be independently distributed as truncations at zero of the normal distribution with mean, μ_{it}, and variance, σ_u^2; where

$$\mu_{it} = \delta_0 + \sum_{m=1}^{7} \delta_m C_{mit} + \delta_p P + \delta_t x_6 \quad (3)$$

where C_m are dummy variables associated with a classification of municipalities made by Statistics Sweden (SCB); very large (C_1), suburban (C_2), large (C_3), medium (C_4), industrial (C_5), rural (C_6), and thinly populated (C_7).

The estimation of the model was done by maximum-likelihood using the program FRONTIER 4.1. For a further description of the estimation procedure see Coelli (1996).

The input saving efficiency measure (Førsund and Hjalmarsson, 1987) of the i'th term in the t'th time period is defined by $E_{it}= exp(-U_{it})$. The inefficiencies reported in this paper are the reciprocal of $E_{it,}$ and have a minimum value of one; higher numbers signify less efficiency. The number gives the proportion by which the actual manning level exceeds the corresponding stochastic frontier level, given the value of all other variables.

The time variable is incorporated both in the basic function and in the inefficiency term; which makes it possible to distinguish between a shift of the frontier function, and movements within the input requirement set. The shift of the frontier function is often called technical or productivity change, while movements within the requirement set are called time inefficiency effects, or catching-up effects. Productivity change is calculated as

$$\frac{\partial lnL}{\partial x_6} = \beta_6 + 2\beta_{66} + \sum_{k=1}^{6} \beta_{6k} lnx_k \qquad (4)$$

while movements within the input requirement set over time is calculated using the method in Battese and Broca (1997) and Battese, Heshmati and Hjalmarsson (1998) to adjust the parameter estimate for x_6 (time) in the inefficiency term for the fact that the time variable is included both in the frontier function and in the inefficiency term.

4 Empirical results

4.1 Hypothesis tests

Formal tests of the hypotheses that the inefficiency effects are absent or that they have simpler distributions are presented in Table 2.[8]

The first null hypothesis suggests a Cobb-Douglas specification instead of the translog function. A Cobb-Douglas input requirement function would imply non-convex transformation surfaces, but this hypothesis is rejected.

[7] The method of how to deal with zeros in the regressors was suggested by Battese (1997). Here x_1=max{no. of traffic accidents; D_1}, and x_3=max{no. of employees in risky industries; D_2}.

[8] The likelihood ratio test statistic is: λ= -2(log-likelihood (H_0) - log-likelihood (H_1)), and is approximately χ^2-distributed with degrees of freedom equal to the number of parameters assumed to be zero in the null hypothesis.

Table 2. Hypothesis tests of Cobb-Douglas specification, time variables, inefficiency and category dummies.

H_0:		log - likelihood	λ-statistic	$\chi^2_{0.05}$ value	Decision
Cobb-Douglas	$\beta_{jk}=0$ j=k=1,...,6	31.4	137.6	32.7	H_0 rejected
No change over time	$\beta_6=\beta_{6k}=0$ j=1,..., 6	96.6	7.1	14.1	H_0 not rejectec
No inefficiency effects	$\gamma=\delta_0=....=\delta_9=\delta_p=\delta_t$	8.6	183.1	18.3	H_0 rejected
Category dummies all = zero	$\delta_1=....=\delta_9$	8.5	183.4	14.1	H_0 rejected

The second null hypothesis that all time variable-parameters all being zero is not rejected at the 5% significance level, which means that there is no statistically significant time-trend in the data. This is somewhat surprising since the descriptive statistics (Table 1) seem to indicate a decline of the manning level over time. This means that the decline of the manning level noted above correspond to a decline of values of the independent variables, i.e. to a movement along the frontier and not to a shift of the frontier itself.

The third null hypothesis, that all inefficiency effects are zero, is rejected, as is the fourth null hypothesis, that all the municipality category parameters in the inefficiency model are zero. This means that these categories are relevant for looking at inefficiency differences.

4.2 Marginal elasticities

The mean marginal elasticities for the variables of the input requirement function are presented in Table 3. (The parameter estimates are presented in Table A1 in the Appendix.)

Population influences the manning level positively, and perhaps increasing slightly over time. It is most influential for very large and large municipalities, and least influential for industrial, rural, and thinly populated municipalities, but the differences are small.

The marginal elasticity is 0.7, implying that a municipality with a 10% higher population has 7% more firemen standing by. This result correspond with the idea that fire protection is a public good, which is one reason for fire and rescue services being public companies. If fire protection is a public good the marginal elasticity with respect to population should be substantially less than one. It is to be remembered that the manning level reflects the suppression side of the fire and rescue service. Earlier studies have found elasticities closer to unity (Brueckner, 1980, and Duncombe and Yinger, 1993).

The marginal elasticities for traffic accidents, total number of fires, and risky industries are all near zero. A municipality with 10% more of one of these variables requires only 2-3% more firemen standing by. Traffic accidents are, however, somewhat influential for very large and large municipalities, and has decreased over time.

That the total number of fires mean marginal effect is close to zero may seem surprising. Even more surprising is that for the different categories the marginal effect is either positive or negative. Negative especially for very large and large municipalities, and positive for suburban municipalities. However, it is not the total number of fires that should influence the number of firemen, but the size of them.

The marginal elasticity of employees in risky industries has increased somewhat over time, the largest effect is in thinly populated municipalities.

The marginal elasticity for area is about 0.15, increasing slightly over time, and is lowest for very large and suburban municipalities, highest for thinly populated municipalities.

Table 3. Marginal elasticities, technical change and inefficiency change; mean per year and category type.

	traffic	total fires	risky industries	popu-lation	area	time variables	ineffici-ency term
Overall:							
mean	0.025	0.021	0.035	0.715	0.152	-0.017	0.308
st. dev.	0.054	0.148	0.031	0.068	0.063	0.007	0.256
max	0.212	0.520	0.115	0.934	0.407	0.009	0.567
min	-0.133	-0.345	-0.079	0.501	-0.148	-0.050	-0.392
Mean per year:							
1989	0.043	0.020	0.031	0.702	0.145	-0.023	0.293
1990	0.036	0.030	0.032	0.702	0.152	-0.021	0.298
1991	0.029	0.029	0.033	0.707	0.154	-0.019	0.303
1992	0.027	0.016	0.035	0.717	0.149	-0.017	0.308
1993	0.017	0.023	0.036	0.718	0.156	-0.014	0.314
1994	0.016	0.005	0.038	0.730	0.148	-0.012	0.319
1995	0.009	0.024	0.041	0.727	0.162	-0.011	0.324
Mean per category:							
very large (n=3)	0.152	-0.138	-0.021	0.801	0.055	-0.013	0.231
suburban (n=22)	0.024	0.141	-0.010	0.687	0.103	-0.017	0.034
large (n=24)	0.095	-0.124	0.038	0.782	0.160	-0.009	-0.370
medium (n=37)	0.048	-0.015	0.036	0.729	0.164	-0.015	0.344
industrial (n=47)	0.008	0.072	0.030	0.676	0.149	-0.018	0.399
rural (n=33)	0.003	0.057	0.038	0.699	0.148	-0.019	0.472
thinly pop. (n=31)	0.010	-0.088	0.081	0.757	0.182	-0.017	0.547
others (n=51)	0.008	0.068	0.031	0.694	0.154	-0.018	0.388

Time has a negative influence on the input requirement, indicating a positive productivity change over time of 1.7% per year, but decreasing over time. Large municipalities have increased their productivity least, while rural municipalities have increased their productivity most. However, as noted above among the hypothesis tests, the possibility that the time variables all equal zero could not be rejected.

The parameter estimate for the variable of part-time versus full-time firemen is 1.04 (Table A1) indicating that, other things equal, part-time fire and rescue services use almost three times as many firemen as do full-time ones.[9]

[9] Percentage difference between part- and full-time is calculated as 100*exp(c)-1=182.3.

4.3 Inefficiency

The inefficiency term in Table 3 shows a slight increase in inefficiency over time. However, the change is not statistically significant (δ_t in Table A1).

For the different categories inefficiency is statistically significantly (Table A1) smaller for suburban and large municipalities, and greater for rural and thinly populated municipalities. Whether the fire and rescue service has part- or full-time firemen, has no effect on inefficiency, other things equal.

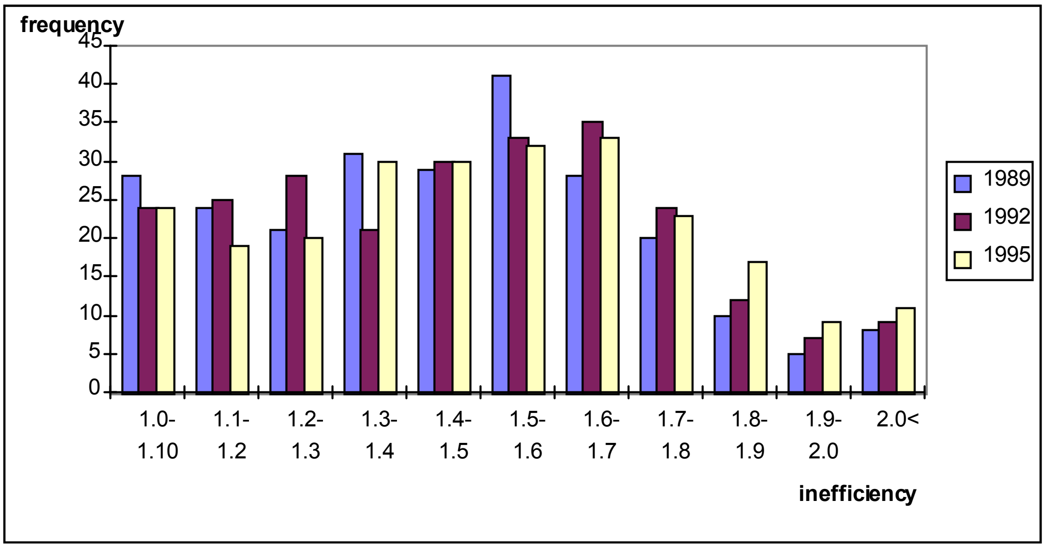
Table 4. Inefficiency per fire service: Mean, standard deviation, min, and max; overall and per category.

	mean (weighted by manning level)	**mean** (weighted by population)	**mean** (arithmetic)	**standard deviation**	**min**	**max**
Overall:	1.453	1.319	1.480	0.274	1.043	2.380
Per category:						
very large	1.351	1.315	1.323	0.116	1.199	1.557
suburb	1.182	1.165	1.186	0.105	1.072	1.596
large	1.078	1.073	1.075	0.021	1.043	1.139
medium	1.487	1.447	1.452	0.181	1.12	1.886
industrial	1.576	1.513	1.531	0.229	1.097	2.228
rural	1.659	1.630	1.635	0.191	1.756	2.22
thinly pop.	1.805	1.758	1.759	0.24	1.232	2.38
others	1.572	1.527	1.513	0.215	1.085	2.364

Table 4 shows the mean inefficiencies per category. The arithmetic mean of all inefficiencies is 1.48 indicating that on average the fire and rescue services use 48% more firemen than the frontier fire and rescue services could do. Weighting the mean by the manning level and by the population give lower values, 1.45 and 1.32 respectively. These values also indicate a substantial input saving potential. The maximum inefficiency is 2.38 (a thinly populated municipality), indicating that 138% too many firemen are used there. The minimum inefficiency is 1.04 (a large municipality), indicating that only 4% too many firemen are used there. Large and suburban municipalities have the lowest inefficiencies, while rural and thinly populated municipalities have the highest. Especially large, but also suburban, and very large municipalities have lower standard deviations; thus it seems like these categories are more aware of the size of the manning level of similar municipalities than other categories.

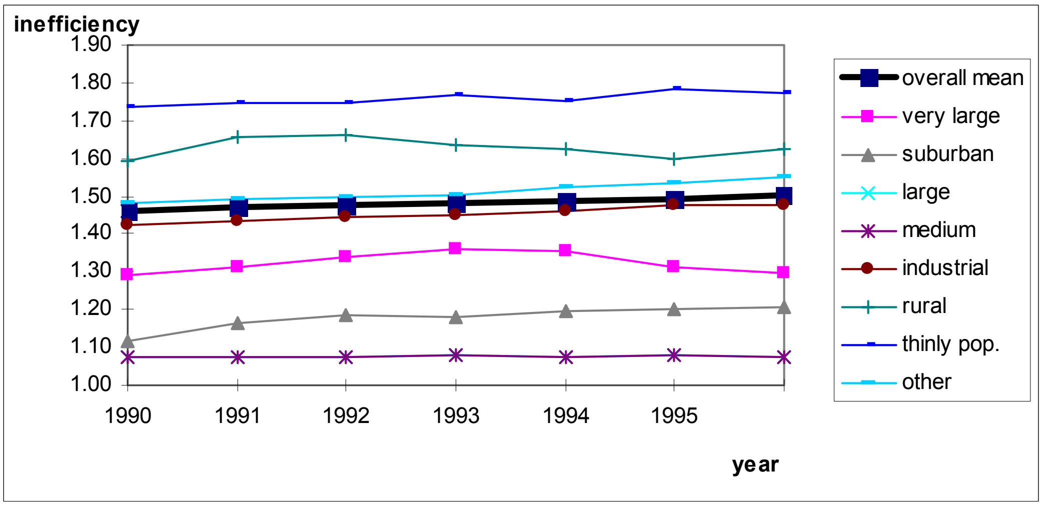
The distribution of inefficiency for the years 1989, 1992, and 1995 is shown in Figure 1. There are fewer most efficient fire and rescue services, and more least inefficient fire and rescue services over the years. The figure confirms the result of the increase in inefficiency over time noted above. The reason for this is hard to explain. The tougher budget constraint of the municipalities since the recession in 1992 should have resulted in the opposite result, but it seems like tradition is still a main determinant of the size of the manning level. Thus it seems like the fire and rescue services close to the frontier has increased their productivity, but that inefficient services has not changed their manning level, thus making them more inefficient.

Figure 1. Distribution of inefficiency for 1989, 1992, and 1995.



Mean inefficiency per category and per year is shown in Figure 2. The time trend does not change very much within each category, and so does not reveal any information for explaining the result above.

Figure 2. Mean inefficiency per category 1989-1995.



5 Summary and conclusions

This study used a stochastic frontier panel data method to estimate an input requirement function for the manning level in fire and rescue services in Sweden. Environmental variables were included in the specification to control for differences in the risk environment. The estimation involved data from 248 fire and rescue services for the period 1989-1995.

The main empirical findings are

- Population is the only substantial determinant of the manning level:
  - 10% higher population leads to 7% higher manning level, i.e. fire suppression does seem to be a somewhat public good.
- The risk variables (traffic accidents, number of fires and risky industries) have practically no influence on the size of the manning level.
- The mean time-trend is negative, indicating an increase in productivity of 1.7% per year over the period, but the hypothesis that all time-trend variables are zero cannot be rejected.
- There are considerable differences in inefficiencies:
  - mean inefficiency is 1.48, indicating an input saving potential of over 30%;

- suburban and large municipalities are most efficient, and also have lowest standard deviations of inefficiency;
- rural and thinly populated municipalities are least efficient.

- There is no statistically significant time effect for mean inefficiency, but there are fewer of the most efficient fire and rescue services, and more of the least efficient over time.

The finding of no significant productivity change and time inefficiency effects are surprising, since a harsher fiscal surrounding since the recession in 1992 might have led to a decrease in the number of firemen in attendance, and thus to efficiency gains. However, it seems like the fire and rescue services have been lucky in the budget cut process.

The substantial differences in inefficiency are not surprising, since the choice of the manning level is mainly based on tradition, and not on a study of the risk of the environment.

It is to be noted that differences in risk-aversity between the municipalities are not taken into account in this study. Frontier units may be regarded as ”too” efficient by other fire and rescue services, i.e. the frontier units use not enough firemen for producing a reasonable safety level. This may also explain why there has been no productivity increase; the frontier is already reached.

References

Ahlbrandt, R. C., 1973, Efficiency in the provision of fire services, *Public Choice*, 16, 1-15.

Aigner, D., C. A. K. Lovell and P. Schmidt, 1977, Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, 6, 21-37.

Battese, G. E., 1997, A note on the estimation of Cobb-Douglas production functions when some explanatory variables have zero values, *Journal of Agriculture Economics*,48(2), 250-252.

Battese, G. E. and S. S. Broca, 1997, Functional forms of stochastic frontier production functions and models for technical inefficiency effects: a comparative study for wheat farmers in Pakistan, *Journal of Productivity Analysis*, 8, 395-414

Battese, G. E. and T. J. Coelli, 1995, A model for technical inefficiency effects in a stochastic frontier production function for panel data, *Empirical Economics*, 20, 325-332.

Battese, G. E., A. Heshmati and L. Hjalmarsson, 1998, Efficiency of labour use in the Swedish banking industry: a stochastic frontier approach, *Working paper no. 289*, Stockholm School of Economics, Sweden

Bjurek H., L. Hjalmarsson and F. Førsund, 1990, Deterministic parametric and nonparametric estimation of efficiency in service production, *Journal of Econometrics*, 46, 213-227.

Bjurek H. and L. Hjalmarsson, 1995, Productivity in multiple output public service: a quadratic frontier function and Malmquist index approach, *Journal of Public Economics*, 56, 447-460.

Bouckaert, G., 1992, Productivity analysis in the public sector: the case of the fire service, *International Review of Administrative Sciences*, 58, 175-200.

Brueckner, J., 1980, Congested public goods: The case of fire protection, *Journal of Public Economics*, 15, 45-58.

Coelli, T. J., 1996, A guide to Frontier version 4.1, *CEPA working paper 96/07*, University of New England, Armidale, Australia.

Coombes, M., and M. Charlton, 1994, Population distribution as a factor in the costs of fire services, *Environment and Planning C: Government and Policy*, 12, 53-70.

Cornwall C., and P. Schmidt, 1996, Production frontiers and efficiency, ch. 32 in L. Matyas and P. Sevestre (eds.), *The Econometrics of Panel Data*, 2nd ed., Kluwer Academic Publishers.

Ds 1994:24, Den offentliga sektorns produktivitetsutveckling 1980-1992 (Productivity in the public sector 1980-1992), *Rapport till expertgruppen för offentlig ekonomi*, Stockholm.

Duncombe, W., and D. Brudney, 1995, The optimal mix of volunteer and paid staff in local governments: an application to municipal fire departments, *Public Finance Quarterly*, 23(3), 356-384.

Duncombe, W. and J. Yinger, 1993, An analysis of returns to scale in public production, with an application to fire protection, *Journal of Public Economics*, 52, 49-72.

Farrell, M. J., 1957,The measurement of productive efficiency, *Journal of the Royal Statistical Society*, ser. A, 120.III, 211-281.

Førsund, F. and L. Hjalmarsson, 1987, *Analysis of industrial structure: a putty-clay approach*, Industriens utredningsinstitut, Stockholm.

Greene, W. H., 1993, The econometric approach to efficiency analysis, ch. 2 in H. O. Fried, C. A. K. Lovell and S. S. Schmidt (eds.), *The Measurement of Productive Efficiency*, Oxford University Press.

Greene, W. H., 1997, Frontier production functions, ch. 3 in M. H. Pesaran and P. Schmidt (eds.) *Handbook of Applied Econometrics, Vol. 2: Microeconomics*, Blackwell Publishers.

Heshmati, A., 1994, *Estimating Technical Efficiency, Productivity Growth and Selectivity Bias Using Rotating Panel Data: An Application to Swedish Agriculture*, Ekonomiska studier no. 50 (PhD dissertation), Göteborg University, Sweden.

Hirsch, W. Z., 1959, Expenditure implications of metropolitan growth and consolidation, *Review of Economics and Statistics*, 41, 232-241.

