ETLA
ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B 00120 Helsinki Finland Tel. 358-9-609 900
Telefax 358-9-601 753 World Wide Web: http://www.etla.fi/

Keskusteluaiheita - Discussion papers

No. 570

Tanja Kirjavainen* - Heikki A. Loikkanen**

EFFICIENCY DIFFERENCES OF FINNISH SENIOR SECONDARY SCHOOLS: AN APPLICATION OF DEA AND TOBIT-ANALYSIS***

* Researcher, University of Tampere and Helsinki School of Economics and Business Administration

** Professor of Local Public Economics, University of Tampere

*** This is part of the "Effectiveness of schooling"-research program of the Academy of Finland.
A paper to be presented in the Meeting of INFORMS (Institute for Operations Research and the Management Sciences), November 3-6, 1996, in Atlanta; Georgia, USA.

ISSN 0781-6847

02.10.1996

Kirjavainen, Tanja - **Loikkanen**, Heikki A., **EFFICIENCY DIFFERENCES OF FINNISH SENIOR SECONDARY SCHOOLS: AN APPLICATION OF DEA AND TOBIT-ANALYSIS**. Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 1996, 40 p. (Keskusteluaiheita, Discussion Papers, ISSN 0781-6847; No. 570)

ABSTRACT: We study efficiency differences among Finnish senior secondary schools by Data Envelopment Analysis (DEA). Four model variants were used. In the simplest one both input and output variables were quantitative, in the most extensive ones also quantified qualitative variables were included. Average efficiencies in the most extensive models were 82-84 per cent, ranging from 44 to 100 per cent. When in addition to inputs related to the schools, parents educational level was treated as an input, average efficiency scores increased to 91 per cent. By ranking schools according to their efficiency scores, schools in the topmost and the lowest quartiles tended to maintain their ranking in alternative models whereas the centrally located ones were more mobile. The results also show that the rankings of schools by matriculation examination scores (an output variable) differed markedly from rankings by efficiency. As a second stage after DEA analysis, we explained the degree of inefficiency (100-efficiency score) by a statistical Tobit model. Schools with small classes were inefficient whereas school size did not affect efficiency. Somewhat surprisingly, private schools were inefficient relative to public schools, and grant ratios, i.e. ratios of state grants to accepted or actual educational expenditures had in some models a positive and significant effect on efficiency. When parents' education level was not as an input in DEA, but it was included in the Tobit model, it had a positive relation to efficiency.

KEY WORDS: education, efficiency, data envelopment analysis (JEL 121)

Kirjavainen, Tanja - **Loikkanen**, Heikki A., **EFFICIENCY DIFFERENCES OF FINNISH SENIOR SECONDARY SCHOOLS: AN APPLICATION OF DEA AND TOBIT-ANALYSIS**. Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 1996, 40 s. (Keskusteluaiheita, Discussion Papers, ISSN 0781-6847; No. 570)

TIIVISTELMÄ: Tutkimme suomalaisten lukioiden tehokkuuseroja Data Envelopment Analysis (DEA) -menetelmää soveltaen. Neljästä mallivaihtoehdosta suppeimmassa kaikki panos- ja tuotosmuttujat ovat kvantitatiivisia, laajemmissa malleissa niihin sisältyy myös kvantifioituja laadullisia muuttujia. Laajimmissa malleissa keskimääräiset tehokkuusluvut olivat 82-84 prosenttia vaihdellen 44 prosentista 100 prosenttiin. Kun kouluun liittyvien panostekijöiden lisäksi vanhempienm koulutustaso otettiin panostekijäksi, keskimääräiset tehokkuusluvut kasvoivat 91 prosenttiin. Kun koulut asetettiin tehokkuusluvun mukaiseen järjestykseen, ylimpään ja alimpaan kvartiiliin sijoittuneet koulut säilyttivät yleensä paikkansa, mutta keskimmäisissä kvartiileissa sijaistsevat olivat "liikkuvampia". Tulokset osoittivat myös, että koulujen paremmuusjärjestys ylioppilaskirjoitusmenestyksen perusteella (yksi tuotosmitoista) poikkeaa huomattavasti tehokkusluvun mukaisesta järjesteyksestä. DEA:n jälkeisenä toisena vaiheena lukioiden tehottomuuslukuja (100-tehokkuusluku) selitettiin tilastollisella Tobit mallilla. Luokkakooltaan pienet koulut olivat tehottomia, mutta koulukoolla ei ollut merkitsevää vaikutusta. Yllättäen yksityislukiot olivat julkisia kouluja tehottomampia ja lukioiden valtionavun suhteella joko koulujen ns. hyväksyttyihin tai todellisiin menoihin oli joissakin malleissa positiivinen suhde tehokkuuteen. Kun vanhempien koulutustaso ei ollut panostekijänä DEA vaiheessa, mutta sisältyi Tobit mallin selittäjiin, sillä oli positiivinen suhde koulun tehokkuuteen.