Kristensen, Ole (1983), Public versus private provision of governmental services: The case of Danish fire protection services, *Urban Studies*, 20, 1-9.

Kumbhakar, S. and L. Hjalmarsson, 1995a, Labour-use efficiency in Swedish social insurance offices, *Journal of Applied Econometrics*, 10, 33-47.

Kumbhakar, S. and L. Hjalmarsson, 1995b, Decomposing technical change with panel data: an application to the public sector, *Scandinavian Journal of Economics,* 311-323.

Lee, Y., and P. Schmidt, 1993, A production frontier model with flexible temporal variation in technical efficiency, ch 8 in H. O. Fried, C. A. K. Lovell and S. S. Schmidt (eds.), *The Measurement of Productive Efficiency*, Oxford University Press.

Lovell, C. A. K., 1993, Production frontiers and productive efficiency, ch. 1 in H. O. Fried, C. A. K. Lovell and S. S. Schmidt (eds.), *The Measurement of Productive Efficiency*, Oxford University Press.

Meussen, W., and J. van den Broeck, 1977, Efficiency estimation from Cobb-Douglas production functions with composed error, *International Economic Review*, 18(2), 435-444.

Southwick, L. and R. Butler, 1985, Fire department demand and supply in large cities, *Applied Economics*; 17; 1043-1064.

Sträng, D., and B. Öström, 1994, Brandförsvarets kostnader i Norden, Nederländerna och Storbritannien (Fire service costs in the Nordic, the Netherlands, and Great Britain), Foa Rapport A 10056-1.3, Försvarets forskningsanstalt, Stockholm.

Appendix

Table A1. Parameter estimates.

Variable	Parameter	Estimate	Standard error	Significant 5%
Function:				
Intercept	β_0	-5.04018	1.11266	¤
Traffic accidents=x_1	β_1	-0.23653	0.16974	
Total fires=x_2	β_2	0.70142	0.15990	¤
Risk industries=x_3	β_3	0.01357	0.06102	
Population=x_4	β_4	0.28972	0.22957	
Area=x_5	β_5	-0.09772	0.10641	
Year=x_6	β_6	-0.06836	0.05090	
(Traffic accidents)^2	β_{11}	0.01028	0.01048	
(Total fires)^2	β_{22}	-0.07439	0.02043	¤
(Risk industries)^2	β_{33}	0.01342	0.00366	¤
(Population)^2	β_{44}	-0.00038	0.01328	
(Area)^2	β_{55}	0.00773	0.00306	¤
(Year)^2	β_{66}	0.00100	0.00167	
Traffic accidents * Total fires	β_{12}	0.01666	0.01244	
Traffic accidents * Risk industries	β_{13}	0.00323	0.00340	
Traffic accidents * Population	β_{14}	-0.00364	0.01187	
Traffic accidents * Area	β_{15}	0.00798	0.00459	
Traffic accidents * Year	β_{16}	-0.00230	0.00276	
Total fires * Risk industries	β_{23}	-0.01064	0.00394	¤
Total fires * Population	β_{24}	0.03107	0.01108	¤
Total fires * Area	β_{25}	-0.04584	0.00682	¤
Total fires * Year	β_{26}	-0.00131	0.00306	
Risk industries * Population	β_{34}	-0.00950	0.00398	¤
Risk industries * Area	β_{35}	0.01145	0.00188	¤
Risk industries * Year	β_{36}	0.00092	0.00097	
Population * Area	β_{45}	0.01915	0.00745	¤
Population * Year	β_{46}	0.00290	0.00292	
Area * Year	β_{56}	0.00047	0.00119	
D_1 - variable	θ_1	-0.08511	0.11112	
D_2 - variable	θ_2	0.05140	0.09614	
Part-time /full-time =P	ζ	1.03793	0.07872	¤
Inefficiency Term:				
Intercept	δ_0	0.39374	0.09210	¤
very large = C_1	δ_1	-0.18300	0.14315	
Suburban = C_2	δ_2	-0.36293	0.06235	¤
Large = C_3	δ_3	-0.76489	0.22800	¤
Medium = C_4	δ_4	-0.05047	0.02863	
Industrial = C_5	δ_5	0.01132	0.02253	
Rural = C_6	δ_6	0.08788	0.02380	¤
Thinly populated = C_7	δ_7	0.16122	0.03542	¤
Part-time/full-time =P	δ_p	-0.02936	0.10050	
Year = x_6	δ_t	0.00527	0.01039	
Variance Parameters:				
Sigma-squared = $\sigma_v^2 + \sigma_u^2$	σ^2	0.06018	0.00276	
Gamma = σ_u^2 / σ_s^2	γ	0.58314	0.07229	
Log-likelihood:		100.1841		

Paper 3

Measuring the efficiency of Swedish fire services' stand-by level

by
Henrik Jaldell

ABSTRACT

In this paper the DEA-model is used to find efficiency scores and returns to scale corrected for environmental variables. The paper studies the stand-by level of Swedish fire services. This level has two output dimensions 1) the turn-out time (the faster the better), measured as number of people reached within five and ten minutes, and 2) the suppressing power, measured by the total number of firemen turning out (the more the better). The empirical results show that the long-run input saving potential is about 30%, and that, surprisingly, many fire services operate under decreasing returns to scale.

Keywords: fire service, data envelopment analysis, public sector, efficiency, stand-by

1 Introduction

The number of studies of efficiency of the public sector by using production frontier techniques has increased in recent years. However, studies that evaluate fire services have been rare. The only one found is by Bouckaert (1992), which analyses Belgium fire services. The reason is probably because it is difficult to specify an operational output measure, due to both the specific production process and to the lack of data for calculations. The main problem of the production process arises from the fact that the fire service has two distinct objectives, prevention and suppression of fires, and that, in examining the suppression aspect, there are two intermediate outputs in providing the public with the fire safety they demand. The first is the stand-by level, which is produced no matter what happens, and the second is the outcome of the turn-outs. This efficiency study will deal with the stand-by level only.

The main purpose of this paper is to measure the production efficiency of fire services in Sweden. Efficiency is calculated by using the data envelopment analysis (DEA). This method, which was developed by Charnes, Cooper, and Rhodes (1978), was intended to measure Farrell's (1957) efficiency measures without specifying a parametric form.[1] Due to the lack of data on input costs for each individual fire service in Sweden and to the presence of multiple outputs, the non-parametric programming technique of DEA for measuring the frontier has been chosen instead of a parametric econometric technique, like the stochastic frontier.

There are two important outcomes of the stand-by level: 1) the turn-out time (the faster the better), and 2) the suppressing power (the more the better). The first outcome is here measured as the number of people reached within five and ten minutes and the second is measured by the total number of firemen that turn out. These outputs are compared to inputs by using data from 1998. The specification used here is a long-run model, since the number of people reached can only be changed by shifting either the location of the fire stations or changing the whole fire crew from part- to full-time.

[1] Recent presentations of the DEA model can be found in Fried, Lovell and Schmidt (1993), Färe, Grosskopf and Lovell (1994), and Charnes, Cooper, Lewin and Seiford (1994).

Efficiency measures of the public service must often be corrected for environmental variables since the public services have a given task to fulfil no matter the environment. There are two main ways of incorporating the environmental variables into the non-parametric efficiency analysis 1) using a two-stage method by first measuring efficiency with DEA and then regressing them on the environmental variables 2) including the environmental variables directly into the linear programming problem as fixed inputs and outputs. The technical problems of both methods have been thoroughly discussed (see e.g., Lovell, 1993, Coelli, Rao, and Battese, 1998, and Ruggiero, 1998). In this paper both methods will be combined into a three-stage procedure.

The empirical findings, i.e. the efficiency scores and classifications of returns to scale, are presented and discussed with an emphasis on differences between different kinds of municipalities (large, suburban, industrial, thinly populated etc.).

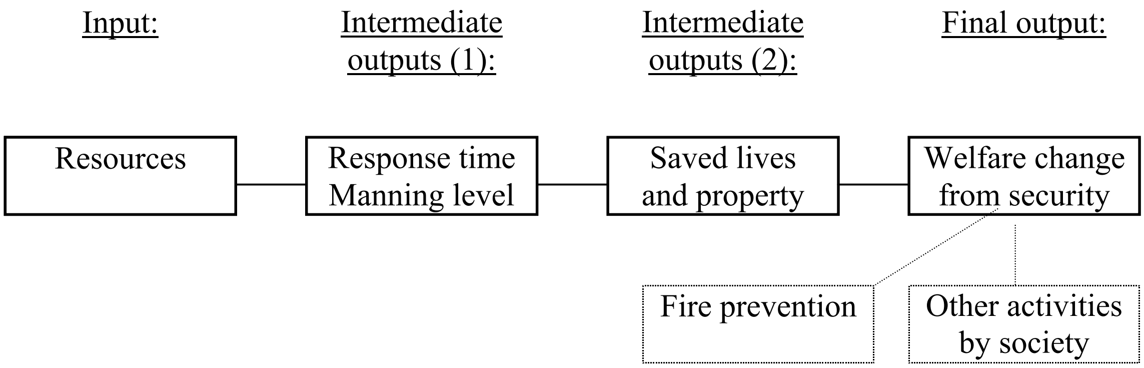
The problem of specifying the output of the fire service is discussed in section 2. The DEA-model is presented in section 3, and the data is used in section 4. In section 5, the results from the DEA-models with and without environmental variables can be found.

2 Output definition

The two main activities of the fire service are to prevent fires and other accidents from happening, and if they happen, to suppress them as effectively as possible. Extending the two-stage model for public sector output suggested by Bradford, Malt and Oates's (1969), fire suppression is described here as a three-stage procedure with two levels of intermediate outputs and one final output, as shown in Figure 1 below.

The fire service receives resources from the municipality to prevent and suppress fires. The first intermediate outputs for fire suppression are the average response time, and the number of firemen initially reaching the fire. Disregarding costs, the shorter the response time the better, and the more firemen arriving at the fire the better. If more firemen are needed to control the fire, it is also important that the reinforcement does not take too long to arrive. For a full-time crew in Sweden it should take about 90 seconds to be on the way, while for a part-time crew it takes between 5 and 6 minutes, since the part-time crew is at work or at home. Part-time firemen usually have other full-time jobs, but do a 24-hour on-call service every third or fourth week. The fire service must decide how many firemen are needed and whether to use full- or part-time firemen, but also, in the long run, how many stations to have, where to locate them, and how to divide the firemen between the stations.

Figure 1. A schematic description of fire suppression activity.



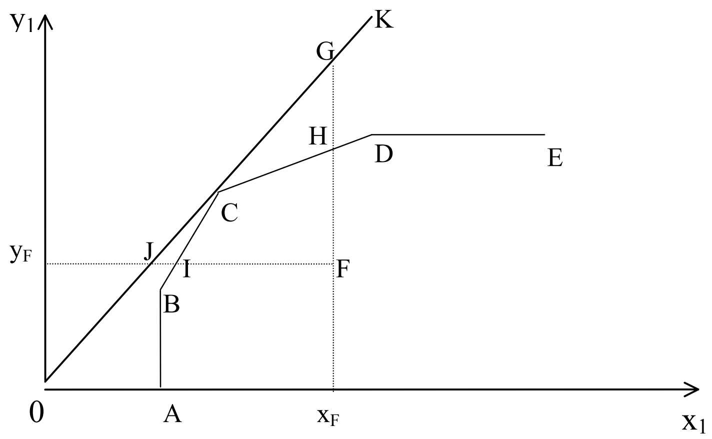
The 240 fire services in Sweden differ considerably in terms of population levels, from 2,800 to 700,000 inhabitants, and the area, from 9 km^2 to 30,000 km^2. Some fire services have several fire stations with many full-time firemen, others have only one station and only part-time firemen, and the rest have a combination of full-time and part-time firemen. Altogether there are over 800 fire and rescue stations. The cost of a force of full-time firemen is about six times higher than that of part-time firemen. The tradition is that larger municipalities have full-time firemen, while smaller municipalities have part-time firemen. The number of firemen should ideally depend on the risk situation. However, in reality the number mostly depends on population and tradition (Jaldell, 2002a).

When an alarm is received at the fire station all firemen often turn out, since they have little chance of knowing how the fire has developed until they reach it. The first intermediate outputs (response time and number of firemen) are the inputs in the production of the second intermediate output; how well the fire service succeeds in suppressing the fire.

The final output is the utility of security or welfare of the inhabitants of the municipality, which depends on all the activities of the fire service (including fire prevention), as well as other factors such as tradition, environmental risk, weather and building conditions.

This paper focuses on the first intermediate output, the stand-by level, and its relation to the costs. It is reasonable to assume that there is a positive correlation between the first

Figure 2. Production frontier with one output and one input.



and second intermediate output, and also between the second intermediate output and the final output.[2]

Studies of productivity differences between fire services are rare. Bouckaert (1992) used a sort of frontier analysis to study productivity differences between Belgian fire services. Using a cost function Ahlbrandt (1973) studied productivity differences between private and public fire services in the state of Washington, USA. Kristensen

[2] The relation between the second intermediate output and its inputs (i.e. the outputs in this paper) has been analysed in Jaldell (2002b).

(1983) did the same for Denmark by comparing cost figures using a discriminant analysis. [3]

3 The model

3.1 Production set

The production set in the data envelopment analysis is characterised by a convex hull, and the technology can be assumed to be constant returns to scale (CRS), non-increasing returns to scale (NIRS), or variable returns to scale (VRS). There are two frontier technologies in Figure 2, the one representing CRS is *0CK,* and the one with VRS is *ABCDEF.* These could be combined to one representing NIRS; *0CDE*. Using Figure 2 we can define the efficiency measures, having one output variable, y_1, and one input variable, x_1. Consider a unit located at point *F*. Since this unit is not a frontier unit, it is possible to either increase output for the same input used, or to decrease input for the same output used, by using the best practise technology described by the frontier. Unit F is thus inefficient.

In many public services outputs are exogenous; the services have a given task to fulfil. Since inputs are endogenous the problem then is to minimise resources. In that case, we should use the input saving measure, i.e. the relative reduction in the amount of inputs needed to produce the observed output using frontier function technique. The definition put forward by Førsund and Hjalmarsson (1974, 1979) for measuring the efficiency of the input saving measure in the VRS is

$$E_I = y_F I / y_F F.$$

If E_I=1 the unit is on the frontier, and is called efficient.

[3] Other studies that use cost and production function to examine the fire service are also rare, but one example is Duncombe and Yinger (1993).

To obtain the input saving efficiency measure under VRS, the following linear program problem must be solved for every unit. For unit *F* the (primal) problem is:

$$E_{1F} = \min_{\lambda_F} \theta_F \tag{1}$$

subject to

$$y_{rF} \leq \sum_{j}^{N} \lambda_{Fj} y_{rj}, \qquad r = 1,\ldots, m \tag{1a}$$

$$\theta_f x_{iF} \geq \sum_{j}^{N} \lambda_{Fj} x_{ij}, \qquad i = 1,\ldots, n \tag{1b}$$

$$\sum_{j}^{N} \lambda_{Fj} = 1 \tag{1c}$$

$$\lambda_{Fj} \geq 0, \quad \forall\, j = 1,\ldots, N \tag{1d}$$

where *m* is the number of outputs, *n* is the number of inputs, and *N* the number of units. Restriction (1a) implies that the reference unit on the frontier, the peer unit (I in Figure 2), must produce at least as much as unit *F*, while restriction (1b) implies that the efficiency adjusted volume of input used by unit *F* must at least amount to the input volume used by the reference unit. Restriction (1c) is the condition for VRS. If this restriction is omitted CRS is implied.

In some output variables there is a limit to how much efficiency can be improved with output increasing efforts. Since the public service is often the single supplier, there may be a maximum level of output that can be produced. In the case of the fire service the population reached within x minutes cannot be higher than the actual population in the specific municipality. Similar problems exists in other public sectors, such as schooling (number of school children) and health care (number of patients). This restriction is also similar to a capacity constraint in the manufacturing industry. In the short-run a plant cannot produce more than the capacity limit.

One way of handling this problem is to use an output increasing measure and add one additional restriction to the linear programming problem (1), so that the number of people reached can never exceed the actual population in the municipality. Another solution, the one chosen here, is to calculate input saving efficiency measures only.

3.2 Returns to scale

In the VRS scale efficiency, i.e. how close a unit actually is to optimal scale, is defined as
$E_3 = y_F J / y_F F$.

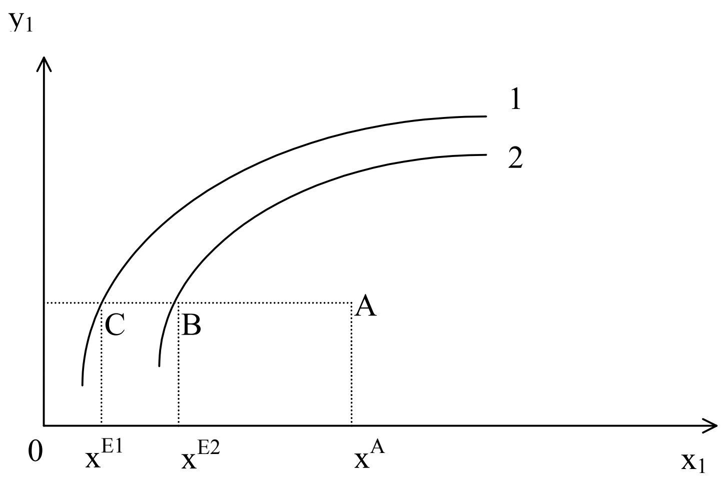
It is also possible to classify the units into different returns to scale classes. If the reference unit on the frontier lies between *ABC,* the units are classified as having increasing returns to scale (IRS), if it lies on *C*, the units can be said to have constant returns to scale, and if it lies on *CDE*, they have decreasing returns to scale (DRS).[4]

However, the restriction on the amount of output that can be produced also affects the interpretation of the returns to scale. For example, a fire service is found to be technically efficient assuming VRS technology, but it may be scale inefficient due to increasing returns to scale. The economic implication is that technically, the resources are used efficiently, but it would be more cost effective to operate under a larger scale. However, since output is restricted, the only way to do this is by merging with the nearby municipality's fire service. A measure of increasing returns to scale with maximum output, is then a signal to co-operate with others. If a fire service operating under decreasing returns to scale is already co-operating with others, then the policy implication would be to split.

3.3 Environmental factors

Since the public sector is influenced by environmental factors outside the control of the public manager, the ”ordinary” DEA-scores may not be an appropriate indicator of efficiency. A fire service in a highly populated area can reach more people than the one located in a thinly populated area. The question then is whether this results in an unfair ranking of the different fire services, or not. One problem with not taking environmental factors into account is that the decision-makers consider these factors when examining and comparing efficiency scores. However, strictly speaking reaching more people with the same costs is more efficient. From Figure 3, it can be seen that if unit A, belongs to a harsher environment, it should be compared to unit B, which also faces a harsher environment, rather than unit C which lies on a less harsh production frontier.

Figure 3. Efficiency under different environments.



Unfortunately, there is no given way of incorporating the environmental variables into the data envelopment analysis. Coelli, Rao and Battese (1998), for example, list several different methods. The main difference between the methods is whether the environmental variables should be included as a restriction solved within the linear programming problem, or if the linear programming problem and the analysis of environmental variables should be divided into two different stages (see below about the two-stage procedure).

Banker and Morey (1986) suggested that the impact of environmental and other non-discretionary variables should be solved within the linear programming problem by adding one additional restriction to equation (1). The restriction is

$$z_{kF} \geq \Sigma_j \lambda_{Fj} z_{kj}, \qquad (1e)$$

where z_k, k=1,…K, are the non-discretionary variables (inputs in this case). [5]

---

[4] The method could give rise to different returns to scale classifications depending on whether output increasing or input saving measures is used. In this paper, however, only input saving measures are used.

[5] Ruggiero (1996) criticised this approach, because it may lead to that the reference set includes units with better surroundings than the unit in question. Instead he proposed another way of setting up the linear

Another problem is to choose the environmental variables to be included in the analysis. The literature on production frontiers has described an ex post way, but no ex ante way. Ex ante, it is reasonable to assume that the environmental variables to choose should affect the efficiency scores only, not the returns to scale of the frontier technology.

Ex post, environmental variables can be checked for their connection with the efficiency scores from the basic model. This is the second step in the two-stage method presented in the literature (e.g. Lovell, 1993, and Coelli et al, 1998). In this step the efficiency scores are regressed on the interesting environmental variables. Since the dependent variable is restricted between 0 and 1, it must be transformed, or a limited dependent procedure must be used.[6] The first step is to measure efficiency without environmental variables.

In this paper, there will also be a third step: calculating efficiency scores once more in the way suggested by Banker and Morey (1986), i.e. including the environmental variables that have significant and substantial marginal effects upon the efficiency scores as fixed variables.[7]

4 Data

The interesting input variables here are costs of labour, capital, material etc. These costs are, however, not easy to homogenise among the fire services, since different municipalities use different accounting principles. For example the municipalities handle pensions in different ways; while some let the different public services take care of the pension fees others do this centrally. The same problem arises for capital costs especially property costs. Some fire services have to pay rents for the buildings to the local authority, while in other municipalities there is no rent. The rents paid may also be

---

program so that the units are only compared to units with as least as harsh surroundings. Unfortunately only one environmental variable can be used.

[6] A second variant of the two-stage method is to reverse the order, as suggested by Grosskopf et al (1997). First regress the outputs or inputs on the environmental variables, and then use the residuals as outputs/inputs in an efficiency analysis using some frontier technique. The environmental variables must then be chosen ex ante.

[7] A variant of the third step is to immediately use the residuals from the second step regression as corrected efficiency scores (Ray, 1991, and Ruggiero, 1998).

set on different ground. They can either be market rents or average cost rents. Another problem encountered in analysing labour costs is that focusing on one specific year only may lead to wrong conclusions due to for example high sick leaves for that year.

Since no centrally collected data exists in Sweden, costs for each fire service have been aggregated in a simplified way. Total labour costs for one year are equal to the number of full-time and part-time firemen in stand-by times 8760 hours (=one year of labour) weighted by an average cost figure of SEK 156 for full-time firemen and SEK 27 for part-time firemen.[8] The only capital variable available is the number of full- and part-time stations. However, this variable has not been used mainly because it is neither a continuous variable nor a homogenous variable. Furthermore labour costs account for about 70% of total costs, and this variable is probably highly correlated with labour cost.[9] Even if it is possible to collect data on the truck fleet for each fire service, different trucks from different years are difficult to compare as long as there is no good way of calculating costs.

Output has two dimensions: 1) the turn-out time, measured by the number of people reached within x minutes, and 2) the suppressing power, measured by the total number of firemen reaching the fire within x minutes. In the data on population reached, there is a choice between which minute intervals to be used. Data existed for each minute from 1 up to 30. Two specific periods, five and ten minutes, have been chosen.

Data on the number of firemen reaching the fire within five and ten minutes does not exist. Instead, the total number of firemen at stand-by in a municipality has been used. Many nearby fire services have agreements of covering up if help is needed, but this factor has not been taken into account. It may seem strange to include the number of men as output, because normally it is an input measure. In this case, there are two distinct types of firemen, part-time and full-time. Their costs are different (reflected in the input variable) since part-time firemen take about five minutes longer to turn-out

[8] Adapted from Svenska Kommunförbundet (1994), but adjusted from SEK 25 to SEK 27 correcting for a change in relative wages.

[9] Testing the inclusion of this variable in the proceeding measurings of efficiency also gave rise to another problem; about one third of the fire services were found efficient.

(reflected in the coverage variables). At the scene of the fire the difference between them is small, which is reflected in the total number of firemen variable.

The number of people reached depends on the structure of the municipality. A fire service in a municipality with a higher population density can reach more people within five minutes than one in a municipality with a smaller population density. The same is true for a fire service in a municipality with a higher population and for municipalities with smaller areas. These three variables will be tested for influencing the efficiency scores. Another interesting variable is the number of population centres, but it is not certain how this variable influences the efficiency scores. More population centres in a municipality call for more fire stations, which implies that the firemen can reach more people faster. However, more fire stations imply an increase in the number of firemen, which in turn leads to higher labour costs.

The data on the number of people a fire service can reach with five firemen has been received from the Swedish Defence Research Institute (FOA) which used simulation techniques to measure this variable (Sträng, 1999). It is broken down on each municipality in Sweden. Data on the number of firemen is received from the Swedish Rescue Services Agency (Räddningsverket), and is broken down on each fire service in Sweden. The environmental variables have been collected from Statistics Sweden (SCB). There are 289 municipalities in Sweden, but only 240 fire services, since some municipalities have chosen to co-operate using a single fire service management.

A population centre is officially defined as having a population of at least 200 and at most 200 metres between the houses. Since a population of 200 is probably too low a number to have a fire brigade, a population of 1000 has been chosen instead. The descriptive statistics of the variables are presented in Table 1.

Table 1. Mean, standard deviation, maximum, and minimum of variables.

n=238	Code	Mean	Standard deviation	Maximum	Minimum
Outputs					
Population reached within 5 minutes	V	8065.2	26577.3	340673	0
Population reached within 10 minutes	X	28257.1	64220.0	714379	0
Total number of firemen	MT	17.0	13.0	93	5
Input					
Total cost of firemen, SEK	MC	974.6	1352.6	12186	135
Environmental variables					
Population	POP	36660.7	68114.6	714449	2856
Area, km^2	ARE	1712.1	2715.6	19446	19
Population per km^2	PD	96.6	347.6	3820.6	0.28
Number of population centres with at least a population of 1000	PC	2.97	2.35	13	1

5 Empirical results

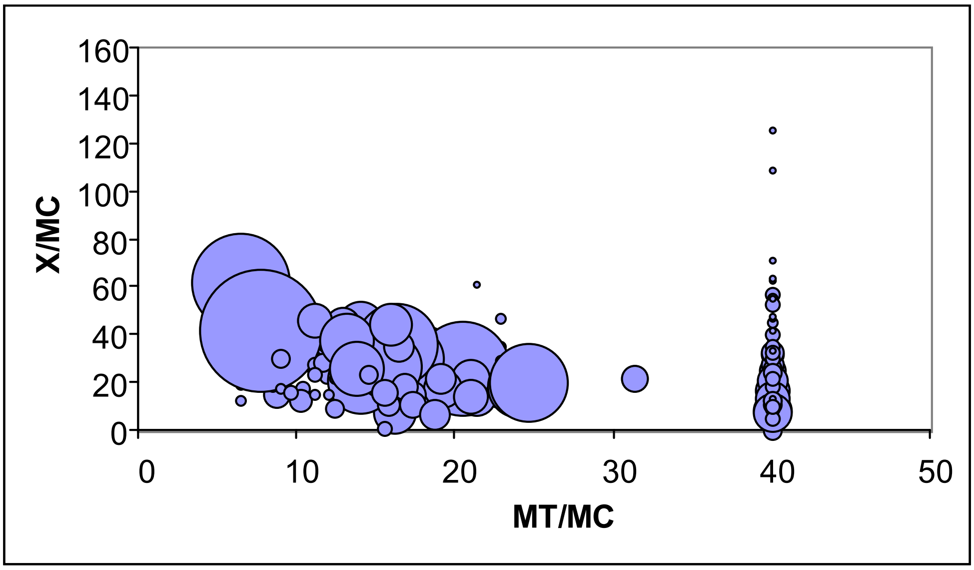
The analysis has been done in three stages. First, efficiency scores have been calculated using DEA without environmental variables; second, the resulting efficiency scores have been used to check for the relevant environmental variables; third, efficiency scores have been calculated again, but this time including the relevant environmental variables.

5.1 First stage: DEA without environmental variables

The main model used here has three output variables 1) number of people reached within five minutes, 2) number of people reached within ten minutes, and 3) total number of firemen as a proxy for suppressing power. Only input saving efficiency scores have been calculated, and a variable returns to scale (VRS) technology has been assumed.[10]

One problem with the model is that all fire services using part-time firemen only will have an efficiency score of 1. This can be seen in Figure 4, where the output coefficients

Figure 4. Plot of two outputs: the number of firemen (MT) vs. population reached within 10 minutes (X). Both are divided by input.

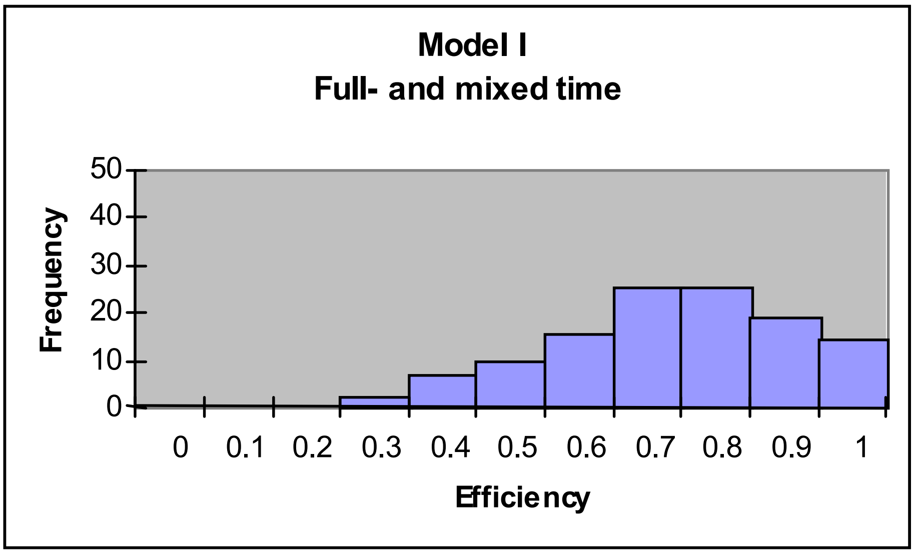


(output/input) for the number of people reached within ten minutes is given on the vertical axis, and the total number of firemen on the horizontal axis. The size of the circles is proportional to the total number of firemen. All part-time fire services lies on a vertical line at approximately 40 since the cost per fireman is equal, and thus they would all be classified as efficient. The reason is that when calculating the efficiency scores with DEA, slacks are not taken into account. The efficiency scores using the three-output model are therefore calculated only for the full- and mixed time fire services, but using all fire services in the model (Model I). The reason for including part-time fire services as reference units is that these are possible (and often better) alternatives. To make it possible to compare all fire services, a two-output model using people reached within five and ten minutes only, as output has also been calculated (Model II).

The distribution of the efficiency scores is shown in the histograms in Figure 5. The shape of the two models is very different. Model I has a normal shape skewed to the left with a peak around 0.7, while Model II has a completely different shape which peaks around 0.4.

[10] Because of zeros in the output variables there were some unconstrained solutions. To avoid this problem one has been added to all observations. This transformation should not influence the ranking of the units by efficiency scores according to Ali and Seiford (1990).

Figure 5. Efficiency distributions.



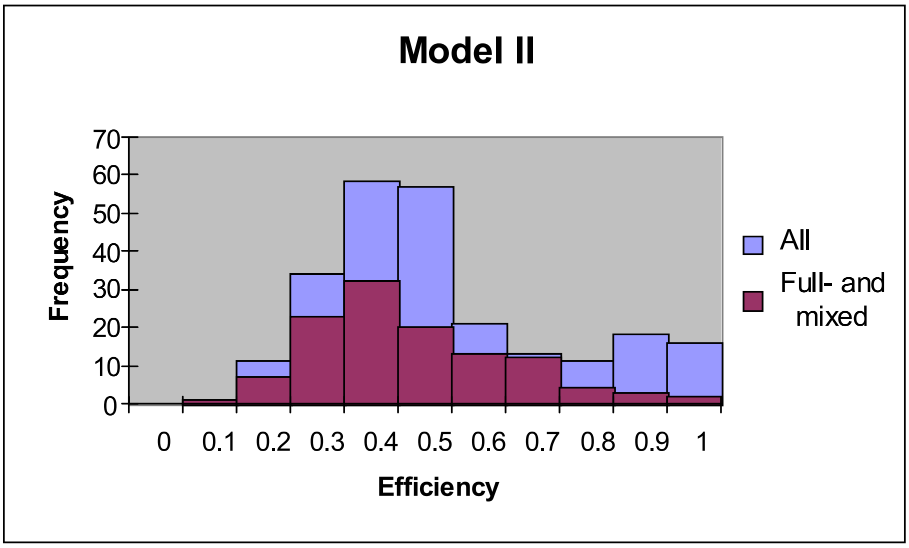


Table 2. Results from the DEA-model without environmental variables.

	Model with full- and mixed	Model with all
Model	**I**	**II**
Units	117 [a]	240
Input	MC	MC
Outputs	V, X, MT	V, X
Minimum	0.288	0.073
Median	0.700	0.428
Mean	0.689	0.484
Number of efficient units	6	16
Efficiency scores for		
Full-time (8)	0.755	0.751
Mixed (109)	0.684	0.396
Part-time (123)	-	0.544
Day/night mix (12)	0.677	0.394
Co-operation (24)	0.809	0.465
Increasing returns to scale [b]	5	189
Constant returns to scale [b]	28	21
Decreasing returns to scale [b]	84	30

[a] Efficiency scores, and returns to scale calculated with respect to all 240 units.
[b] If the scale efficiency is higher than 0.95, returns to scale is classified as constant.

Table 2 lists the descriptive and order statistics for the efficiency scores of each model. As can be seen in the table, mean efficiency is 0.69 for the full- and mixed fire services in Model I, indicating an input saving potential of 31%. In Model II mean efficiency is lower, 0.48. Full-time fire services have higher efficiency scores than mixed and part-time fire services, about 0.75. Considering the response time outcome mean efficiency is 0.54 for part-time fire services in Model II, indicating an input saving potential of 46%. Fire services with full-time firemen during daytime and part-time firemen during the night have higher efficiency scores, and so do fire services covering several municipalities. For Model II, there is no increased efficiency for these fire services, probably due to the fact that the second dimension, number of total firemen, is not included.

The last rows in Table 2 shows the classification of returns to scale. If scale efficiency is greater than 0.95, then the fire service can been classified as having constant returns to scale. One reason for the producing of fire service in Sweden as a publicly owned good is attributed to economies of scale. However, there are quite many fire services operating under decreasing returns to scale. Constant and decreasing returns to scale dominate in Model I, while increasing returns to scale dominate in Model II. However, since returns to scale is about what happens to *all* variables, Model II is not very relevant. By including the second dimension of output, the total number of firemen, the hypothesis about increasing returns to scale for fire service must be rejected.

It should also be noted that Model I gives different rankings for mixed-time fire services than Model II, but not for full-time fire services. Comparing the two models, the Spearman ranking coefficient is 1.0 for the full-time fire services, but only 0.497 for the mixed-time fire services,

5.2 Second stage: Search for environmental variables

To find out the type of environmental variables that should be used to correct the efficiency scores, an ex post regression analysis has been done using the efficiency scores obtained in the first stage as dependent variables. Since the efficiency scores must lie between 0 and 1, a Tobit-model with censoring on both tails has been used.[11] The marginal effects of the variables for the two models are given in Table 3. A star indicates that the parameter is statistically significant at the 5% level.

Table 3. Marginal effects of the Tobit regressions.

	Model with full- and mixed	Model with all
	I	**II**
Units	117	240
Input	MC	MC
Outputs	V, X, MT	V, X
Constant	0.543*	0.567*
Population	0.0000119*	0.0000122*
Area	0.0000168	-0.00160*
Number of population centres with a population of at least 1000	0.0127	-0.0400*
Population per square km	0.000671	0.00208*

* Statistically significant at the 5% level.

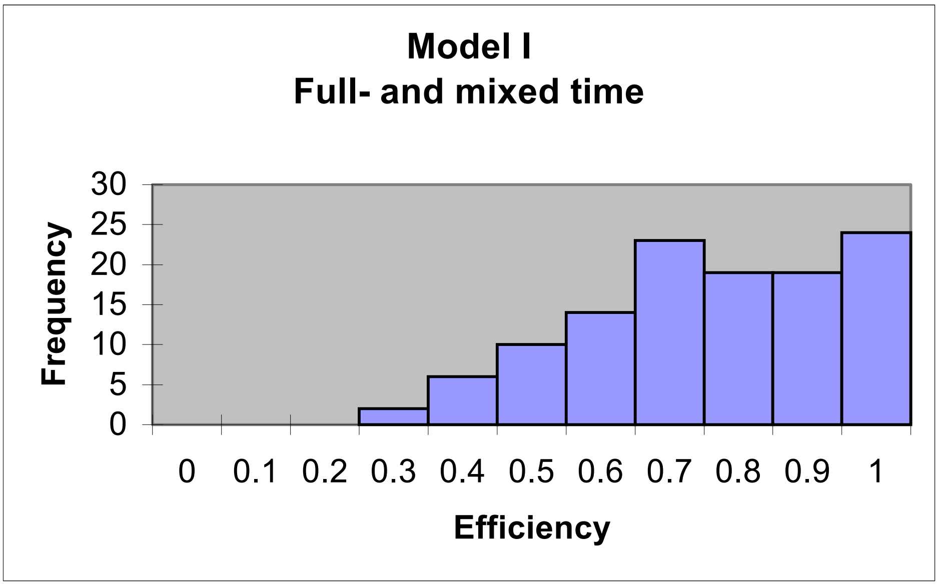
[11] Lundvall (1999) discussed the appropriateness of using the Tobit model, since it assumes that there exists a latent dependent variable that can take values greater than one. He compared the Tobit model with OLS, and a logit model, but found no major differences.

Population is statistically significant for both models, while all environmental variables are significant for Model II. All statistically significant variables have the hypothesised sign, that is, a greater population, a smaller area, and a higher population density. It also turns out that less population centres increase efficiency. The number of population centres thus influences output more than input.

5.3 Third stage: DEA including environmental variables

Following Banker and Morey (1986), the efficiency scores for the two models will again be calculated, by including the relevant environmental variables as fixed input variables. The distributions of the efficiency scores are presented in the histograms in Figure 6. The distributions are similar to those without environmental variables, the exceptions being that the columns have shifted to the right, and that there are more fire services that are efficient. A summary of the efficiency scores is presented in Table 5.

Figure 6. Efficiency distributions.



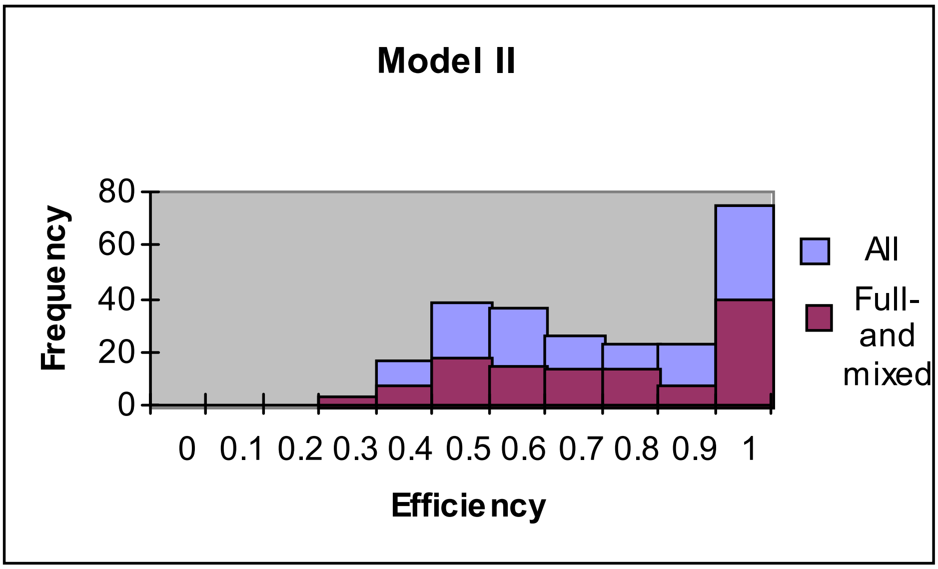


Table 4. Results from the DEA-model including environmental variables.