AVAINSANAT: koulutus, tehokkuus, data envelopment analyysi (JEL 121)

Yhteenveto

Tutkimuksessa arvioidaan empiirisesti 291 suomalaisen lukion tehokkuuseroja. Tuotosten (suoritteiden) ja panosten välistä suhdetta analysoidaan Data Envelopment Analysis (DEA) -nimistä lineaariseen ohjelmointiin perustuvaa menetelmää käyttäen. DEA-menetelmän eräänä etuna erityisesti julkisten markkinahinnattomien palvelujen arvioinnissa on se, että sitä sovellettaessa monituoteyritysten kaltaisten palveluyksiköiden panos- ja tuotospuolta ei tarvitse aggregoida yhdeksi rahamääräiseksi suureeksi. Riittää, kun on määrällistä panos- ja tuotostietoa toimipaikoittain (lukioittain). Tutkimuksessa lukioiden toimintaa arvioitiin neljällä muuttujiltaan erilaisella mallilla. Suppeimmassa tapauksessa panostekijät olivat lukion opetus- ja muiden tuntien määriä, laajimmassa panoksina oli myös oppilasainesta, opettajien kokemus- ja koulutustasoa mittaavia muuttujia. Suppeimmassa mallissa tuotospuoli koostui kustakin lukioista vuonna 1991 ylioppilaaksi päässeiden määrästä ja vuosina 1989-91 (lukioaika) luokaltaan päässeiden määrästä. Laajimmassa mallissa tuotosmuuttujina oli myös ylioppilaskirjoitusmenestystä mittaavia pistemääriä.

DEA erottelee havaintoaineistosta tehokkaat (tehokkuusluku 1 tai 100 %) ja laskee tehottomille nollan ja yhden (0-100 %:n) välille sijoittuvan tehokkuusluvun, jonka voi tulkita seuraavasti: jos tehokkuusluku on esimerkiksi 0.85 jollekin lukiolle, sen pitäisi pystyä saavuttamaan sama suoritetaso 85 prosentilla nykyisistä resursseistaan, jos se olisi yhtä tehokas kuin havaintoaineiston tehokkaimmat yksiköt. Laajimmissa malleissa keskimääräiset tehokkuusluvut olivat 82-84 prosenttia vaihdellen 44 prosentista 100 prosenttiin. Kun kouluun liittyvien panostekijöiden lisäksi vanhempien koulutustaso otettiin panostekijäksi, keskimääräiset tehokkuusluvut kasvoivat 91 prosenttiin. Kun koulut asetettiin tehokkuusluvun mukaiseen järjestykseen, ylimpään ja alimpaan kvartiiliin sijoittuneet koulut säilyttivät yleensä paikkansa, mutta keskimmäisissä kvartiileissa sijainneet lukiot muuttivat asemaansa. Tulokset osoittivat myös, että koulujen paremmuusjärjestys ylioppilaskirjoitus-menestyksen perusteella (yksi tuotosmitoista) poikkeaa huomattavasti tehokkuusluvun mukaisesta järjestyksestä.

Tehokkuusanalyysin avulla saatuja tehottomuuslukuja (100-tehokkuusluku) selitettiin myös tilastollisella Tobit mallilla toisessa vaiheessa. Tulosten perusteella luokkakooltaan pienet koulut olivat tehottomia, mutta koulukoolla ei ollut merkitsevää vaikutusta. Yllättäen yksityislukiot olivat julkisia kouluja tehottomampia ja lukioiden valtionavun suhteella joko koulujen ns. hyväksyttyihin tai todellisiin menoihin oli joissakin malleissa positiivinen vaikutus tehokkuuteen. Kun vanhempien koulutustaso ei ollut panostekijänä DEA malleissa vaan yhtenä selittäjänä Tobit malleissa, sillä oli positiivinen vaikutus koulun tehokkuuteen.

CONTENTS

1 INTRODUCTION

Evaluation of schools is a complicated issue and there are various ways to approach the question of what is a good school. Especially in countries like Finland where the whole school system with rare exceptions is public (financed by the local and the central government) from primary school to universities, parents, pupils as well as mass media most often pay attention to output or achievement in evaluating schools. For instance in the case of senior secondary schools good schools are the ones in which pupils do well in the national matriculation examinations. Paying attention to output variables would only be fair if the resources of schools were identical. Budget officials and people responsible for the financing of schooling tend to pay attention to the expenditures. For them a good school has low expenditures per pupil which would make sense if the outputs of schools were identical.

In actual fact educational institutions, private or public, differ most often both in terms of inputs that they employ and their outputs and the challenge is to evaluate their performance in this kind of multi-dimensional setting. Thus there are questions related to defining and measuring both input and output variables, and choosing a method which would make it possible to measure the relation between inputs and outputs in a meaningful manner. As for output, one choice to be made is related to timing as the benefits of education spread over time. One can use school time indicators, measures related to the end of school or variables related to success (e.g. earnings) in later life for which education can be viewed as a key input. On the input side, in addition to defining what are important school resources, one has to consider the role of environment outside the schools such as pupils' family and community variables.

In addition to the choice of variables, one faces the challenge of how to measure performance. More specifically, if one wants to study efficiency as a weighted sum of outputs relative to a weighted sum of inputs, one key problem common to quite a few public services is the lack of market prices which could be used as weights. Thus, unless it is possible to use some cost based or otherwise determined weights, it would be helpful if the method employed would not require weights determined in advance, but rather would solve them as a part of the analysis.

The purpose of this paper is to study the efficiency of Finnish senior secondary schools with

Data Envelopment Analysis (DEA) which is an application of linear programming. The basic idea of the approach is to view schools as productive units which use multiple inputs and outputs. The method produces measures of schools' (relative) efficiency by deriving a frontier production function (efficiency frontier) and measuring the distance of observations to the frontier to get their efficiency scores. Observations on the frontier get an efficiency score of one (or 100 per cent) and those below the frontier get scores below one (below 100 per cent) depending on their location. To apply DEA, no weight of the input or output variables are needed. This is one of the reasons why the use of the method has spread tremendously especially in the evaluation of efficiency in the public service sectors including analyses of courts, health centers, ferries etc. As we will point out in our brief summary of previous studies, education was one of the earliest areas where the method was applied.