	Model with full- and mixed	Model with all
	I	II
Units	117 [a]	240
Input	MC	MC
Outputs	V, X, MT	V, X
Environmental variables		
fixed inputs	POP	POP, PD
fixed outputs		ARE, PC
Minimum	0.288	0.271
Median	0.719	0.704
Mean	0.716	0.713
Number of efficient units	16	65
Ranking correlation with DEA without environmental variables	0.876	0.532
Ranking correlation between the models for full- and mixed time	0.526	
Very large (3)	0.981	1.000
Suburban (22)	0.693	0.827
Large (25)	0.814	0.875
Medium (35)	0.661	0.697
Industrial (46)	0.664	0.689
Rural (31)	0.670	0.623
Thinly populated (29)	0.491	0.746
Others (49)	0.723	0.633
Full-time (8)	0.826	0.928
Mixed(109)	0.709	0.706
Part-time (123)	-	0.705
Day/night mix (12)	0.738	0.720
Co-operation (24)	0.835	0.822
IRS[b]	1	89
CRS[b]	69	112
DRS[b]	46	39
Mean efficiency assuming CRS	0.628	0.626

[a] Efficiency scores, and returns to scale calculated with respect to all 240 units.
[b] If the scale efficiency is higher than 0.95, returns to scale is classified as constant.

As expected, the efficiency scores and the number of efficient fire services in both models increase by including environmental variables. The input saving potential is 28% for the full- and mixed fire services, and the number of efficient units are now 16 compared to 6 before. For Model II the difference is greater. Mean efficiency has increased from 0.48 to 0.71, and the number of efficient units from 16 to 65. Compared to the models without environmental variables; the Spearman ranking correlations are 0.86 for Model I and 0.53 for Model II. The inclusion of the environmental variables thus affects Model II more than Model I.

In order to carry out further investigations, fire services that are more or less efficient have been divided according to the classification of municipalities in Sweden (according to Statistics Sweden): very large, suburban, large, medium, industrial, rural, thinly populated, and others. Fire services covering more than one municipality have been classified according to the municipality with the highest population.

Despite the fact that population is included as a fixed input it can be seen that very large and large municipalities have the highest efficiency scores in both models. Thinly populated municipalities have the lowest efficiency scores in Model I. Rural municipalities have, together with others, the lowest efficiency scores in Model II. Full-time fire services have higher efficiency scores than mixed fire services in both Model I and II.

Municipalities that co-operate by sharing a fire service are more efficient than the average fire service. Fire services that use different kind of firemen during day and night have, surprisingly, almost the same efficiency as the average fire service. By using only part-time firemen at night the labour cost savings do not seem to turn up as efficiency gains. Note that these results are different compared to the first stage and they are thus affected by the inclusion of the environmental variables.

Compared to the models without environmental variables the efficiency scores for full-time and mixed time fire services in Model I are almost the same for all categories. For

Model II, the efficiency scores mostly increased for large, medium, thinly populated and suburban municipalities.

As can be seen in the table the number of fire services for each class of returns to scale has changed quite a lot from the models without environmental factors. There are now more fire services classified as having constant returns to scale. However, it is not clear whether the production frontier has changed.

Let us have a closer look at the most and least efficient fire services in order to see if it is possible to identify common attributes. Table 5 lists the 16 efficient fire services and the 10 least efficient fire services in Model I.

Table 5. Sixteen most efficient and ten least efficient fire services in Model I.

Unit	Eff.	No. of times peer	Returns to scale	Cost	Population	Area	Category	Ten-minute variable	Full-time fire-men	Part-time fire-men
101	1	2	CRS	1872	86367	29	Suburban	86367	12	0
180	1	1	CRS	11544	714449	187	Very large	714379	74	0
186	1	1	IRS	936	39337	30	Suburban	39236	6	0
191	1	1	CRS	2190	33516	327	Suburban	31963	13	6
382	1	1	DRS	1925	21328	1451	Others	14642	8.36	23
662	1	1	DRS	1401	29689	1141	Industrial	24462	5	23
682	1	1	DRS	1401	30114	934	Medium	24631	5	23
687	1	1	CRS	834	17808	404	Others	15425	5	2
883	1	1	DRS	1590	39066	1870	Medium	29609	5	30
1383	1	1	DRS	1746	51992	873	Medium	43681	6	30
1411	1	1	DRS	12186	565104	1206	Very large	508893	75	18
1413	1	1	DRS	3594	149382	3897	Large	90722	13	58
1495	1	15	CRS	495	18571	439	Others	15373	1.79	8
1701	1	1	DRS	3398	95230	4540	Industrial	77138	16.07	33
2284	1	1	DRS	2043	57457	6418	Large	36420	6	41
2482	1	6	DRS	2496	74471	6838	Large	48066	7	52
Median				**1898**	**45664.5**	**1037.5**		**37828**	**6.5**	**23**
1984	0.288	-	CRS	724	14310	324	Industrial	11446	4	3.71
1982	0.295	-	CRS	732	13176	270	Industrial	11210	4	4
2062	0.328	-	CRS	576	20687	2827	Medium	12243	3	4
2581	0.337	-	CRS	1152	40783	3086	Medium	10717	6	8
1282	0.345	-	DRS	1404	37126	142	Medium	35922	9	0
1284	0.358	-	CRS	603	22845	144	Medium	12255	3	5
120	0.385	-	CRS	861	26190	439	Suburban	17724	5	3
882	0.385	-	CRS	1776	27153	1047	Medium	22187	10	8
1261	0.403	-	CRS	915	24106	153	Suburban	23009	5	5
1407	0.414	-	CRS	535	11165	26	Suburban	10046	2.429	5.786
Median				**796.5**	**23475.5**	**297**		**12249**	**4.5**	**4.5**
Median All 117				**1293**	**36214**	**1047**		**26604**	**5**	**13**

The 16 efficient fire services look like a good sample. Both large and small municipalities are included in relation to variables such as cost, population and area. The efficient fire services also belong to many different municipality categories. These results are expected since the production frontier should envelop all fire services. However, the efficient fire services have proportionally more part-time firemen than the rest.

The ten least efficient fire services have low populations, but cover small areas. They belong to medium, industrial, and, more surprisingly, to suburban municipalities. A possible way of increasing the efficiency of the suburban fire services would be to co-operate with neighbouring fire services. It can also be seen that the least efficient fire services employ a higher degree of more costly full-time firemen. Therefore, another way of increasing efficiency is to substitute from full-time firemen to part-time firemen. These fire services have also a low coverage after 10 minutes.

Furthermore efficiency can be improved through studying the units that are peers for the inefficient units most times. Table 6 below shows the fire services that are most of the times peers for both models. The peer weight (λ_{Fj} in equation 1) can be either very low, or very high, but since this weight also depends on the sizes of the variables any peer weight is considered.

For Model I it is interesting to see that only four, out of the sixteen efficient units, are among the most popular peers (nos. 101, 180, 1495 and 2482). The reason why all peers are not among the most efficient fire services is that in Table 5 only full- and mixed time fire services are presented, but part-time fire services could also be among the peers. The median values for the peers are higher than the median values for all units. The fire service in the largest municipality is one of the peers, but there are also five peers that are part-time fire services. There are also three suburban fire services among these peers. Thus the most popular peers are either very large or small, and therefore the hypothetical peer (target) on the frontier is most often a weighted average of large and small fire services. This is a pedagogical problem since medium-sized fire services do not have real peers to compare themselves to. However, the conclusion must again be that changing the manning from full- to part-time is the easiest way to increase efficiency.

In Model II we first notice that the peers are very different from Model I, with the exception of fire service no. 1230. The median of the most popular peers in Model II are also smaller than those of the peers for Model I, which is not surprising since the peers for all 240 fire services are presented. There are three fire services with the minimum cost of 135. There are also two large fire services included among the most popular peers. It is also interesting to note that almost all of the most popular peers operate under constant returns to scale.

Table 6. Most popular peers for Model I and II.

Unit	Model I, no. of times peer	Model II, no. of times peer	Returns to scale	Cost	Popul-ation	Area	Category	Ten-minute variable	Full-time fireme n	Part-time fire-men
180	76		CRS	11544	714449	187	Very large	714379	74	0
1495	39		CRS	495	18571	439	Others	15373	1.786	8
2161	33		CRS	648	20500	579	Others	12412	0	24
1230	29		CRS	162	19145	108	Suburban	18811	0	6
801	26		CRS	702	25341	1342	Rural	11143	0	26
1440	23		CRS	324	25281	318	Suburban	17005	0	12
2482	22		DRS	2496	74471	6838	Large	48066	7	52
2309	20		CRS	756	14672	6211	Thinly populated	5898	0	28
2361	17		CRS	756	11957	11405	Thinly populated	5225	0	28
101	15		CRS	1872	86367	29	Suburban	86367	12	0
Median Model I				**729**	**22890.5**	**509**		**16189**	**0**	**18**
2061		78	CRS	162	12519	953	Industrial	8251	0	6
2506		58	CRS	189	3573	12945	Thinly populated	2135	0	7
2101		51	CRS	3171	146428	4250	Large	117321	16	25
834		49	CRS	135	7733	469	Rural	2692	0	5
1962		45	CRS	135	6393	422	Industrial	5214	0	5
581		42	CRS	2697	123176	1491	Large	106797	14	19
2513		40	CRS	135	4479	2790	Thinly populated	1724	0	5
1472		37	CRS	135	11023	220	Industrial	8881	0	5
1256		29	CRS	270	14841	434	Industrial	5508	0	10
1230		25	CRS	162	19145	108	Suburban	18811	0	6
Median Model II				**162**	**11771**	**711**		**6879.5**	**0**	**6**
Median all 240 units				**459**	**16054**	**827**		**10122**	**0**	**11**

Table 7. Relation between returns to scale and size.

No. of units	Q1	Q2	Q3	Q4
Model I				
IRS	1	0	0	0
CRS	2	1	1	1
DRS	2	2	3	3
Model II				
IRS	11	3	0	0
CRS	6	10	11	10
DRS	0	3	5	6

Q1=25% with least cost
Q4=25% with highest cost

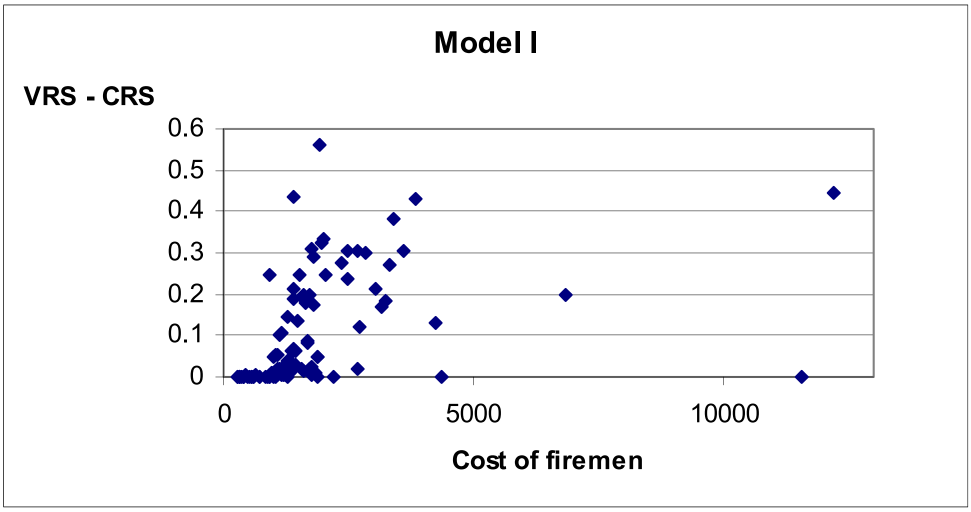
Can optimal scale (CRS) be related to a specific size for the fire service considering the stand-by level? The classification of returns to scale for the frontier units in the two models are listed, with their sizes (measured as cost), divided into four quartiles in Table 7. In both models it is clear that shifting from small to medium and then to large units changes the returns to scale from increasing to constant and to decreasing returns to scale. However, the constant returns to scale units include small, medium and large units, and therefore no optimal size can be found.[12]

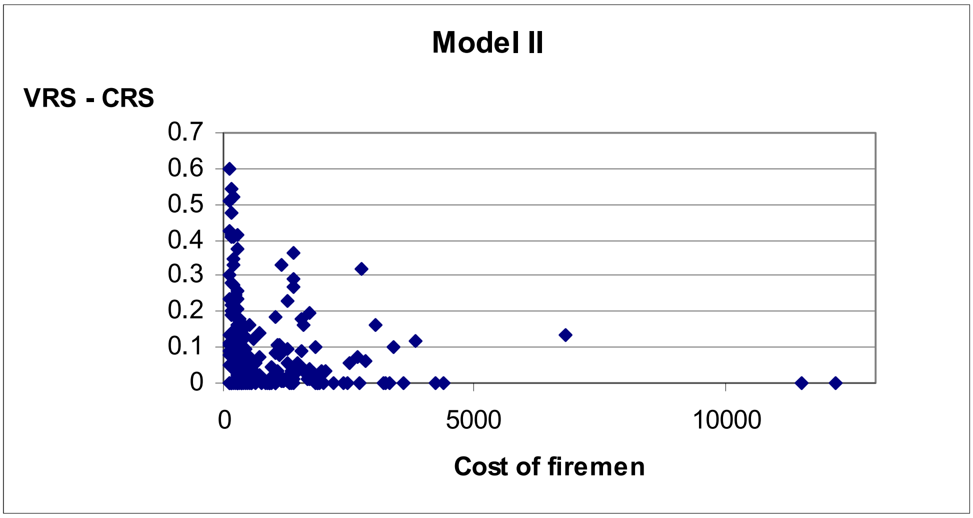
Changing the technological assumption from VRS to CRS could result in much smaller efficiency scores for very large and very small of the most efficient fire services. Thus, these fire services may be situated on the horizontal and vertical part of the production frontier in Figure 2. It can therefore be interesting to see how the efficiency measures will differ if another assumption about the technology is made. Figure 7 plots the difference between the efficiency scores assuming a CRS-technology instead of a VRS-technology and the size of the fire services.[13] The patterns are different in the two models. In Model I the difference is smallest for small fire services, but there seems to be a positive correlation between the difference and the size. In Model II, the differences are instead largest for small fire services. The difference probably depends

[12] The problem of measuring optimal scale in DEA has been discussed by Førsund and Hjalmarsson (1996 and 2002).
[13] Notice that this number is close to 1-scale efficiency.

on the fact that most small fire services are part-time fire services, and are therefore not included in Model I.

Figure 7. Difference between VRS and CRS assumption vs. size





6 Summary and conclusions

Empirical results in the long-run model that uses DEA to analyse the efficiency for the stand-by level among the Swedish fire services show that:

- Taking environmental aspects into account the input saving potential is about 30% for fire services in the long run.
- Fire services in very large and large municipalities have higher efficiency scores, while those in thinly populated municipalities have lower.
- Full-time fire services have higher efficiency scores than fire services using both full- and part-time firemen.
- Municipalities that co-operate are more efficient, but fire services with a full-time crew during the day and part-time crew during the night are not more efficient.
- When comparing returns to scale for the frontier units to their sizes, no optimal size of production could be found,.
- Since fire services are publicly run in Sweden, increasing returns to scale for most units were expected, but not found.

Despite the fact that input cost was calculated in a simple way, and that both dimensions of output for part-time fire services were not possible to include, this study can be interesting to fire service managers because:

- It clarifies the difference between just studying output or input measures, and studying productivity measures. It makes more sense to study differences in productivity than just differences in either output produced or inputs used, because output depends on input.
- It gives the managers a hint of what they get for their resources. They produce two dimensions of output for the stand-by level; the turn-out time (the faster the better), and the suppressing power (the more firemen the better). Both outputs should be included when comparing productivity with other fire services.
- The efficiency scores make it possible for them to find out what other fire services that are better, and how much better they are, a sort of benchmarking.

- The returns to scale classification for the efficient fire services will indicate whether the size of the fire service is appropriate.

In the future there may be better statistics about costs of the production factors for the Swedish fire services. Future research should then concentrate on trying to include more input variables, such as capital and administrative inputs.

References

Ahlbrandt, R. C., 1973, Efficiency in the provision of fire services, *Public Choice*, 16, 1-15.

Ali, A. I., and L. Seiford, 1990, Translation invariance in data envelopment analysis, *Operations Research Papers*, 9, 403-405.

Banker, R. D., and R. C. Morey, 1986, Efficiency analysis for exogenously fixed inputs and outputs, *Operations Research*, 34, 513-521.

Banker, R. D, and R. M. Thrall, 1992, Estimation of returns to scale using data envelopment analysis, *European Journal of Operational Research*, 62, 74-84.

Bouckaert, G., 1992, Productivity analysis in the public sector: the case of the fire service, *International Review of Administrative Sciences*, 58, 175-200.

Bradford, D., R. Malt and W. Oates, 1969, The rising cost of local public services, *National Tax Journal*, 22, 185-202.

Charnes A., W. Cooper, A. Lewin, and L. Seiford, 1994, *Data envelopment analysis: Theory, methodology and applications*, Kluwer Academic Publishers, Boston.

Charnes, A., W. W. Cooper and E. Rhodes, 1978, Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2, 429-444.

Coelli, T., 1996, A guide to DEAP version 2.1: A data envelopment analysis (computer) program, *CEPA working paper 96/08*, Department of Econometrics, University of New England, Armidale.

Coelli, T., D. S. P. Rao and G. E. Battese, 1998, *An introduction to efficiency and productivity analysis*, Kluwer Academic Publishers, Dordrecht.

Duncombe, W. and J. Yinger, 1993, An analysis of returns to scale in public production, with an application to fire protection, *Journal of Public Economics*, 52, 49-72.

Färe R., S Grosskopf and C.A.K. Lovell, 1994, *Production Frontiers*, Cambridge.

Farrell, M. J., 1957,The measurement of productive efficiency, *Journal of the Royal Statistical Society*, ser. A, 120.III, 211-281.

Fried, H. O., C. A. K. Lovell and S. S. Schmidt (eds.), 1993, *The Measurement of Productive Efficiency*, Oxford University Press.

Førsund, F. and L. Hjalmarsson, 1974, On the measurement of productive efficiency, *Swedish Journal of Economics,* 76(2), 141-154

Førsund, F. and L. Hjalmarsson, 1979, Generalised Farrell measures of efficiency: an application to milk processing in Swedish dairy plants, *Economic Journal*, 89, 294-315.

Førsund, F. and L. Hjalmarsson, 1996, Measuring returns to scale of piecewise linear multiple output technologies: Theory and empirical evidence*, Memorandum 223*, Department of Economics, Göteborg University.

Førsund, F. and L. Hjalmarsson, 2002, Are all scales optimal in DEA? Theory and empirical evidence*, Working Paper no. 66,* Department of Economics, Göteborg University.

Grosskopf, S., K. Hayes, L. Taylor, and W. Weber, 1997, Budget-constrained frontier measures of fiscal equality and efficiency in schooling, *Review of Economics and Statistics*, 79(1), 117-124

Jaldell, H., 2002a, Swedish fire and rescue services' manning levels - a stochastic frontier analysis using panel data, *Paper 2 in this thesis*.

Jaldell, H., 2002b, Measuring performance differences using an ordinal output variable: The case of Swedish fire services, *Paper 5 in this thesis*.

Kristensen, Ole, 1983, Public versus private provision of governmental services: The case of Danish fire protection services, *Urban Studies*, 20, 1-9.

Lovell, C. A. K., 1993, Production frontiers and productive efficiency, ch. 1 in Fried, Lovell and Schmidt (1993).

Lundvall, K., 1999, *Essays on manufacturing production in a developing economy: Kenya 1992-94*, Ekonomiska studier 96, Göteborg University.

Ray, S. C., 1991, Resource use efficiency in public schools. A study of Connecticut data, *Management Science*, 37, 1620-1628.

Ruggiero J., 1996, On the measurement of technical efficiency in the public sector, *European Journal of Operational Research*, 553-565.

Ruggiero J., 1998, Non-discretionary inputs in data envelopment analysis*, European Journal of Operational Research*, 11, 461-469.

Sträng, D., 1999*, Insatstid för räddningstjänst vid skydd mot olyckor*, SRV/FOA, P21-317/99.

Svenska Kommunförbundet, 1994, *Räddningstänstens personalkostnader* (Swedish Association of Local Authorities*, Labour costs of the fire service*), Stockholm.

4

Paper 4

Productivity change of Swedish fire services between 1992 and 1998

by
Henrik Jaldell

ABSTRACT

In this paper, Malmquist productivity indexes are used to find out how productivity has changed among Swedish fire services between 1992 and 1998. The paper studies the stand-by level which has two output dimensions 1) the turn-out time (the faster the better), measured as the number of people reached within five and ten minutes, and 2) the suppressing power, measured by the total number of firemen turning out (the more the better). The empirical results show that productivity has decreased for full- and mixed-time fire services. Less input used has resulted in less output produced.

Keywords: fire service, public sector, Malmquist, productivity, stand-by

1 Introduction

Swedish municipalities have faced tremendous fiscal stress since the recession in 1992. Between 1993-1996 aggregated net account was about SEK -30 billion for the municipality sector, which resulted in budget cuts for most of the local public services. This has in turn resulted in less input employed by the municipality services, which may have resulted in less output produced. However, the fiscal stress may have given rise to productivity increases, and in that case output may instead have been stable or perhaps increased.

The purpose of this paper is to analyse how the fiscal stress has affected one of the rarely studied local public services: The fire service. Have the budget cuts led to less input used, and has this resulted in less output produced? In other words: Has productivity changed? No other studies of productivity change over time for fire services have been found.

One problem with studying the public sector is to choose an appropriate output measure. Bradford et al (1969) emphasised that it was important to distinguish between directly produced intermediate outputs and final output. Extending their model, this paper shows how fire suppression could be seen as a three-step process. The first step is the stand-by level, where suppressing power (measured as the number of firemen on duty) and the response time (which is measured by the number of people reached within x minutes) are the first intermediate outputs. The outcome from the turn-outs where saved lives and values of fires are the second intermediate outputs, and the welfare resulting from all activities of the fire service is the final output.

In this paper, the change in the relation between the first intermediate outputs and inputs from 1992 to 1998, is studied. Productivity is measured by using input based Malmqvist indexes that are calculated by using a non-parametric linear programming technique, the so-called data envelopment analysis (DEA).[1] The advantages of the Malmquist

[1] The DEA method was first proposed by Charnes, Cooper and Rhodes (1978) and is thoroughly described in Charnes, Cooper, Lewin and Seiford (1994), and connected to economic production theory in Färe, Grosskopf and Lovell (1994).

productivity index over other indexes, such as Fisher and Törnqvist indexes, are that prices are not needed and no behavioural assumptions such as cost minimisation have to be made. Another advantage of the Malmquist productivity index is that it makes it possible to decompose total productivity change into two parts 1) change in efficiency, which may be called the catching-up-effect, and 2) change of the production frontier, known as technical change. Caves, Christensen and Diewert (1992) first proposed using Malmqvist indexes for productivity measurement, and Färe, Grosskopf, Lindgren and Roos (1994) modified the index so that inefficiency could be taken into account. The advantage over parametric and econometric estimations is that no assumptions about functional forms and about statistical noise have to be made. The disadvantage is that statistical tests are hard to perform.

The Malmquist productivity indexes are introduced in section 2. The input and output variables are described in section 3, and the empirical results can be found in section 4. Section 5 summarises and concludes.

2 The model

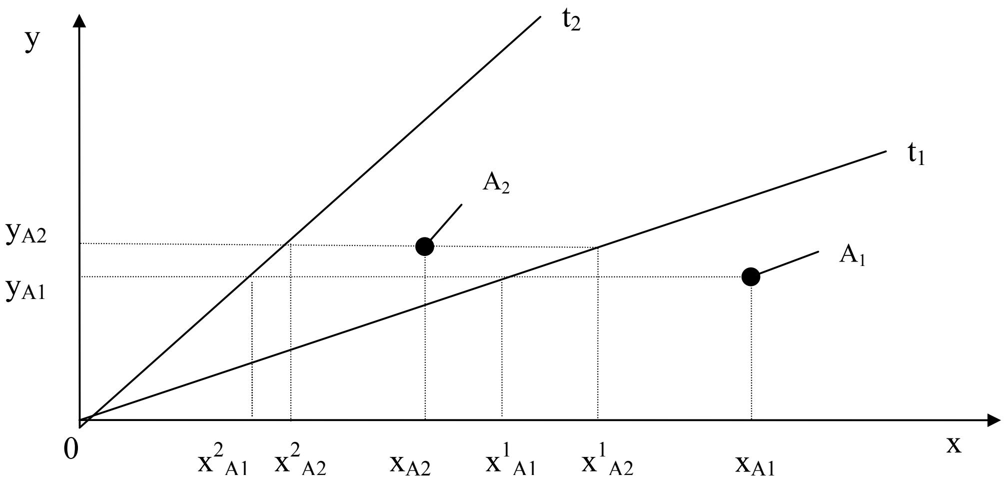
The Farrell (1957) input saving efficiency measure for a unit *i* observed in period *s* is

$$E^i_{st}=min_\theta\{\theta:(y^i_t,\ \theta x^i_t)\in P^s\} \qquad (1)$$

where y^i_t is the output vector, x^i_t the input vector, and P^s is the production possibility set in period *s*. If the efficiency score, $E^i_{st,}$ is less than one, the unit is inefficient compared to the technology in period *s*, while a value greater than one means that the technology in period *s* is not feasible in period *s*.

Comparing efficiency scores to different technologies over time, the Malmquist productivity indexes can be calculated.[2] In Figure 1, where we have two production technologies, t_1 and t_2 for two time periods, the structure of the index can be seen. Production unit *A* has decreased its input from x_{A1} to x_{A2}, and increased its output from y_{A1} to y_{A2}.

Figure 1. Decomposition of productivity change with input based Malmquist productivity indexes.



Given the technology frontier in period 1 the Farrell input saving efficiency for unit A in period 1 is calculated as $E_{11}=x^1_{A1}/x_{A1}$. Given the observation in period 2 the efficiency score is $E_{12}=x^1_{A2}/x_{A2}$. The input-oriented Malmquist productivity index given the technology in period 1 is then

$$M_1 = \frac{x^1_{A2}/x_{A2}}{x^1_{A1}/x_{A1}} = \frac{E_{12}}{E_{11}}. \qquad (2)$$

Alternatively, one may be interested in comparisons to the technology in period 2. The input-oriented Malmquist productivity index given the technology in period 2 for unit *A* is then

[2] Färe, Grosskopf, Lindgren and Roos (1994) used distance functions to define the Malmqvist productivity index, but Farrell efficiency measures can also be used.

$$M_2 = \frac{x_{A2}^2 / x_{A2}}{x_{A1}^2 / x_{A1}} = \frac{E_{22}}{E_{21}}. \qquad (3)$$

If M_1 and M_2 are greater than one, productivity has increased. If they are less than one, productivity has decreased, and if they are equal to one, productivity has not changed. Since these two measures do not normally give the same result, it is common practice, in the spirit of the Fisher index, to use the geometric mean of the two indexes $M = \sqrt{M_1 M_2}$. The index can now be decomposed into one catching-up component, *CU*, and one technology change component, *TC*:

$$M = \sqrt{M_1 M_2} = \sqrt{\frac{E_{12}}{E_{11}} \frac{E_{22}}{E_{21}}} = \frac{E_{22}}{E_{11}} \cdot \sqrt{\frac{E_{12}}{E_{22}} \frac{E_{11}}{E_{21}}} = CU \cdot TC \qquad (4)$$

The Farrell efficiency scores are calculated by means of the data envelopment analysis (see Appendix). The reason for using input saving measures only is that in the fire service, there is a limit to the output increasing potential; the number of population reached within x minutes cannot exceed the population of the municipality.

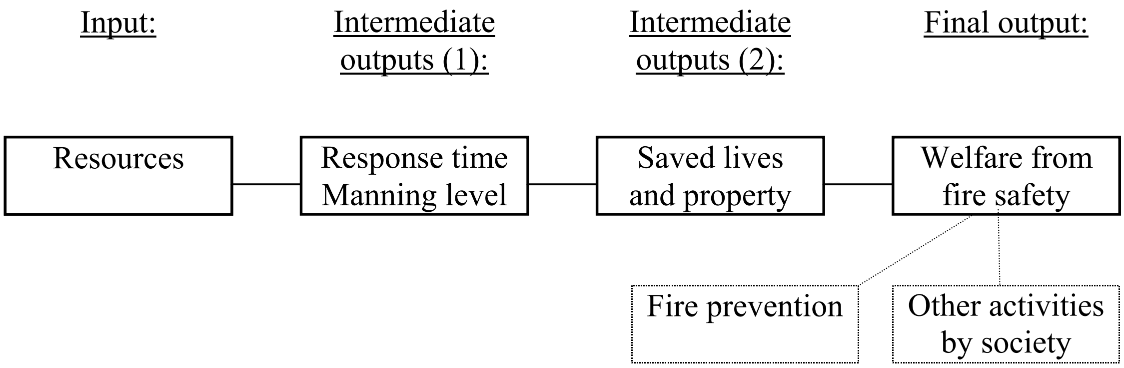
The catching-up component could be further decomposed into a pure efficiency change and a scale efficiency change. However, the output and input indexes are not well defined for a variable returns to scale technology (Grifell-Tatjé and Lovell, 1995, Bjurek, 1996, and Bjurek, et al, 1998), so scale efficiency change is not calculated here. Former studies, using the same data as that used here, have defined fire service technology as a non-constant returns to scale technology (Jaldell, 2002). However, the problem of assuming constant returns to scale is not very severe since this is a productivity analysis of changes over time, and therefore the results will not be very different depending on the scale assumption.

3 Data

This analysis studies productivity change between 1992 and 1998, but with observations for these two years only. There are 289 municipalities, but 240 fire services only, since some municipalities have a single fire service management.

The two main activities of the fire service are to prevent fires and other accidents from happening, and if they happen anyway, to suppress them as effectively as possible. This paper only considers the suppressing activity. Extending the two-step model for public sector output suggested by Bradford, Malt and Oates's (1969), fire suppression is here described as a three-step procedure with two levels of intermediate outputs and one final output, as in Figure 2 below.

Figure 2. A schematic description of the fire suppression activity.



The fire service receives resources from the municipality to prevent and suppress fires. The first intermediate outputs for fire suppression is the stand-by level, which can be measured as the average response time to get to the fires and the number of firemen initially reaching the fire. Disregarding costs, the shorter the response time the better, and the more firemen arriving at the fire the better. If more firemen are needed to control the fire, it is also important that the reinforcement does not take a very long time. The fire service must decide how many firemen are needed and what kind to use (full- or part-time), and, in the long run, how many stations to have and how to divide the firemen between the stations. For a full-time crew in Sweden it should take about 90 seconds to be on the way, while for a part-time crew it takes between 5 and 6 minutes, since the part-time crew is at work or at home. Part-time firemen usually have other full-time jobs, but do a 24-hour on-call service every third or fourth week.

The first intermediate outputs (response time and number of firemen) are then inputs in the production of the second intermediate output; how well the fire service succeeds in suppressing the fire. The final output is the welfare of the inhabitants of the municipality brought by all activities of the fire service (including fire prevention), and other factors such as tradition, environmental risk, weather and building conditions.

This paper focuses on the first intermediate output, the stand-by level, and its relation to the costs. It is reasonable to assume that there is a positive correlation between the first and second intermediate output, and between the second intermediate output and the final output. That is, the findings about productivity in this paper can be extended to the final output.

The interesting input variables are costs of labour, capital, material etc. However, these costs are not easy to homogenise among the fire services, since different municipalities use different accounting principles to handle them.

To avoid these problems, and since no centrally collected data exists in Sweden, costs have been aggregated in a simplified way for each fire service. Total labour costs for one year are equal to the number of full-time and part-time firemen on duty times 8760 hours (=one year) weighted by an average cost figure of 156 kr for full-time firemen and 25 kr for part-time firemen in 1992.[3] For 1998 the cost figures used are 156 kr for full-time firemen and 27 for part-time, which demonstrates that the figure for part-time firemen has relatively increased. The only capital variable available is the number of full- and part-time stations. However, this variable has not been used since it is not a continuous and homogeneous variable. In addition labour costs account for about 70% of total costs and this variable is highly correlated with labour cost.[4] Even if it is possible to collect data on the truck fleet for each fire service, different trucks from different years are difficult to compare as long as a good way of calculating costs is not used.

Output has two dimensions 1) the number of people reached within x minutes, and 2) the total number of firemen reaching the fire within x minutes. For the data on

[3] Svenska Kommunförbundet (1994)

[4] Testing the inclusion of this variable in the efficiency calculations also gave another problem; about one third of the fire services were found efficient, thus reducing the discriminatory power substantially.

population reached there was a choice between which minute intervals to use. Data existed for each minute from 1 up to 30, but two specific intervals only, 5 and 10 minutes, have been chosen. Data on the number of firemen that can reach the fire within 5 and 10 minutes does not exist. Instead, only the total number of firemen on duty in a municipality has been used. It may seem strange to include the number of men as output because normally, this is an input measure. However, in this case, there are two distinct types of firemen, part-time and full-time. Their costs are different (reflected in the input variable), and part-time firemen take about five minutes longer to turn-out (reflected in the coverage variables), but at the place of the fire the difference between them is small (reflected in the total number of firemen variable).

The data on how the number of people a fire can service reach with five firemen has been received from the Swedish Defence Research Institute (FOA), which used simulation techniques to measure this variable (Sträng, 1999, and Andersson, 2001). It is broken down on each municipality in Sweden. A change in population of course influences these variables. However, the data for both years has been calculated using the population in 1998. The data on the number of firemen is received from the Swedish Rescue Services Agency (Räddningsverket), and it is broken down on each fire service in Sweden. All 240 fire services in Sweden, except two, are studied. [5]

The descriptive statistics of the variables for the two years are given in Table 1. In the case of output it can be seen that people reached within 5 minutes have decreased by 2%, while those reached within 10 minutes have slightly increased. The third output, total number of firemen, has decreased by almost 9%. Looking at the composition of number of firemen, the number of full-time firemen is about the same, while that of part-time firemen has decreased.

[5] Two fire service are missing due to lack of data.

Table 1. Mean and standard deviation of the variables.

n=238		1998		1992	
	Code	Mean	Std. deviation	Mean	Std. deviation
Output					
Population reached within 5 minutes	V	8099.0	26679.0	8252.9	26801.1
Population reached within 10 minutes	X	28376.0	64478.0	28299.0	64675.8
Total number of firemen	MT	17.1	13.0	18.7	14.7
Number of part-time firemen		13.1	8.8	14.8	10.2
Number of full-time firemen		4.0	8.3	3.9	8.5
Input					
Total cost of firemen	MC	978.7	1357.1	980.3	1412.6

The cost of firemen is slightly lower for 1998 than for 1992. The change in productivity between 1992 and 1998 is thus not clear because of the divergent changes of the output variables. However, it seems as if the fire services have come out well from the budget cuts in the Swedish municipalities.

4 Empirical results

Two models have been used in calculating the indexes. Model I presents the results of the full- and mixed time fire services, but they are calculated using all fire services, including part-time fire services as reference units. This model uses both dimensions of output, number of people reached and total number of firemen. The reason why the results of the part-time fire services are not presented is because they all have an efficiency score of one due to the simplified way of constructing the input variable. To make comparisons possible for part-time fire services also, Model II includes all fire services, but uses only one dimension of output, number of people reached within 5 and 10 minutes.[6] The mean results for the models are presented in Table 2 below.[7]

In Model I with the results for the full- and mixed-time fire services only, total productivity has decreased by 8% between 1992 and 1998. The decomposition into a catching-up component and a technology change component reveals both are less than one. Thus it seems that both the efficient fire services, represented by technology change, and the less efficient ones, represented by the catching-up component, have

[6] Two units were deleted from the mean calculations in Model II because of extremely high productivity numbers, which were due to a very high percentage change in number of people reached.

become less productive. The savings in inputs used have thus resulted in even less output produced.

In Model II the total number of firemen as output has not been included and total productivity has increased by 2.5%. The decomposition shows that there has been no technical change and that the catching-up component is positive. Thus, it seems that productivity has increased. The savings in inputs used have not resulted in a reduction in the number of people reached within five and ten minutes.

For full-time fire services productivity has not changed but there has been a positive catching-up effect and a negative technical change. For mixed-time fire services the results are very close to the average. Part-time fire services have had a somewhat lower productivity increase than the average unit. Since the technical change component is only greater than one for the part-time fire services, these are the only units pushing the frontier outward.

[7] Calculations have been done using DEAP version 2.1, which is described in Coelli (1996).

Table 2. Arithmetic means from the Malmquist productivity calculations.

		Model with full- and mixed	Model with all
	Model	I	II
	Units	117[a]	236[b]
	Outputs	V, X, MT	V, X
	Input	MC	MC
All units	Catching-up part	0.958	1.029
	Technical change	0.956	0.997
	Total productivity	0.917	1.025
Full-time (8)	Catching-up part	1.016	1.016
	Technical change	0.980	0.978
	Total productivity	0.995	0.997
Mixed-time (108)	Catching-up part	0.954	1.045
	Technical change	0.954	0.983
	Total productivity	0.911	1.026
Part-time (120)	Catching-up part	-	1.016
	Technical change	-	1.010
	Total productivity	-	1.026
Co-operation (24)	Catching-up part	1.000	1.045
	Technical change	0.960	0.976
	Total productivity	0.960	1.018
Day/Night (12)	Catching-up part	1.070	1.308
	Technical change	0.948	0.989
	Total productivity	1.019	1.294
Ranking correlation, full and mixed	Catching-up part	0.711	
	Technical change	-0.578	
	Total productivity	0.677	

[a] Efficiency scores calculated with respect to all 238 units.
[b] Means calculated using 236 units. Two outliers are removed.

A hypothesis that can be made is that municipalities, which co-operate by running a single fire service, have been more productive than the others. Total productivity has decreased less for these fire services in Model I. Thus, it seems that the savings from co-operation have been made on both the administration and the operation of the fire service.

Another hypothesis is that fire services using full-time firemen at daytime and part-time firemen at night have been more productive. The results also show that these fire services have a positive total productivity change due to catching-up with the frontier units.