In this study we have constructed cross-section data on Finnish senior secondary schools in 1988-91 both from registers and by our own survey. As DEA is a non-parametric method, there are no classical statistical tests which could be used to evaluate the models employed. This is why we present results on efficiency distributions of alternative models and study how robust are the efficiency score rankings of schools. If the rankings change a lot with the addition of a priori meaningful input and output variables, the simplest indicators of efficiency ('partial productivities') are questionable. In addition to that we also investigate whether the efficiency in two most extensive models is related to some factors that are not included in the DEA-models like the scale and financing of schools.

This paper is organized as follows. In section 2 we present a selective summary of previous studies using DEA on measuring efficiency of schools. In section 3 we describe the nature of DEA analysis and how this method measures efficiency. Data, variables and model types are described in section 4. Results of the study are presented in section 5. First we describe in the form of graphs and figures the efficiency distributions emerging from DEA analysis. Thereafter the stability of efficiency distributions from alternative models is studied. Finally we discuss the results of explaining the efficiency differences with certain variables related to i.e. the scale of operation and financing of the schools. Section 6 offers some conclusions.

2 PREVIOUS STUDIES

Given the amount of inputs, theoretical production function defines the (Pareto) efficient set of outputs, i.e. it is not possible to increase the quantity of any output without decreasing the quantity of at least one other output. Correspondingly, given outputs, it is not possible to decrease the quantity of any input without increasing the quantity of at least one other input. Inefficiency manifests itself as a deviation from the production function. Thus empirical study of differences in productive efficiency involves two basic steps. First, one has to determine an empirical educational production function. Second, one has to construct a distance measure such that the efficiency of an individual observation depends on its distance from the empirical production function.

There are two alternative ways to determine the educational production function empirically. Most common way is to apply statistical methods by either estimating the frontier production function using regression analysis or other related methods. Residuals of the estimated model are then used to define measures of efficiency to each unit. These techniques have been reviewed by i.e. Lovell *et al.* (1993). In the field of education they have been applied by Barrow (1991) for measuring the efficiency of the Local Education Authority (LEA) in United Kingdom. [1]

The other possibility of defining the efficiency frontier is to use non-parametric methods that are based on linear programming. Data Envelopment Analysis (DEA) is an example of such methods. These methods have also been applied in different forms to define the educational production function. One of the major advantages of this approach is that it is fairly easy to incorporate several outputs into the analysis. In the following we shall survey previous educational input-output studies confining ourselves to non-statistical (non-parametric) studies which are similar to our own application.

The first study applying DEA studied the program and managerial efficiency of a federally

---

[1] Most of the empirical educational production function studies have, however, been analyses where average (instead of efficient frontier) relations between inputs and outputs have been estimated. For a review of studies of educational production functions using statistical methods see for example Hanushek (1986) and Cohn and Geske (1990).

sponsored program called Program Follow Through (PFT) charged with providing remedial assistance to educationally disadvantaged early primary school students (Charnes *et al.*, 1981). Thereafter there have been several studies that have applied DEA in measuring the efficiency of schools.

Most often studies applying DEA examine the efficiency differences of certain group of schools and test the applicability of DEA for detecting the differences. Examples of such studies are Bessent *et al.* (1982), Bessent *et al.* (1983), Bessent *et al.* (1984), Ludwin and Guthrie (1989), Färe *et al.* (1989) using the U.S. school data, and Jesson *et al.* (1987) and Smith and Mayston (1987) who studied the efficiency of school districts (LEAs) in United Kingdom. Other European studies are Bonesrønning's and Rattsø's (1992 and 1994) efficiency analysis of Norwegian high schools. The conclusion of these studies usually is that DEA is applicable to efficiency measurement of schools in the sense that it detects differences between schools and the results are fairly robust.

The results of DEA have also been compared to efficiency scores obtained by using more conventional regression analysis, where the efficiency scores are calculated from the residuals. Examples of these studies are Mayston and Smith (1988) and Sengupta and Sfeir (1986). In both of them the writers concluded that the efficiency rankings were different depending on the method and for this reason the method of analysis mattered. Sengupta and Sfeir (1986) also noted that the results obtained by using DEA were fairly robust.

Usually the input variables used in the DEA models are such that they are controllable by the school or school district depending on the level of analysis. However, one of the most significant and robust results of input-output studies have been that student socioeconomic status affects student achievement. This is a factor that is not controllable by the school but still it influences its results. There are two studies that have taken this fact into account (Ray, 1991 and McCarty and Yaisawarng, 1993). In these studies DEA have been applied to calculate the efficiency scores by using variables that are controllable by the school. Thereafter these efficiency differences were explained by students' socioeconomic status using either regression analysis (Ray, 1991) or Tobit models[2] (McCarty and Yaisawarng

---

[2] The use of Tobit-models is more appropriate resulting non-biased estimates since the dependent variable is restricted to vary between 0 and 1.

1993). In this case the corrected residuals measure the efficiency of each unit.