To further investigate the relationship between the two components of the index and total productivity, all three productivity indexes for Model I have been plotted against each other in Figure 3. The first plot (a) shows the relationship between technical change and the catching-up component, while the second plot (b) shows the relationship between technical change and total productivity. First, one notices that the plots are very similar. All units have had a negative technical change, which is due to the fact that the frontier mainly consists of part-time fire services. In addition, most full-time and mixed time fire services have both a negative technical change and a negative total productivity. In both plots there is a positive correlation between the indexes. The third plot (c) shows an almost perfect positive correlation between the catching-up component and total productivity. Thus total productivity is almost entirely due to fire services catching-up with the frontier and not to the shifts of the frontier.

Figure 3 also reveals that there is one fire service with larger productivity change than the rest. The reason for the high total productivity index and catching-up component is a very high increase in the population reached within 5 minutes, from 0 to 12460. The reason for this increase is that the number of full-time firemen in the first crew has increased from 4 to 5 men during daytime.

The arithmetic means for different municipality categories are shown in Figure 4. Rural municipalities have had the highest total productivity change in Model I. This is also the only category with a positive productivity change. Suburban and industrial municipalities have had the lowest productivity change. The catching-up component is larger than the technical change component for very large, large, rural, thinly populated and other municipalities. The opposite is true for suburban, medium and industrial municipalities. As noted above, it is the catching-up component that mainly contributes to total productivity change. Technical change has been negative for all categories, that is, the production frontier has shifted inwards for all different municipalities.

In Model II, other municipalities have had the highest total productivity change, while the differences between the rest of the categories are small. The catching-up component is above one for all categories except suburban and medium-sized municipalities.

Figure 3. Comparison between Malmquist total productivity (M), technical change (TC), and catching-up (CU) for Model I.

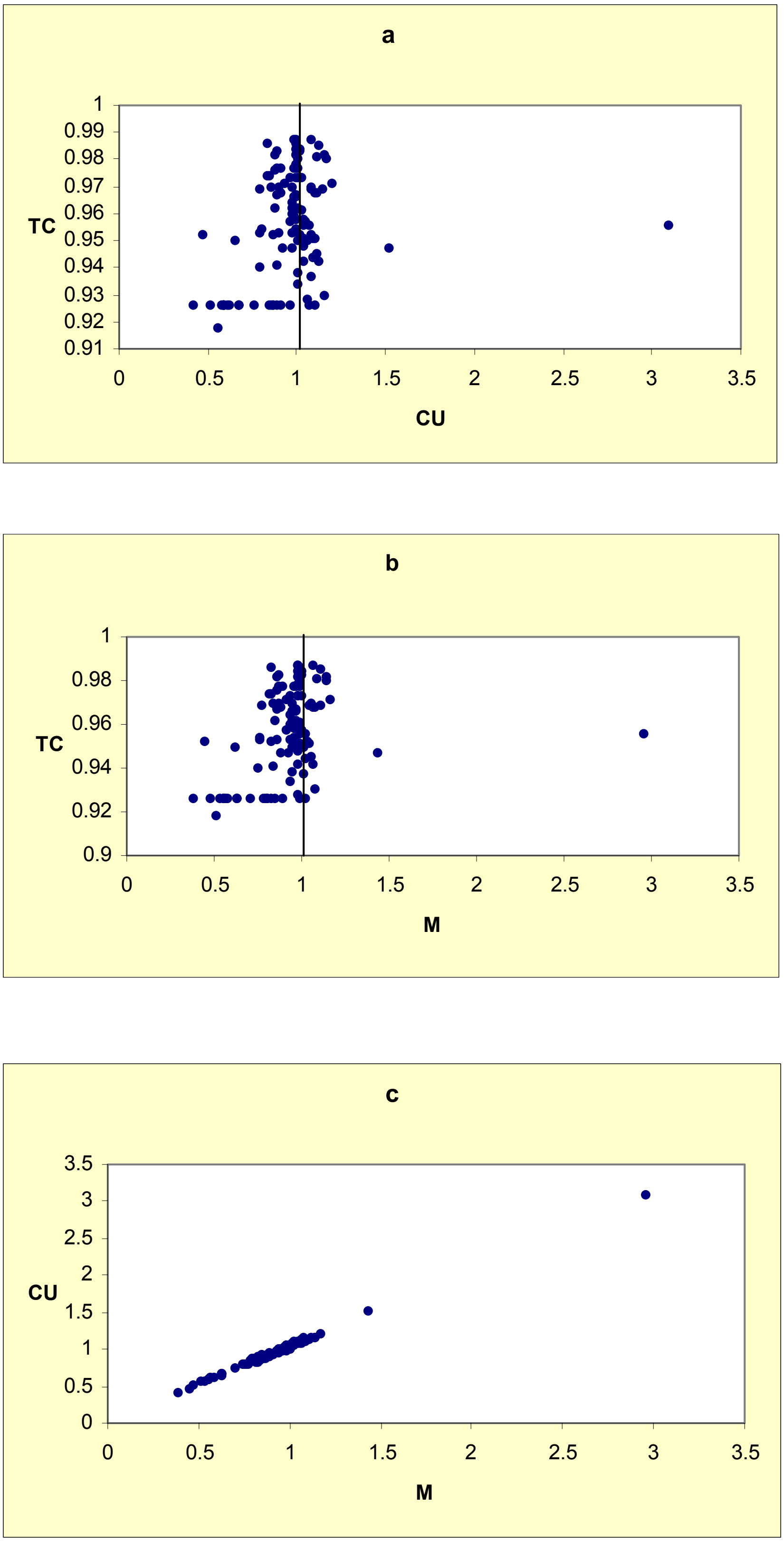
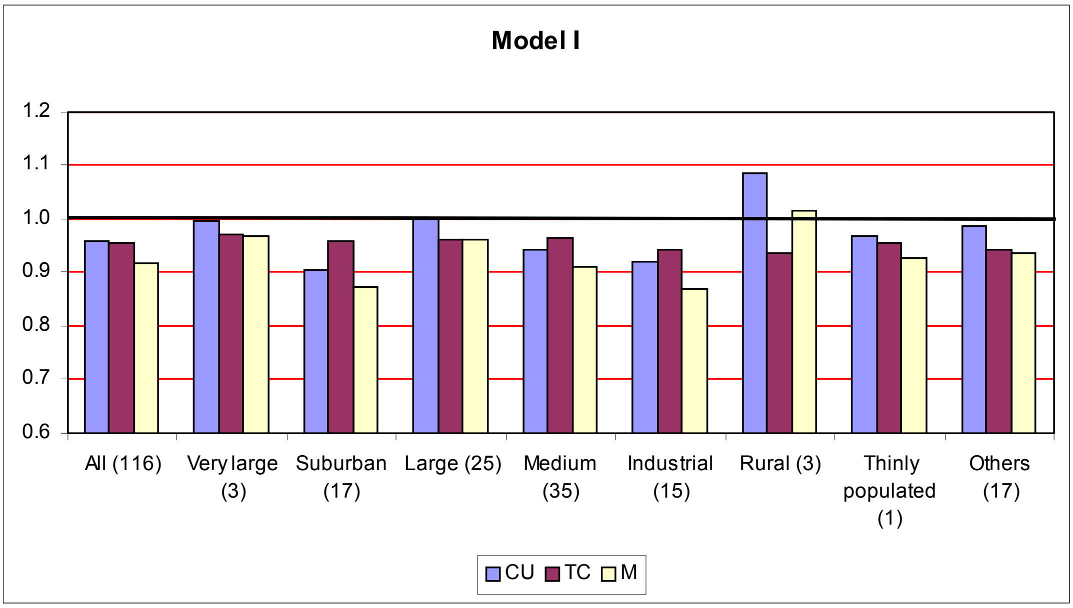


Figure 4. Catching up (CU), technical change (TC), and total productivity (M) for different municipality categories.



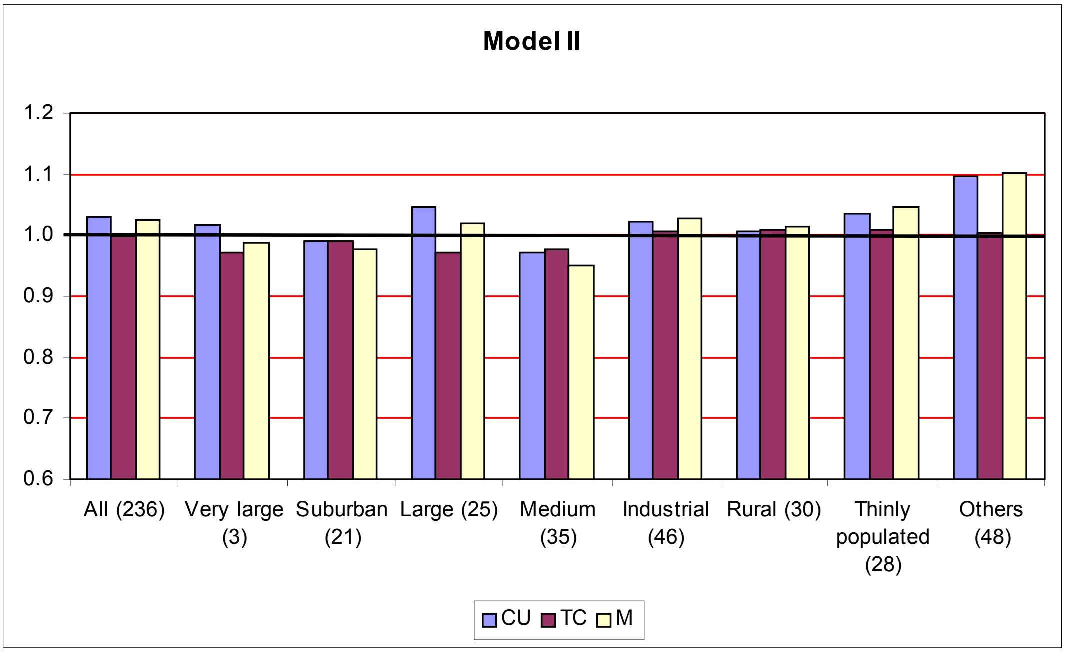


Table 3. Correlation between productivity indexes and input and output factors.

	% change of five-minute variable	% change of ten-minute variable	% change of full-time firemen	% change of part-time firemen	% change of total number of firemen	% change of total cost
Model I [a]						
CU	0.071 (0.449)	0.036 (0.702)	-0.569 (0.000)	0.303 (0.001)	0.017 (0.854)	-0.730 (0.000)
TC	-0.108 (0.252)	-0.144 (0.124)	-0.343 (0.000)	0.172 (0.067)	0.065 (0.492)	-0.472 (0.000)
M	0.054 (0.565)	0.019 (0.839)	-0.574 (0.000)	0.308 (0.001)	0.025 (0.794)	-0.739 (0.000)
Model II [b]						
CU	0.376 (0.000)	0.513 (0.000)	-0.198 (0.002)	-0.350 (0.000)	-0.460 (0.000)	-0.512 (0.000)
TC	-0.050 (0.444)	0.069 (0.289)	0.085 (0.193)	0.109 (0.095)	0.062 (0.346)	0.178 (0.006)
M	0.372 (0.000)	0.521 (0.000)	-0.195 (0.003)	-0.345 (0.000)	-0.459 (0.000)	-0.506 (0.000)

In parenthesis probability values for not rejecting H_0: Correlation=0.
[a] One outlier removed in calculations.
[b] Three outliers removed in calculations.

Another interesting issue is to find out the nature of the relationship between the productivity index scores and the changes of the output and input variables. Table 3 shows the correlation coefficients between the productivity numbers and the output and input factors for both models.[8]

In Model I, total productivity change, *M*, is negatively correlated with the change of the input variable (total cost). The same is true for the catching-up component, *CU*, and the technical change of the frontier, *TC*. That is, the less costly the fire service has turned, the more productive it has become. Productivity is positively correlated with the change of full-time firemen, and negatively with part-time firemen. Increased productivity is thus connected with switching from full-time firemen to part-time firemen. The change of the output variables has not affected productivity statistically significantly. Thus, it seems that choosing better locations for the fire stations does not influence productivity for this model.

[8] The reason for not doing regression analyses with the productivity indexes as dependent variables and the input and output variables as independent factors is that the correlations between the latter are very high.

For Model II the results are somewhat different. Total productivity change, *M*, and the catching-up component, *CU*, is now positively correlated with the coverage variables and the change of firemen. The negative correlation with the change in the number of firemen probably depends on the fact that this dimension is an input factor in this model and not an output factor as in Model I. The technical change of the frontier is, surprisingly, positively correlated with the input variable, which can be due to the fact that the number of frontier units are small.

When doing a productivity analysis on the public sector it is natural to ask whether public choice variables have any influence on the productivity index scores or not. Duncombe, Miner and Ruggiero (1997) discuss how public choice and bureaucratic models could be incorporated into a productivity analysis. Their conclusion is that there are four areas that should be interesting:
1) *Competition*. More competition should lead to higher efficiency.
2*) The size of the municipality*. A larger municipality could be less efficient. Another factor is the fiscal stress of the municipality.
3) *External factors*. Examples include wealth, education of the inhabitants, political interest and ideology.
4) *Internal characteristics*. Examples include training and the age distribution of the employees.

Which variables can then be chosen as proxy variables for the factors above? 1) Concerning competition, there is none among Swedish fire service since they all have local monopolies. 2) Concerning the size of the municipality, perhaps it is the relative size of the fire service compared to the total municipality that is interesting. The factors chosen here are the relative size of the fire services budget, the debt of the total municipality, and total number of the employed in the total public sectors in the municipality. 3) Concerning external factors, Bosch and Suarez-Pandiello (1995) lists several political factors that could influence local spending. The factors chosen here are the ideology of the political majority, political ignorance exercised as non-voting behaviour, tax income, and grant from central government. 4) Concerning internal characteristics, data on different training and age distribution are not available. However, the fire services are not only financed by taxes, but they also sell some services. The income of these services will be used as a factor for internal characteristics.

The above factors were used as independent variables in OLS-regressions with the productivity numbers as dependent variables. Five of the above variables were found to be statistically significant variables for at least one of the productivity numbers. The results from the regressions with these five variables are shown in Table 4.

Table 4. Public choice factors affecting productivity numbers.

Variable	**Mean of variable**	**Category**	**Total productivity (M)**	**Catching-up effect (CU)**	**Technical change (TC)**
Intercept			*0.8078 (0.1880)*	*0.8840 (0.1947)*	*0.9117 (0.0139)*
Total cost for municipality 1998, per capita	30.93	2	*0.0332 (0.0154)*	*0.0314 (0.0160)*	*0.0036 (0.0011)*
Share of population not voting 1994, percent	0.149	3	0.9683 (1.0927)	0.7642 (1.1320)	*0.2599 (0.0810)*
Government grants to municipality 1998, per capita	4.613	3	*-0.0493 (0.0218)*	*-0.0446 (0.0226)*	*-0.0070 (0.0016)*
Tax income for municipality, per capita	20.31	3	*-0.0450 (0.0193)*	*-0.0436 (0.0200)*	*-0.0038 (0.0014)*
Fire service external income 1993, per capita	113.13	4	*0.00071 (0.00019)*	*0.0070 (0.00019)*	*0.000041 (0.000014)*
R-square			0.157	0.1456	0.2372

Standard errors in parenthesis
Significant at 5% level in italics
Statistical source: Statistics Sweden

It seems that productivity change has been influenced by factors attributed to all three relevant categories (2, 3, 4). Perhaps surprisingly, higher total cost per capita for the municipality has led to a higher total productivity change, and so has external income for the fire service. A higher income could imply that more effort is placed upon lucrative activities instead of fire preventing and suppressing activities, but the results indicates that external activities are complements rather than substitutes to ordinary activities. Less surprisingly, government grants to the municipality and tax income have led to a lower total productivity change. The parameter estimates for the catching-up component are similar. For the technical change component, the above factors have the same signs, but the economical significance is lower for all since the absolute numbers are lower. One additional variable from category 3 also shows up as statistically significant; share of population not voting (surprisingly with a positive effect).

5 Summary and conclusions

The main question in this analysis has been whether the harsher budget stress that most municipalities in Sweden faced during the 90s has resulted in an increased productivity for the fire services. This question is analysed by studying the stand-by level with both dimensions of output, turn-out time and suppressing power, and using the Malmquist productivity index. The empirical results show that

- Productivity has decreased from 1992 to 1998 for full- and mixed-time fire services. Less input used has resulted in even less output produced.
- Both efficient and less efficient fire services have decreased their productivity.
- Rural fire services have had the highest productivity change, while suburban and industrial have had the lowest.
- Municipalities using full-time firemen at daytime and part-time at night have had a higher productivity change.
- There is a negative correlation between productivity change and change in input use. Thus, it thus seems that the best way of increasing productivity is to change from more expensive full-time firemen to less expensive part-time firemen.
- The decompositions show that the catching-up component is closely correlated with total productivity.
- Factors attributed to the size of the municipality, to external factors and to internal characteristics all statistically significantly affect total productivity.

How can this study benefit fire managers?

- They become aware that productivity changes depend both on changes of outputs and inputs, and not only on one of them. This makes it possible to compare their fire service's productivity with similar units, and not only to other units inputs and outputs.
- They become aware that total productivity change can be decomposed into a technical change and a catching-up component. This makes it possible to compare their fire service to what has happened to frontier units.
- They learn that considering one dimension of the stand-by activity of the fire service, i.e the number of people that can be covered within x minutes, will result in a different conclusion about productivity changes. Thus, fire managers become aware that it is important to use both dimensions.

In the future, there will hopefully be better statistics about costs of the production factors among the Swedish fire services. Future research should then concentrate on trying to include more input variables, such as capital and administrative inputs.

References

Andersson, C., 2001, *Insatstider för räddningsinsatser vid brand i bostad*, SRV/FOA, P21-375/01.

Bjurek, H., 1996, The Malmquist Total Factor Productivity Index, *Scandinavian Journal of Economics*, 98(2), 303-313

Bjurek, H., F. Førsund, and L. Hjalmarsson, 1998, Malmquist productivity indexes: an empirical comparison, in *Index numbers: Essays in honour of Sten Malmquist*, Eds. R. Färe, S. Grosskopf, and R. Russell, Kluwer Publishers.

Bosch, N. and J. Suarez-Pandiello, 1995, *Seven hypotheses about public choice and local spending*, Public Finance, 50(1), 36-50.

Bradford, D., R. Malt and W. Oates, 1969, The rising cost of local public services, *National Tax Journal*, 22, 185-202.

Charnes A., W. Cooper, A. Lewin, and L. Seiford, 1994, *Data envelopment analysis: Theory, methodology and applications*, Kluwer Academic Publishers, Boston.

Charnes, A., W. W. Cooper and E. Rhodes, 1978, Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2, 429-444.

Caves, D. W., L. R. Christensen and W. Diewert, 1992, The economic theory of index numbers and the measurement of input, output, and productivity, *Econometrica*, 50, 1393-1414.

Coelli, T. J., 1996, A Guide to DEAP Version 2.1 A Data Envelopment Analysis (Computer) Program, *CEPA Working Paper 96/8*, Department of Econometrics, University of New England.

Duncombe, W., J. Miner and J. Ruggiero, 1997, Empirical evaluation of bureaucratic models of inefficiency, *Public Choice*, 93, 1-18.

Färe R., S Grosskopf and C.A.K. Lovell, 1994, *Production Frontiers*, Cambridge.

Färe R., S. Grosskopf, Lindgren and P. Roos, 1994, Productivity developments in Swedish hospitals: a Malmquist output index approach, in A. Charnes, W. Cooper, A. Lewin, and L. Seiford, 1994, *Data envelopment analysis: Theory, methodology and applications*, Kluwer Academic Publishers, Boston.