McCarty and Yaisawarng (1993) also tested the differences in results of incorporating the variable measuring the socioeconomic status into the original DEA model. According to them, the two modeling alternatives produced similar results in the sense that the efficiency rankings in both cases were positively and significantly correlated. There were, however, notable differences in the rankings which were mainly related to the intercorrelatedness of the controllable and uncontrollable variables and to possible outlier cases.

As we pointed out in the introduction, it is not quite obvious what are the inputs and outputs of educational processes and at what stage (timing) should they be measured. Because of this one should pay attention to the robustness of results. However, quite a few earlier studies only report the final results without information on how they depend on the choice of variables. In most of the studies the selection of variables seems to be based more on data availability than any other reason[3]. Our analysis is done by using quite large data set (in most of the earlier studies the data sets used were considerably smaller). We have 291 senior secondary schools in our analysis. In our study we also test the robustness of the DEA results by using four different models. In addition to that, we present models explaining the efficiency scores with variables related to the scale, financing and student body of the school and test whether these results are related to the choice of variables in the DEA-models.

[3] This is also the problem of many production function studies using statistical methods (see e.g. Hanushek, 1979).

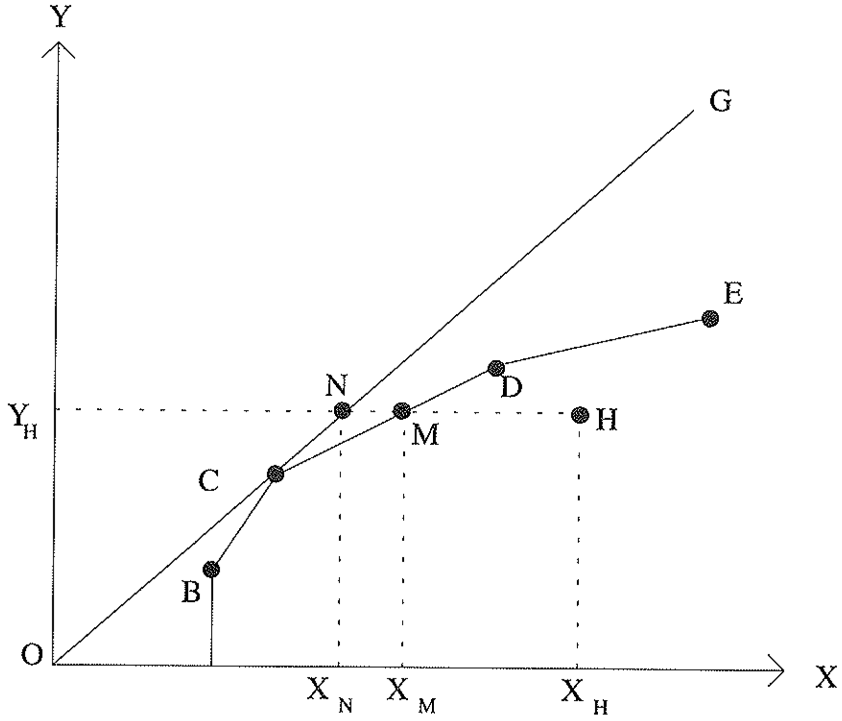
3 DATA ENVELOPMENT ANALYSIS

The purpose of this section is first to illustrate in a simple case what is meant by efficiency and how it is measured in DEA applications. Thereafter we present mathematical formulation of DEA problem. To see how the frontier production function (efficiency frontier) is determined non-parametrically by the DEA method let us consider a simplified educational sector consisting of schools which use one input to produce a single output. A cross-section picture of the sector is depicted in figure 1 where points B, C, D, E and H are the locations of schools in input (X) and output (Y) space. We shall first consider the determination of efficiency measure when the assumption of constant returns to scale are assumed to apply for the efficiency frontier. This means that we are looking from the data for the unit with highest productivity, i.e. maximum output to input ratio. In figure 1 the school at point C has the highest productivity as a line drawn from the origin to the observed points has the greatest slope in case of observation C. In this case line OG passing through point C determines the efficient production technology. All the other points are inefficient because their productivity is lower.

Having determined the efficiency frontier, the next step is to define measures of efficiency. Assuming the observation(s) on the efficiency frontier (in our case C) to be fully efficient, they are given an efficiency score of one (or 100 per cent). Observations under the efficiency frontier are inefficient, the degree of which depends on the extent to which their productivity is below that of point C. The efficiency score for school H can be determined as the ratio of efficient use of input X to the actual use of input X, i.e. $X_N/X_{H,}$, keeping output constant[4].

[4] alternatively, it can be determined as the ratio of actual output to potential output, keeping input constant. Under the assumption of constant returns to scale give the same efficiency scores. This is not true under the assumption of variable returns to scale. In this study, we calculate the efficiency scores assuming that schools are minimizing their use of inputs. Therefore we do not discuss the determination of efficiency scores in the case where schools are assumed to maximize their outputs.

Figure 1. Determination of efficiency frontier when constant and variable returns to scale hold.



Assuming that variable returns to scale hold, the efficiency frontier is a piecewise linear curve that passes through the points B, C, D and E. In this case only school H is inefficient. Taking output as given, efficient use of input X for school H is obtained at point M at the efficiency frontier. Accordingly, the efficiency score for school H is obtained as X_M/X_H.[5]

Mathematically, the efficiency score for school 0, assuming that schools minimize the use of inputs given outputs, is determined by solving a linear optimization problem (c.f. Charnes *et al.*, 1978). Let us consider n schools where school j uses the amount of x_{ij} of input i and produces the amount of y_{rj} of output r. We assume that $x_{ij} \geq 0$, $y_{rj} \geq 0$ and that each school uses at least one input to produce at least one output. By denoting the input weights by v_r (r=1,...,s) and output weights by u_i (i=1,...,m) the optimization problem can be formulated as

[5] In addition to technical efficiency which is depicted by the piecewise linear frontier function under the variable returns to scale assumption, in this case it is also possible to define a separate scale efficiency concept and to consider whether the schools are operating at the regions of increasing, constant or decreasing returns to scale (Banker *et al.*, 1984). We also studied these factors in our earlier report (Kirjavainen and Loikkanen, 1993) and the results are obtainable from the authors upon request.

follows assuming constant returns to scale:

$$\max_{\mu,\nu} \qquad w_0 = \sum_{r=1}^{s} \mu_r y_{r0} \tag{1}$$

$$\text{s.t.} \qquad \sum_{i=1}^{m} \nu_i x_{i0} = 1 \tag{2}$$

$$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \nu_i x_{ij} \leq 0 \qquad j = 1,\ldots,n; \tag{3}$$

$$\mu_r,\ \nu_i \geq \varepsilon \qquad r = 1,\ldots,s;\ i = 1,\ldots,m. \tag{4}$$

where ε is a small positive constant.

The maximizing problem is called the multiplier problem and it determines the efficiency score of school 0 by maximizing the sum of its weighted outputs (1) so that the sum of its weighted inputs equals one (2), and so that the weighted outputs of all schools subtracted by the weighted inputs of all schools is less than or equal to zero (3). This setting implies that the schools are either at the efficiency frontier or below it and the efficiency scores vary between 0 and 1 (or in terms of percentages the scores range from 0 to 100).

If we assume variable returns to scale according to Banker *et al.* (1984) the target function of the multiplier problem (1) as well as the second restriction (3) is modified by adding a constant term ω which determines values for the supporting hyperplanes passing through the dominating set of school 0. The values of ω specify whether the school 0 operates in the area of decreasing (ω>0), constant (ω=0) or increasing returns to scale (ω<0).

By denoting the input weights of school 0 by θ and the input and output weights of other schools by λ_j (j=1,..,n) we can write the dual of maximizing problem when constant returns to scale prevail as follows:

$$\min_{\theta,\lambda,\, s_r^+,\, e_i^-} \qquad z_0 = \theta - \varepsilon \sum_{r=1}^{s} s_r^+ - \varepsilon \sum_{i=1}^{m} e_i^- \tag{5}$$

$$\text{s.t.} \qquad \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{r0} \qquad r = 1,\ldots,s; \tag{6}$$

$$\theta\, x_{i0} - \sum_{j=1}^{n} \lambda_j x_{ij} - e_i^- = 0 \qquad i = 1,\ldots,m. \tag{7}$$

$$\lambda_j,\ s_r^+,\ e_i^- \geq 0 \tag{8}$$

Above, s_r^+ and e_i^- are so called slack variables measuring the excess of inputs and outputs. Small positive constant ε quarantees that inputs and outputs are positive and that the slack variables do not influence the target function z_0. The minimizing problem is called the envelopment problem and it determines the efficient use of inputs for school 0 (5) so that the outputs of school 0 are equal to the sum of weighted outputs of other schools (6). In addition, the weighted inputs of school 0 must equal to the weighted inputs of other schools (7). The optimal value of parameter θ in (7) determines the amount school 0 should reduce its use of inputs in order to be at the efficiency frontier and positive values of λ_j determine those schools that dominate school 0, i.e. form its comparative set.

If we assume that there exists variable returns to scale in the production, another restriction is added to the envelopment problem. It is of the following form

$$\sum_{j=1}^{n} \lambda_j = 1 \qquad (9) \quad ,$$

and it ensures that the efficiency frontier is a convex hyperplane.

4 DATA, MODELS AND VARIABLES

4.1 Data

Our study investigates the efficiency differences of Finnish senior secondary schools which provide general education after the nine year comprehensive school. Nearly 60 % of the comprehensive school leaving pupils opt for senior secondary schools. There are about 450 senior secondary schools in Finland and they provide education for about 100 000 students. They are mainly maintained by municipalities and financed by local taxes and state grants to municipalities.

Senior secondary school education is usually completed in three years. The school terminates in a national school-leaving examination, the matriculation examination. Passing of the examination gives general eligibility for university studies and vocational education intended for matriculated students. The examination comprises of four compulsory subjects: the student's mother tongue (Finnish/Swedish/Sami), the second national language of the country (Finnish/Swedish), a foreign language (English, French, German or Russian) and either mathematics or science and humanities. Besides the compulsory subjects the candidate may also take additional subjects. Teachers undertake initial grading of the exams, the final grading is done by a national matriculation examination board.

Our data consists of 291 senior secondary schools all over the Finland. The sample does not include all the senior secondary schools because we were not able to obtain all the necessary information from all schools. The data is cross-sectional and aggregated to the school level. Our key output variables are related to the results of matriculation examinations in the spring of 1991. Because senior secondary school lasts for three years those students who matriculated in the spring of 1991 started their studies in the Fall of 1988. Therefore our data covers the years between 1988-91. Whenever possible, we have measured the input variables as averages over the whole period. One of our output variables, number of students who passed their grade, is an average over three years, too.