Farrell, M., 1957, The measurement of productive efficiency, *Journal of the Royal Statistical Society*, Series A, III, 253-280.

Førsund, F., 1997, The Malmquist Productivity Index, TFP and scale, *Memorandum 233*, Department of Economics, Göteborg University.

Grifell-Tatjé, E., and C. A. K. Lovell, 1995, A note on the Malmquist productivity index, *Economics Letters*, 47, 169-175.

Jaldell, H., 2002, Measuring efficiency for Swedish fire services' stand-by level, Paper 3 this thesis.

Sträng, D., 1999, *Insatstid för räddningstjänst vid skydd mot olyckor*, SRV/FOA, P21-317/99.

Svenska Kommunförbundet, 1994, *Räddningstänstens personalkostnader* (Swedish Association of Local Authorities, Labour costs of the fire service), Stockholm.

Appendix

The efficiency scores calculated using the data envelopment analysis are obtained by solving linear programming problems. For example, to calculate E_{11} in equation (2) for unit A the following linear program must be solved

$$E_{11}^{A} = \min_{\lambda_A} \theta_A \qquad \text{(A1)}$$

subject to

$$y_{rA} \leq \sum_{j}^{N} \lambda_{Aj} y_{rj}, \qquad r = 1,\ldots, m \qquad \text{(A1a)}$$

$$\theta_A x_{iA} \geq \sum_{j}^{N} \lambda_{Aj} x_{ij}, \qquad i = 1,\ldots, n \qquad \text{(A1b)}$$

$$\lambda_{Aj} \geq 0, \quad \forall j = 1,\ldots, N \qquad \text{(A1c)}$$

where *m* is the number of outputs, *n* is the number of inputs, and *N* the number of units. Restriction (A1a) implies that the reference unit on the frontier, the peer unit (x^{1}_{A1}, y_{A1} in Figure 1), must produce at least as much as unit *A*, while restriction (A1b) implies that the efficiency adjusted volume of input used by unit *A* must at least amount to the input volume used by the reference unit.

5

Paper 5

Measuring performance differences using an ordinal output variable: The case of Swedish fire services

by
Henrik Jaldell

ABSTRACT

This paper investigates how to find performance differences in fire services with an ordinal output variable. The reason for not using a continuous output variable is that there is no continuous data available for fire suppression in Sweden. Instead, an ordinal output variable was constructed for fires in detached houses. Performance is measured by adjusting the outputs for inputs, such as number of firemen, using the ordered probit model. No performance differences were found between full-time and part-time firemen. The results also indicate that "team spirit" is more important for performance than the actual number of firemen fighting a fire.

Keywords: fire service, ordinal output, productivity, public sector, ordered probit

1 Introduction

Efficiency and productivity studies of the public sector have increased in popularity over time, mostly because of an interest in rendering it more effective due to the harsher fiscal environment. However, in some cases, such as the fire service, there are problems in finding data for relevant continuous output variables; this may be one reason why the fire service has been neglected in performance studies measuring efficiency and productivity. Another reason is the problem of specifying what the fire service actually produces. A further complication is that in the few fire service studies made by economists there has been no consensus on the definition of output.

The inverse of the insurance pay-out has been used as a continuous output measure for the value saved through fire suppressing by Bouckaert (1992) and Duncombe and Yinger (1993). The problem with using this variable is that it is not totally under the control of the fire service. Once at the fire it is not reasonable to assume that the effort of the fire service depends on the object of fire. The effort of the firemen is probably the same no matter if the house is worth SEK 500,000 or SEK 1,000,000. This makes this variable questionable for use in a study of productivity differences among fire services. Another problem is that data for this variable is not available in Sweden today on a fire service level. However, in the fire services' turn-out reports there is information both on the situation of the fire when the fire service arrived, and on where the fire was extinguished. This information could be used to construct an ordinal output variable.

The purpose of this paper is thus to show how an ordinal output variable can be constructed and used for studying productivity differences among Swedish fire services. The model used here could also be extended to other sectors where continuous output variables are hard to find, such as schools and universities (where grades are discontinuous), or public opinion research. However, no earlier productivity studies using ordinal output variables have been found.

Using the specified ordinal output variable it is possible to compare the performance of different fire services, but the condition of the fire services arrival at a fire may not be the same. For example some always send 5 firemen, while others always send 15. Naturally, the more firemen available the easier it is to extinguish the fire. To control for this factor, the ordinal output variable is used as a left-hand-side variable in an ordered probit model. The right-hand-side variables are the “inputs”, e.g., the number of firemen. The structure of the model thus corresponds to ordinary production function estimation, with the only difference being that the output variable is ordinal.

Since data exists for each turn-out in Sweden, dummy variables can be used to find out differences in performance among the fire services. This is similar to finding the individual effects in panel data models, where the two dimensions are cross-section and time. Here, however, the dimensions are cross-section and number of turn-outs. The differences in performance can be interpreted as productivity differences.

The paper thus includes three novel ideas. First is the construction of the ordinal output variable, which is presented in section 2. Then comes the use and motivation of the ordered probit model in the estimation of a production function in section 3. The third is the way performance or productivity differences are found by the “individual effects” approach, also presented in section 3. The results of the estimations are found in section 4, and conclusions in section 5.

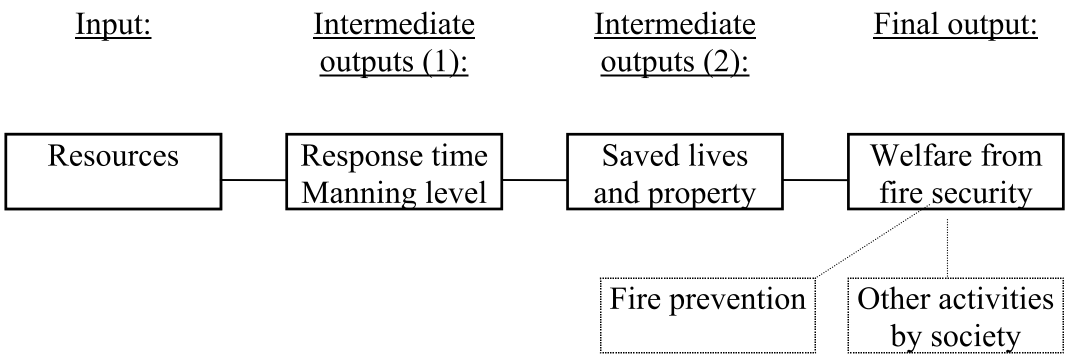
2 Definition and measurement of output and input

2.1 Definition of output

The fire services’ two main activities are to prevent fires and other accidents from happening, and if they happen anyway, to suppress them as effectively as possible. This paper focuses on the fire suppressing activity. Schaenman and Swartz (1974) proposed three measures for fire suppression output: response time, spread of the fire after the first unit's arrival, and value of property loss. These measures are interrelated, however: the value of property loss for example depends on the response time, as does the spread of the fire, since a fire usually grows exponentially. It is therefore necessary to divide output into more precise intermediate steps.

Here, extending the two-step model for public sector output suggested by Bradford, Malt and Oates's (1969), fire suppression is described as a three-step procedure with two levels of intermediate outputs and one final output, as in figure 1 below.

Figure 1. A schematic analysis of fire suppression activity.



The fire service receives resources from the municipality to prevent and suppress fires. The first intermediate outputs in fire suppression are the average response time, and the number of firemen initially reaching the fire. Disregarding costs, the shorter response time the better, and the more men arriving at the fire the better. If more men are needed to control the fire, it is also important that the reinforcement shall not take too long time. The fire service must decide how many are needed and what kind to use (full- or part-time), and, in the long run, how many stations to have and how to divide the firemen between the stations. For a full-time crew it shall take about 90 seconds to be on the way, while for a part-time crew it takes between 5 and 6 minutes, since the part-time crew is at work or at home. Part-time firemen usually have other full-time jobs, but do a 24-hour on-call service every third or fourth week.

The 288 municipalities in Sweden differ considerably both in population, from 2,800 to 700,000 inhabitants, and in area, from 9 km^2 to 30,000 km^2. Some municipalities have several fire stations with many full-time firemen, while others may have only one station and only part-time firemen, and still others have a mix of full-time and part-time firemen. Altogether there are over 800 fire and rescue stations. The cost for a force of full-time firemen is about six times higher than that of part-time firemen. The tradition is that larger municipalities have full-time, while smaller municipalities have part-time firemen. The number of firemen should ideally depend on the risk situation. However, in reality the number depends mostly on population and tradition.

When there is an alarm received at the fire station, for most stations, all firemen in turn out, since they have little chance of knowing how the fire has developed until they reach. To be allowed to use “smoke-divers” a minimum of five firemen are required in the force turning out.

The first intermediate outputs (response time and number of firemen) are then inputs in the production of the second intermediate output; how well the fire service succeeds in suppressing the fire. The main difference between fire suppression and other production processes is that fire is an organic process which changes all the time (Juås, 1994). The fireman’s task is to change this process. The output produced is the difference between the potential course and the result due to the fireman’s work. The ideal measure of this is saved value of property and number of saved lives.

The final output is the feeling of security or welfare of the inhabitants of the municipality, which depends on all the activities by the fire service (including fire prevention), but also on other factors such as tradition, environmental risk, weather and building conditions.

The three levels of output of the fire service have not been thoroughly taken into account before, which has led to the use of diverse output measures in economic analyses. Different researchers have studied different steps. Some have considered allocative efficiency (e.g. Duncombe and Brudney (1995), and Southwick and Butler (1985)); i.e. they studied whether supply of the final output meets demand. Others have estimated the supply or the cost function using proxies for the final output (e.g. Ahlbrandt (1973), Hirsch (1973), Duncombe and Yinger (1993)), i.e., looking at the relation between resources and final output. There are also researchers who have used only the first intermediate output (Wallace, 1977), a mix of the two intermediate outputs, or both intermediate and final outputs (e.g., Bouckaert (1992) and Kristensen (1983)) and related these to the resources. Juås (1994, 1995) calculated the marginal products of firemen and response time in moving from the first to the second intermediate output.

Studies of productivity differences between fire services are rare. Bouckaert (1992) used a graphical analysis for studying productivity differences between the Belgian fire services. Ahlbrandt (1973) and Kristensen (1983) studied productivity differences between private and public fire services in the state of Washington, USA, and in Denmark respectively.

2.2 Construction of output variable

In this paper the second intermediate output is studied. It not only depends on the response time and the number and achievement of the firemen, but also on many other things: the time between fire-start and discovery, and between discovery and alarm, the water supply, the number of fire vehicles and other equipment available, the weather, the building conditions, and random factors like open windows.

The “true” property value saved due to the fire service’s efforts is very hard to measure because of the need to control for the many environmental variables. In a stochastic framework using thousands of turn-outs it should be possible to approximate this value by the negative of the insurance money paid out. However, since there is no centrally collected data on how much insurance companies pay out for each fire, it is very hard work to connect information regarding inputs and output for each single turn-out.

In 1996 a new data-collection system began in Sweden. For each fire turn-out there is information about the condition of the fire when the fire service arrived, and about where it was extinguished. With this information it should be possible to draw some conclusions as to the performance of the fire service. In table 1 below the initial conditions are given in the rows, and where the fire was extinguished is shown in the columns. The υ’s in the table show the different possible outcomes.

In the table there are fourteen different reasonable outcomes; the question is what conclusions can be drawn from them.

Table 1. Condition of fire upon arrival and extent of spread before extinguished.

		Extinguished in:				
		1	2	3	4	5
Condition upon arrival:		Starting article	Starting room	Starting fire-cell	Starting building	Other buildings
1	Fire in the starting article	υ11	υ12	υ13	υ14	υ15
2	Fire in the starting room		υ22	υ23	υ24	υ25
3	Fire in several rooms (same fire-cell)			υ33	υ34	υ35
4	Fire in several fire-cells				υ44	υ45

It should be possible to rank them. Looking at each row it is quite obvious that having an initial condition of fire in the starting article, it is better to extinguish it in the starting article, than letting the fire spread to the starting room, which can be written $\upsilon 11 \succ \upsilon 12$.

This is true for all rows:

$\upsilon 11 \succ \upsilon 12 \succ \upsilon 13 \succ \upsilon 14 \succ \upsilon 15$;

$\upsilon 22 \succ \upsilon 23 \succ \upsilon 24 \succ \upsilon 25$;

$\upsilon 33 \succ \upsilon 34 \succ \upsilon 35$;

$\upsilon 44 \succ \upsilon 45$;

It should also be possible to rank the outcomes vertically. Compare an initial condition of fire in the starting object and fire in the starting room. If both have the same final condition, e.g. both fires are extinguished in the same room, the latter should be better than the first, so $\upsilon 22 \succ \upsilon 12$, and thus for all columns:

$\upsilon 22 \succ \upsilon 12$;

$\upsilon 33 \succ \upsilon 23 \succ \upsilon 13$;

$\upsilon 44 \succ \upsilon 34 \succ 24 \succ \upsilon 14$;

$\upsilon 45 \succ \upsilon 35 \succ 25 \succ 15$.

Combining these two assumptions, one can also see that:

υ22≻υ13≻υ14≻υ15;

υ33≻υ24≻υ25;

υ33≻υ14≻υ15;

υ44≻υ35≻υ25≻υ15.

Using these three ranking schemes, it is possible to rank all outcomes from better to worse. To minimise the influence of environmental effects, this study concentrates on the most homogenous of different turn-out objects, which, according to fire experts, is "fires in detached houses". For detached houses fire-cell and building is the same thing, and thus υ13=υ14, υ23=υ24, and υ33=υ34=υ44. A ranking of the possible outcomes for detached houses is shown in table 2 below. The reason for dividing υ22, υ33, υ11 and υ44 into two levels is that the first two, according to fire expertise, are tougher to fight.

Table 2. Ranking of possible outcomes for detached houses.

Level:			No.	y*=
Highest/Best	υ22	υ33	841	2
	υ11	υ44	1557	1
	υ12	υ23 υ35	580	
	υ13	υ25	56	0
Lowest/Worst	υ15		5	
		Total	3039	

There were 3239 turn-outs to detached houses in 1996 and 3062 in 1997. Many reports were not filled in correctly, and many turn-outs had outcomes not interesting in the analysis (e.g. only smoke), leaving a total of 3039 turn-outs used shown in table 2. Of the five levels in the table the worst ones had few turn-outs. The worst three levels have therefore been aggregated into one. Defining a three level ordinal output variable y*, the worst level (y*=0) includes υ15, υ13, υ25, υ12, υ23, and υ35; the middle (y*=1) level includes υ14 and υ44; and the best (y*=2) level includes υ22 and υ33.

2.3 Inputs

As mentioned above, many different input and environmental variables should be included in the model. The turn-out reports, however, restrict us to only three basic inputs: the response time, the number of full-time and part-time firemen, and the number of extra firemen helping out from nearby fire services.

There is unfortunately neither information on environmental variables, nor on capital inputs, such as fire trucks and water equipment. The lack of data on capital outputs may not be a very severe problem because, first, they are complements to the firemen, and second, they are quite similar over the country. The problem of not including environmental variables has been decreased by, first, only considering turn-outs to fires in detached houses and, second, for the estimation of individual effects, by only including fire services with at least 20 turn-outs.

Table 3. Mean, standard deviation, maximum, and minimum of dependent variables and number of turn-outs per fire service.

n=3039	Mean	Std. dev.	Max.	Min
Response time (seconds)	745.0	455.9	7800	79
No. of own firemen	9.72	5.29	20	1
No. of extra firemen	0.37	1.46	10	0
Full-time (=1)/part-time (=0)	0.33	0.13	1	0
Life saving (yes=1, no=0)	0.016	0.127	1	0
Turn-outs per fire service	12.06	10.76	62	1

The descriptive statistics for the inputs and the number of turn-outs per fire service are presented in table 3. Response time is the time under the control of the fire service, i.e., the time from when the service receives the alarm to when the first man arrives at the fire. As discussed above, longer response times could lead to worse outcomes.

The next variable is the number of own firemen rather than the number of hours spent at the fire that would be more appropriate. Unfortunately, if there is an exchange of firemen, both forces will be added to the total number. Therefore the maximum number of firemen has been restricted to 20. For the number of extra firemen (from nearby fire services) we have a similar problem. Here the maximum has been restricted to 10. More firemen should have a positive influence on output.

It is interesting to study if there is a difference in performance between full-time and part-time firemen. A variable being 1 if there are only full-time firemen and 0 if there are only part-time firemen has been included in the model (if a mixed force is used the number is between 0 and 1). Full-time firemen are better trained, so there could be a positive relation between this variable and output.

The model also includes a dummy variable for life saving. This variable is included since the fire service always tries to rescue lives first. A house may burn down because the fire service concentrates on life saving. There should thus be a negative relation between this variable and the output variable. This dummy variable is equal to 1 if someone has been saved and 0 otherwise. Unfortunately it is also 0 if the house have been searched, but nobody was found.

3 The model

Holding environmental variables constant the production function for a continuous output, y, using an input vector, $\boldsymbol{x}$, is defined as

$$y = f(\boldsymbol{x}) \qquad (1)$$

The problem in our case is that the dependent variable is not a continuous variable, but instead an ordinal variable, y^*, with three levels. Thus the ordinary technique for estimating production functions is not applicable. However, each level of the ordinal variable can be associated with the continuous variable such that y^* belongs to the j'th category of y if $\mu_{j-1}<y<\mu_j$, where the μ's are constants. This is the ordered probit model proposed by McKelvey and Zavoina (1975). The model for our three level output variable (j=0, 1, 2) is [1]

$$y^* = \beta'\boldsymbol{x} + \varepsilon \qquad (2)$$

where $\varepsilon \sim N[0,1]$

$$y^* = 0 \quad if\ y \leq \mu_0,$$
$$= 1 \quad if\ \mu_0 \leq y \leq \mu_1,$$
$$= 2 \quad if\ \mu_1 \leq y.$$

[1] This presentation follows Greene (1993). For a graphical exposition of the ordered probit model see Becker and Kennedy (1992).

Thus there are two more constants (the μ's) which are unknown parameters to be estimated along with the β's (When the model is estimated with an intercept μ_0 is normalised to zero).

The model is estimated by maximum likelihood.[2] Using the normal distribution the three probabilities for the different outcomes are

$$\begin{aligned} &Prob\ (y^*=0) = \Phi(-\beta'\mathbf{x}), \\ &Prob\ (y^*=1) = \Phi(\mu_1-\beta'\mathbf{x}) - \Phi(-\beta'\mathbf{x}), \\ &Prob\ (y^*=2) = 1 - \Phi(\mu_1-\beta'\mathbf{x}), \end{aligned} \tag{3}$$

where Φ is the cumulative normal distribution.

The marginal effects are calculated as

$$\begin{aligned} &\frac{\partial Prob(y^*=0)}{\partial \mathbf{x}} = -\phi(\beta'\mathbf{x})\beta \\ &\frac{\partial Prob(y^*=1)}{\partial \mathbf{x}} = (\phi(-\beta\mathbf{x}) - \phi(\mu_1 - \beta\mathbf{x}))\beta \\ &\frac{\partial Prob(y^*=2)}{\partial \mathbf{x}} = \phi(\mu_1 - \beta'\mathbf{x})\beta \end{aligned} \tag{4}$$

where ϕ is the standard normal density function. This means that the sign of the β-parameters is interpretable as a positive or negative marginal effect for y*=0 and y*=2, but for y*=1 the marginal effect cannot be decided ex ante.

The data has two dimensions, the turn-outs, *t*, and the fire service, *i*. The number of turn-outs per fire service ranges from 1 to 62. Fire service-specific effects are estimated by adding a dummy variable for each fire service with at least 20 turn-outs. Of 253 fire services 37 has been given a dummy variable.