The variables used for explaining the efficiency differences are also averages over the years 1989-91 in case of variables measuring the scale of operation. The variables depicting the financing of the schools are, however, information on 1992. This should not be a problem since the state grant system has not changed during the years of 1988-92. As for other variables e.g. those depicting the environment and student body composition of the schools, the information on the year of 1991 is used.

4.2 DEA-models and variables

In studies of educational production function there are usually two alternative ways of describing the influences of schooling on student achievement (see e.g. Hanushek, 1979). Either one takes into account the cumulative influence of family background, peers, school inputs and innate abilities on student achievement at certain time point or measures these factors during the period student is attending school. We use the latter alternative in our efficiency analysis. This so called value added model of the educational production function at the level of an individual student can be written as

$$A^t_i = f(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)}, I_i, A_i^{t^*}) \tag{10}$$

where A_i^t is a vector of variables measuring student i's achievement at time t, $B_i^{(t-t^*)}$ is the vector of family background influences over the period t^* to t, $P_i^{(t-t^*)}$ is the vector of influences of peers over the period t^* to t, $S_i^{(t-t^*)}$ is the vector of school inputs of ith student over the period t^* to t, I_i is the vector of innate abilities of ith student, and $A_i^{t^*}$ the outcomes of the ith student in earlier period. This formulation evaluates the educational achievements of the student by paying attention not only to inputs controllable by the schools but also taking into account the influences of student's innate abilities, former outcomes, family background and peers.

The value added model is convenient in the sense that it reduces the data requirements. In the applications of the value added model the data may consist of information on individuals or it may be aggregated for example to school level. In our case the units of observation are schools and the performance indicators measure the achievement of pupils in each school.

DEA is not a statistical method with which the theoretically based hypotheses could be tested with classical tests. Therefore, our strategy was the following. We tried to construct variables which would be operational counterparts to at least some of the elements in (10) in a form typically used also in the educational production function literature. Because of the lack of clear criterion for selecting the number of variables to be included in the analysis, we ran four different models. The guiding principle in the construction of models was to proceed from a simple one to more complicated ones. The simplest version included only a few quantitative input and output variables, whereas in the more complicated model quantified measures of qualitative factors were incorporated. This strategy enabled us to test the stability of the results. The input and output variables included in our four models are shown in table 1. The summary statistics of the variables are reported in Appendix 1.

Table 1. Variables used in the DEA-efficiency measurement.

Inputs:	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Teaching hours per week	x	x	x	x
Other than teaching hours per week	x	x	x	x
Experience of teachers			x	x
Education of teachers			x	x
Admission level		x	x	x
Education of students' parents				x
Outputs:				
Number of students who passed their grade (were moved up)	x	x	x	x
Number of graduates	x	x	x	x
Score in compulsory subjects in matriculation examination			x	x
Score in other subjects in matriculation examination			x	x

Our model 1 consists of simplest quantitative input and output variables. According to this model schools were depicted as producers of students who pass their grade or matriculation exam with teaching and non-teaching hours of the school staff as the input variables. We included as output variables both the number of students who passed their grades after first and second year (average of 1989-91) and the number of graduates in matriculation examination because the input variables measured the whole teaching load in each school. The input variables were measured by the number of teaching and non-teaching hours per week (average of 1989-1991).

Pupils to senior secondary schools are chosen by a "cream skimming" procedure, i.e. by choosing the best applicants on the basis of their comprehensive school reports (subjects graded using a scale from four to ten). For this reason the student body and their earlier educational achievements differ from school to school. In model 2 we included as an input a variable controlling the quality of students in the school.[6] It is measured by the admission level of the school (the lowest grade with which the school could be entered). This variable is from the Fall of 1988 and it is multiplied by the number of students that entered the school at that time. Only information of this cut off point could be obtained by our survey, although the average grade of each admitted pupil would have been a better measure.

Model 3 consisted of two additional qualitative input and output variables. Even though teacher characteristics are rarely found to have on impact on student achievement in statistical analyses, we wanted to study the influence of adding both teacher education and experience on efficiency distributions and efficiency rankings. The education of teachers was measured by the number of teachers having at least master's degree. The experience of teachers was measured by giving one point for each five professional year period and then adding these points up in each school. These variables were also averages over the years of 1989-91.

The two additional output variables measured the achievement of students by their scores in matriculation examination (spring of 1991) divided into achievement in compulsory subjects and in additional subjects.[7] The matriculation examination score in each subject has a range from one (improbatur = fail) to six (laudatur). Our school level variables are sums of pupils' scores. By using these two variables instead of one composite one we allowed schools to have a different emphasis on their provision of courses.

---

[6] There are also alternative ways of taking into account the student body of the school. One possibility is to measure the output as a value added in the sense recommended by Hanushek and Taylor (1990). Bønesrønning and Rattsø (1994) used this approach in their analysis. In our case this procedure was not possible because of lack of student level data about the scores in comprehensive school.

[7] As mentioned earlier the grade in compulsory subjects consist of mother tongue, the second national language of the country, foreign language and mathematics or science and humanities. Additional subjects may include grades in foreign language, mathematics, and science and humanities. Finnish speaking candidates take the same examination only in mother tongue and in the second national language. Foreign language examination differs depending on the language student has chosen to study as comprehensive. Also mathematics may be taken either as a comprehensive or as a short course. Science and humanities examination covers a wide range of subjects and students may choose to certain extent subjects they want to answer. Because of these choices the same score in matriculation examination has been achieved in various ways.