[2] The likelihood function can be found in McKelvey and Zavoina (1975) and in Maddala (1983).

The model to be estimated is then

$$y^*_{it} = \beta_0 + \sum_{k=1}^{5} \beta_k x_{kit} + \alpha_i + \varepsilon_{it} \tag{5}$$

where x_{kit} are the five input variables specified above, ε_{it} is white noise (distributed as in equation 2), and the α_i's are individual fire service-specific effects ($\alpha_i=0$ if the number of turn-outs is less than 20).

The model fits into a panel data framework as a one-way fixed effect error component model. The differences from the standard panel data models are that the dependent variable is ordinal, the panel is unbalanced (the fire services have different number of turn-outs), and the fixed effects are not estimated for all fire services.

The linear functional form does not imply that the inputs are assumed to be perfect substitutes. Since the probit model is derived from the cumulative normal distribution it follows the law of variable returns (the textbook case; first increasing and then decreasing marginal products).

4 Empirical results

To test for statistical significance of the dependent variables, the model was first estimated without the fire service dummies. The model was then re-estimated with the significant dependent variables.

4.1 Model without fire service dummy variables

The estimated parameters, their standard errors, and the marginal effects for the inputs (calculated at the mean of the total sample) are shown in columns 1-6 in table 4. The table shows that there are two significant variables at the 5% level in the model without fire service dummy variables: own firemen and extra firemen.

For own firemen the parameter estimate is 0.0717, which cannot be interpreted as a marginal effect, but the positive sign indicates that using more own firemen leads to fewer $y^*=0$ outcomes and more $y^*=2$ outcomes. To know the effect on the $y^*=1$ outcome the marginal effects must be calculated according to equations (4). The interpretation of the marginal effect is that using one additional own fireman leads to 2 percentage units fewer y*=0 outcomes, 0.4 percentage units less y*=1 outcomes, and 2.4 percentage units more y*=2 outcomes.

The parameter estimate for extra firemen is also positive indicating that getting help by firemen from nearby fire services leads to fewer worst (y*=0) outcomes, and more best (y*=2) outcomes. The parameter estimate is smaller than that for own firemen. Since most often the extra firemen arrive later, this indicates a diminishing marginal product for total firemen, which can be seen from the marginal effects: a smaller decrease in number of y*=0 and y*=1 outcomes, and a smaller increase in number of y*=2 outcomes for extra firemen compared to own firemen.

Response time is not a significant variable. The conclusion is not that fires do not have an exponential growth, but rather that the output variable includes outcomes from both short and long response times in all three levels.

The life saving variable is surprisingly positive. The hypothesis was that life saving is first priority, which would lead to a worse outcome for property when life saving occurs. For detached houses, however, there have not been many life savings in 1996 and 1997, and this could be the reason for the odd result. The variable is not significant at the 5% level.

Another hypothesis is that full-time firemen are better than part-time, since they have more training and more experience. The data does indicate a positive relationship between the share of full-time firemen and the outcome, but the variable is not significant. For detached houses it is thus not possible to reject the hypothesis that there is no difference in performance between full-time and part-time firemen.[3] One reason could be that fires in detached houses are not very complicated to fight.

[3] The model was also estimated with different variables for full-time and part-time firemen, but the conclusion (no difference in performance) was the same.

Table 4. Parameter estimates and marginal effects.

	Model without fire service dummies					Model with fire service dummies	
	param. estimate	standard deviation	marginal eff. y*=0	marginal eff. y*=1	marginal eff. y*=2	param. estimate	standard dev.
Response time	-.000031	0.000046	0.000015	0.00000030	-0.0000018	-	
Own firemen	0.0717*	0.00408	-0.0197	-0.0034	0.0237	0.0758*	0.00424
Extra firemen	0.0463*	0.0145	-0.0126	-0.0025	0.0150	0.0467*	0.0147
Full-time(=1)/ part-time(=0)	0.00638	0.0546	-0.0041	-0.0008	0.0049	-	
Life saving	0.2515	0.1685	-0.0613	-0.0267	0.0880	-	
β_0	0.1694					0.1424	
μ_1	1.4869[a]					1.5262	
Log-likelihood	-2944.95					-2879.80	

* Significant at the 5% level.
[a] This is not the estimated number, but follows the presentation of the model in this paper.

The values of the intercept (β_0) and of μ_1 have no economic relevance. The only thing that can be said is that since they differ in size the estimation procedure divides the data into three parts, which is what we wanted it to do.

Marginal effects in the ordered probit model depend on the size of the input itself, i.e. they are not constant over the variable. Since fire services use different numbers of men in their forces, an interesting question is how the marginal effect varies for number of firemen. Table 5 shows the marginal effects for 1 to 20 firemen. The marginal effect for the worst outcome (y*=0) becomes less and less negative as more men are added, while the best outcome (y*=2) becomes more and more positive. The middle outcome (y*=1) goes from positive to negative. It is hard to see in the table, but the model implicitly follows the law of variable returns as discussed above. Marginal products expressed in monetary value and returns to scale could be calculated from table 5 with the help of an average for each level; i.e. information about property value saved is not needed on a turn-out level, but more aggregated data could be used.

Table 5. Marginal effect for one additional own fireman.

One more of the x'th man	y*=0	y*=1	y*=2
1	-0.0279	0.0150	0.0129
2	-0.0274	0.0133	0.0140
3	-0.0267	0.0115	0.0153
4	-0.0260	0.0095	0.0165
5	-0.0251	0.0074	0.0177
6	-0.0242	0.0052	0.0190
7	-0.0231	0.0029	0.0202
8	-0.0220	0.0006	0.0225
9	-0.0208	-0.0017	0.0225
10	-0.0196	-0.0040	0.0236
11	-0.0184	-0.0062	0.0246
12	-0.0172	-0.0084	0.0255
13	-0.0159	-0.0104	0.0263
14	-0.0147	-0.0124	0.0270
15	-0.0135	-0.0141	0.0276
16	-0.0123	-0.0158	0.0281
17	-0.0112	-0.0172	0.0284
18	-0.0101	-0.0184	0.0286
19	-0.0091	-0.0195	0.0286
20	-0.0082	-0.0203	0.0285

4.2 Model with fire service dummy variables

The main reason for estimating the model is to compare the performance of the different fire services, while taking into account the fact that they use different amounts of inputs. This is a sort of productivity analysis, which is performed by keeping the statistically significant variables (own firemen and extra firemen), and then adding dummy variables for the 37 fire services having at least 20 turn-outs in the sample.

Parameter estimates for the firemen variables are presented in the last two columns of table 4. The estimates for both own firemen and extra firemen are still positive and significant, and in fact very little different from the first model; the marginal effects are also similar.

Fire service parameter estimates and marginal effects of the dummy variables are presented in table 6. The marginal effects are calculated at the mean of the total sample for all fire services, not at the mean of the specific fire service.

The most productive fire service is number 173, with a 25% higher probability for a best outcome compared to the average fire service, and about 12.5% lower probabilities for the other two outcomes. The least productive fire service in this model is number 213 with a 38% higher probability for a worst outcome.

The differences seem large, but the difference between the property value saved of a $y^*=0$ and a $y^*=2$ outcome is unknown, and may be small in monetary values. Furthermore managerial efficiency is very important in the fire service. As an official report puts it: "It is possible that the differences in knowledge and personal enthusiasm of the management of the fire service give a larger effect than for other services of the municipalities" (SOU 1981:82, p. 119, own translation). As can be seen in tables 5 and 6 the marginal effects between the fire services differ more than do the marginal effects for one additional own fireman. The conclusion may be that "team spirit" is as important as the actual number of firemen.

Table 6. Estimates and marginal effects for the fire service dummy variables.

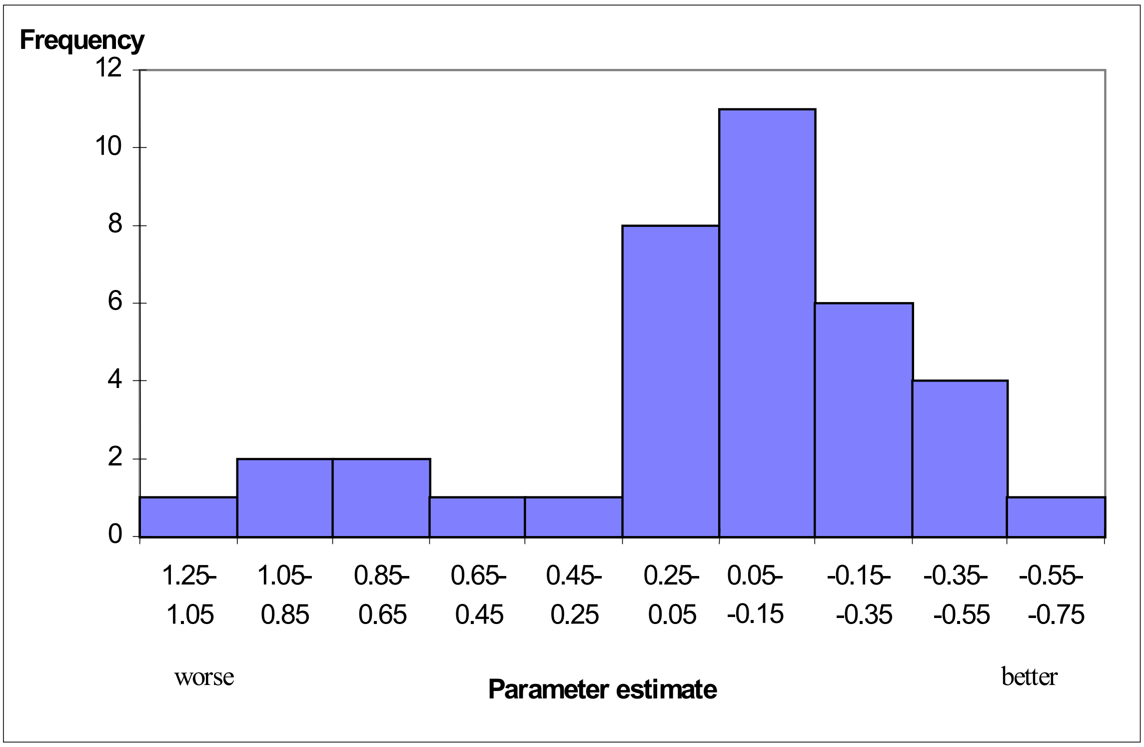
Fire service	Parameter estimate[a]	Marginal effect y*=0	Marginal effect y*=1	Marginal effect y*=2	Rank	No of y*=0	No of y*=1	No of y*=2	Total no. of turn-outs	Popula-tion
6	0.043	0.012	0.002	-0.014	22	5	15	13	33	50-100
8	0.679	0.229	-0.060	-0.169	33	28	11	18	57	100-
34	-0.302	-0.070	-0.038	0.107	6	2	14	9	25	100-
35	0.174	0.050	0.004	-0.054	26	5	8	7	20	100-
43	-0.189	-0.046	-0.019	0.065	10	2	36	19	57	100-
47	-0.106	-0.027	-0.009	0.036	14	6	21	8	35	20-30
60	-0.024	-0.006	-0.002	0.008	20	2	18	3	23	15-20
69	-0.251	-0.059	-0.029	0.088	9	1	39	10	50	50-100
71	0.108	0.030	0.004	-0.034	25	4	23	9	36	50-100
72	0.186	0.054	0.003	-0.057	28	5	9	8	22	20-30
83	-0.284	-0.066	-0.034	0.100	7	8	20	20	48	50-100
85	-0.273	-0.064	-0.032	0.096	8	3	24	9	36	30-50
86	0.054	0.015	0.003	-0.017	23	3	14	4	21	30-50
107	0.269	0.080	0.000	-0.080	31	5	14	5	24	50-100
109	-0.442	-0.095	-0.066	0.161	4	1	11	9	21	30-50
110	0.220	0.064	0.002	-0.067	30	4	11	6	21	50-100
118	0.178	0.051	0.004	-0.055	27	2	15	7	24	10-15
139	-0.095	-0.024	-0.008	0.032	16	4	24	12	40	100-
153	-0.065	-0.017	-0.005	0.022	19	6	15	9	30	50-100
158	0.866	0.303	-0.106	-0.197	35	11	7	3	21	10-15
167	0.060	0.017	0.003	-0.019	24	7	32	5	44	30-50
173	-0.671	-0.126	-0.125	0.252	1	0	16	12	28	100-
178	-0.073	-0.019	-0.006	0.024	18	7	21	12	40	20-30
184	-0.406	-0.089	-0.058	0.147	5	2	8	11	21	100-
188	-0.476	-0.100	-0.074	0.174	3	4	20	15	39	30-50
199	-0.091	-0.023	-0.007	0.031	17	2	19	8	29	30-50
203	0.528	0.171	-0.030	-0.141	32	8	13	1	22	30-50
206	-0.130	-0.033	-0.012	0.044	12	4	29	7	40	20-30
207	-0.491	-0.102	-0.078	0.181	2	1	10	11	22	20-30
208	0.781	0.269	-0.084	-0.185	34	14	7	3	24	20-30
209	-0.187	-0.046	-0.019	0.064	11	2	24	6	32	30-50
210	-0.112	-0.028	-0.010	0.038	13	8	32	12	52	100-
213	1.069	0.383	-0.164	-0.220	37	25	10	4	39	20-30
214	0.198	0.058	0.003	-0.061	29	7	16	3	26	20-30
216	0.011	0.003	0.001	-0.004	21	7	40	15	62	100-
224	-0.103	-0.026	-0.009	0.035	15	2	18	2	22	50-100
235	0.871	0.305	-0.107	-0.198	36	10	9	3	22	100-

[a] The more negative, the better.

The importance of not just comparing output levels without taking input into account can be seen by comparing fire service number 43 with number 69. They have a similar number of y*=0 and y*=1 outcomes, but fire service number 43 has almost twice as many y*=2 (best) outcomes. However, since fire service number 43 uses more firemen, number 69 is considered more productive and thus has a better ranking.

Unfortunately, the fire service-specific effects can neither be interpreted as a technical efficiency number as in the production frontier literature, nor can the scale of production be considered. However, the histogram in figure 2 shows that the fire service-specific effects are skewed to the left (worse), indicating that more fire services are closer to the best fire service than to the worst one.

Figure 2. Distribution of the parameter estimates of fire-service effects.



There exists a positive correlation between productivity (the negative of the parameter estimate for the fire service-specific effect) and population, and also between productivity and the number of turn-outs. The Pearson correlation coefficient is -0.175 between the size of the parameter estimate and the population (using group means); i.e., the higher population, the more productive. The correlation coefficient is -0.108 between the size of the parameter estimates and the number of turn-outs; i.e. the more turn-outs, the more productive. The reason for these results may be that more fires give better training, and that fire services in larger municipalities have better training facilities.

5 Conclusions

So far the fire services in Sweden have not been able to compare their fire suppression performance. This paper is a first step in the measurement of productivity for fire suppression. The output variable constructed in this paper can be used by individual fire services to measure their performance.

This paper also describes how marginal effects and productivity differences can be estimated with an ordinal dependent variable, using the ordered probit model. The reason for using the ordered probit model is that information on a continuous output variable is not available.

Individual effects were calculated for fire services with 20 or more turn-outs (only turn-outs to detached houses were considered). Since reports only exist for two years (1996 and 1997) only 37 of 253 possible fire service-specific effects could be estimated. The results show diversity in the fire service-specific effects, which indicate that “team spirit”, is possibly as important as the number of firemen. There is also a positive correlation between productivity and population and between productivity and number of turn-outs. The data does not reveal any difference in performance between full- and part-time firemen.

The drawbacks with using an ordinal output variable are that marginal products and returns to scale cannot be calculated. To use the results for calculating marginal products, we would have to assume some monetary values for levels of the dependent variable.

The specification of the output variable can of course be criticised for not being exact. The different levels may be different in different fire objects, and the difference between $y^*=0$ and $y^*=1$, compared to the difference between $y^*=1$ and $y^*=2$, is not known. However, in a statistical framework with over 3000 turn-outs it should be possible to draw some conclusions. Would the results have been different if a continuous variable was available? This question is also related to the question whether a house is harder to fight if it is worth 1.5 millions instead of 500,000; if not, then the measure used here is a better measure of performance.

The findings in this paper could be empirically extended in at least three ways: First, the specified output variable, y^*, could be checked with some turn-outs where data on amounts of the insurance money paid out are available. Hopefully the correlation between y^* and the money paid out would be negative. These values could also be used to calculate the value marginal product of additional firemen and returns to scale. Second, this paper only studies turn-outs to detached houses. A natural extension is to other fire objects such as blocks of houses, industrial buildings, and public buildings. Third, it would be interesting to investigate the hypothesis of the positive correlation between training and performance.

It would also be interesting to investigate theoretically if the productivity differences used here could be translated into input-saving and output-increasing efficiency, corresponding to these measures in the production frontier literature.

References

Ahlbrandt, R. C., 1973, Efficiency in the provision of fire services, *Public Choice*, 16, 1-15.

Becker, W. E., and P. E. Kennedy, 1992, A graphical exposition of the ordered probit, *Econometric Theory*, 8, 127-131.

Bouckaert, G.,1992, Productivity analysis in the public sector: The case of the fire service, *International Review of Administrative Sciences*, 58, 175-200.

Bradford, D., R. Malt, and W. Oates, 1969, The rising cost of local public services: Some evidence and reflections, *National Tax Journal*, 22(2), 185-202.

Coelli, T., D. S. P. Rao, and G. Battese, 1998, *An introduction to efficiency and productivity analysis*, Kluwer Academic Press, The Netherlands.

Duncombe, W., and Yinger, J., 1993, An analysis of returns to scale in public production, with an application to fire protection, *Journal of Public Economics*, 52, 49-72.

Duncombe, W., and D. Brudney, 1995, The optimal mix of volunteer and paid staff in local governments: an application to municipal fire departments, *Public Finance Quarterly*, 23(3), 356-384.

Fried, H. O., C. A. K. Lovell and S. S. Schmidt (eds.), 1993, *The Measurement of Productive Efficiency*, Oxford University Press.

Greene, W. H., 1993, *Econometric Analysis*, MacMillan Publishing, New York.

Hirsch, W. Z., 1973, *Urban Economic Analysis*, McGraw-Hill, New York.

Juås. B., 1994, Räddningstjänst vid byggnadsbränder (Fire service cover for fires in buildings), *Forskningsrapport 94:7,* Högskolan i Karlstad.

Juås. B., 1995, Tidsfaktorns betydelse vid räddningstjänstens insatser, *Forskningsrapport 95:15,* Högskolan i Karlstad, (summarised in B. Mattsson and B. Juås, 1997, The importance of the time factor in fire and rescue service operations in Sweden, *Accident Analysis and Prevention*, 29 (6), 849-857).

Kristensen, Ole, 1983, Public versus private provision of governmental services: The case of Danish fire protection services, *Urban Studies*, 20, 1-9.

Maddala, G. S., 1983, *Limited -dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

McKelvey, R., and W. Zavoina, 1975, A statistical model for the analysis of ordinal level dependent variables, *Journal of Mathematical Sociology*, 4:103-20.

Schaenman, P. S., and J. Swartz, 1974, *Measuring Fire Protection Productivity in Local Government: Some Initial Thoughts*, National Fire Protection Association, Boston.

SOU 1981:82, 1981, *Samhällets räddningstjänst: principbetänkande av räddningstjänstkomittéen.*

Southwick, L., and R. Butler, 1985, Fire department demand and supply in large cities, *Applied Economics*, 17, 1043-1064.

Wallace, R. J., 1977, Productivity measurement in the fire service, *Public Productivity Review*, 2(3), 12-36.