Even though there exist mixed results in the literature concerning the role of various inputs on educational achievement, socioeconomic background is quite consistently found to affect the success of students (see e.g. Hanushek, 1986). In model 4 we studied the role of students' socioeconomic status by the educational level of their parents even though this variable is not a pure input factor because it is not controlled by the school. The variable is constructed by first calculating the average educational level of both biological parents of those students who matriculated in the spring of 1991 and then multiplying it by the number of matriculated students. The educational level is measured by giving points for degrees[8].

Finally, we shall comment the role of scale variables in our models. Most of our variables both on the input and the output side are related to the size of the schools. Class size which is a ratio type variable, is not included directly in the variable list. Class size, however, affects the results somewhat indirectly as the number of graduates and the number of teaching hours are included and their ratio depends on class size.

---

[8] Following points were given to each degree: 1,5 = no other degrees than comprehensive school diploma. 3 = lowest vocational degree (approximately 10-11 years of schooling). 4 = medium vocational degree (approximately 12 years of schooling). 5 = highest vocational degree, not a university degree (13-14 years of schooling). 6 = bachelor's degree. 7 = master's degree. 8 = post graduate degree.

5 RESULTS

In this section we shall present the empirical results of this study. In presenting DEA results, we measure and interpret efficiency scores assuming output given so that scores below one (or 100 per cent) indicate savings possibilities in the use of inputs. As pointed out in section 2, under CRS the efficiency scores are the same if we had chosen to keep inputs constant and measured efficiency in the output increasing direction. However, under VRS this choice matters.

In addition to having four models with different inputs and outputs, we shall present results both assuming constant returns to scale (CRS) and variable returns to scale (VRS). In this connection, it is worth noting that the efficiency scores for each observation under VRS are typically smaller than (at maximum equal to) those under CRS (c.f. figure 1). Also, as the number of variables in a DEA model increases, the efficiency scores either increase or remain the same tending to increase average efficiency scores. Thus e.g. average efficiency scores of different models cannot be compared to each other as such.

In section 5.1 we describe the results of different DEA models. In section 5.2 we pay attention to the stability of the rankings of schools according to efficiency scores. Stability of rankings is important especially for policy purposes if there is limited information on educational processes. Then, we would know that efficient (and inefficient) schools remain similarly judged irrespective of how detailed information we have on them.

In section 5.3 we first illustrate how efficiency is related (or unrelated) to a number of variables of interest like the class size and school size from the input side, and to matriculation examination results from the output side. This is to illustrate that efficiency as a ratio of weighted outputs and weighted inputs is a different concept than any typical input or output indicator. In the end of this section we also present Tobit models in which efficiency scores derived from the most extensive models are explained by variables related to the scale of operation, student body composition, and financing of the schools.

5.1 Efficiency distributions

Table 2 offers basic information on the distribution of efficiency scores obtained from our four models both assuming CRS and VRS. Although the efficiency scores of different models are not readily comparable, it seems that the efficiency differences among the Finnish senior secondary schools were quite considerable regardless of the model used. In model 1 where only quantitative factors were included, the efficiency scores ranged from 17 per cent to 100 per cent when CRS was assumed. Even though the differences diminished in models with more inputs and outputs, in models 2 (admission level added) and 3 (teachers' experience and education as input and two matriculation examination scores as outputs added) the efficiency scores still ranged from 44 per cent to 100 per cent. In model 4, which consisted of variables of model 3 and educational level of pupils' parents, the efficiency varied from 60 to 100 per cent.

Table 2. The average efficiency, minimum and maximum of efficiency scores, and percentage share of efficient schools in each of the models.

	Model 1		**Model 2**		**Model 3**		**Model 4**	
	CRS	**VRS**	**CRS**	**VRS**	**CRS**	**VRS**	**CRS**	**VRS**
Mean	66,26	72,23	77,48	80,00	81,89	84,12	91,30	93,73
Minimum	16,99	40,61	38,14	49,03	43,82	58,40	59,74	59,81
Maximum	100	100	100	100	100	100	100	100
%-share of efficient schools	1,4	5,8	3,1	7,6	10,0	16,5	21,0	33,3

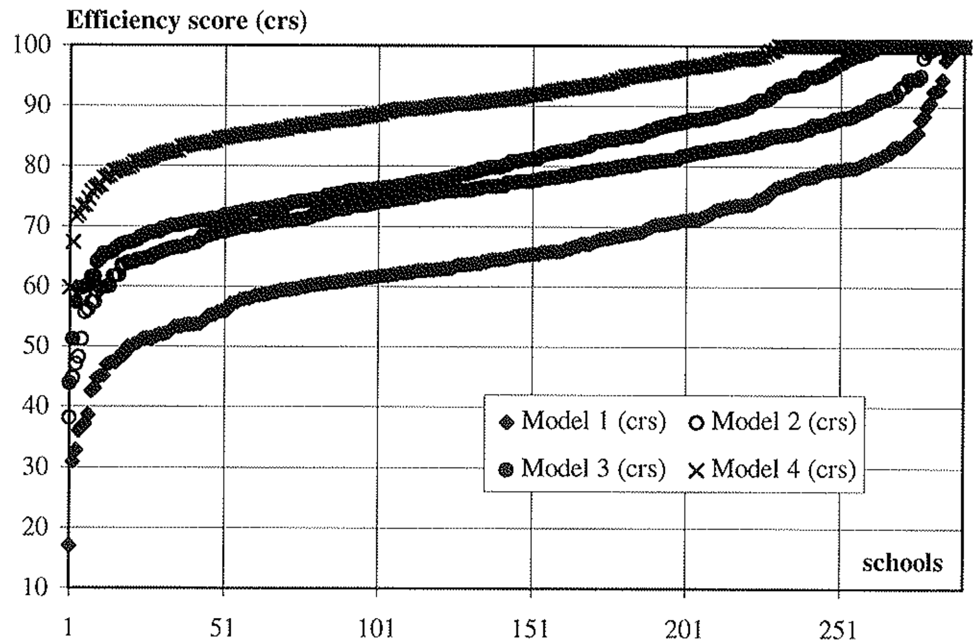
The average efficiency in model 1 was 66 per cent assuming CRS indicating an average savings potential of 34 (= 100-66) per cent in the use of resources. With addition of variable measuring student quality (admission level in model 2), it increased to 77 per cent. The average efficiency was 82 per cent in model 3, leaving only a savings potential of 18 per cent. Thus the difference between results of models 2 and 3 in terms of variation and averages are relatively small.

What is quite remarkable is the increase of average efficiency when parents' educational level is included as an input to the educational process. The average efficiency increases from 82 per cent in model 3 to 91 per cent in model 4. Finally, we note that the share of fully efficient schools (score 100 per cent) i.e. schools determining the frontier, increased from only 1,4 per cent in model 1 to 10 per cent in model 3 and all the way to 21 per cent in model 4.

When VRS was assumed, as expected the variation of efficiency scores was smaller. The minimum values and the average efficiencies were higher in all models. The differences between CRS and VRS results are considerable in model 1 but the become already rather small in case of model 3. Especially in model 4 there is a very small difference between CRS and VRS results which is not surprising given that the efficiency scores are so high under CRS. Referring to figure 1, if all observations are close to CRS efficiency frontier OG, also the VRS efficiency frontier must be close to OG implying also that there are not large scale inefficiencies involved in case of model 4.

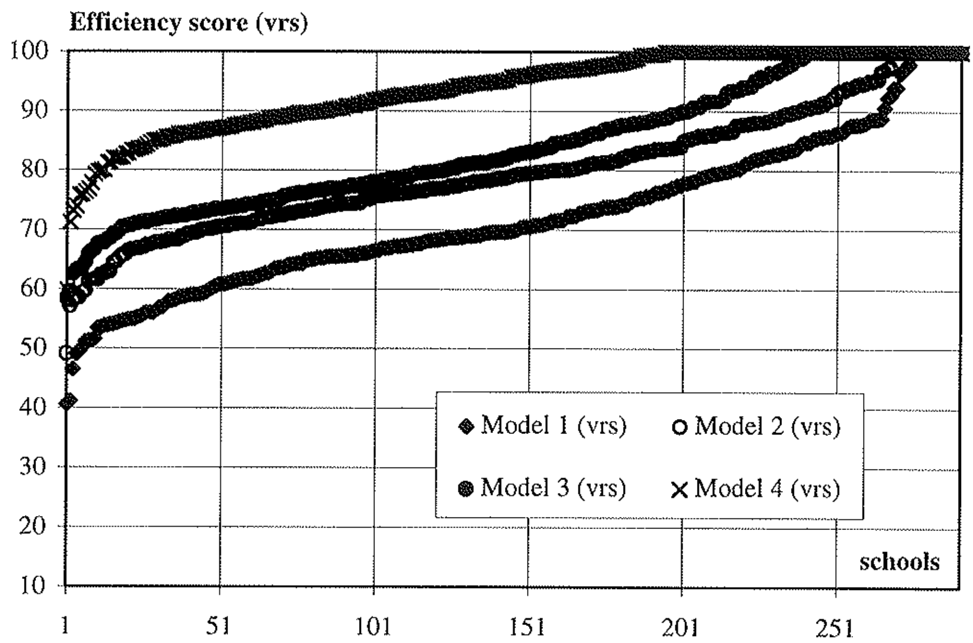
The results of different models can also be compared by depicting the efficiency distributions as shown in figure 2. There, the schools are ranked according to their efficiency score from lowest (number 1) to highest (291) in each model. The ranking numbers of schools are in the vertical axis and the efficiency scores in the horizontal axis. The efficiency distributions in figure 2 are based on the results of models 1-4 assuming CRS.

Figure 2. The efficiency distributions of models 1-4 when CRS is assumed.



The efficiency scores of model 1 are clearly below the other three models. The addition of the variable (admission level in model 2) measuring the quality of students shifted the efficiency distribution remarkably upwards whereas the efficiency distributions of model 2 and 3 are close to each other. Thus, the addition of four different variables measuring quality did not have a large influence on the efficiency score distribution. The addition of variable measuring the educational level of students' parents (model 4) shifted again the distribution clearly upwards. The same general pattern of changes in efficiency distributions emerged when VRS was assumed (figure 3) with the difference that the level of efficiency scores was higher and the number of efficient schools (score 100 per cent) was greater than under CRS in each model.

Figure 3. The efficiency distributions of models 1-4 when VRS is assumed.



5.2 The stability of results

There are different factors which may affect the stability of the DEA results. First, the frontier may be partly based on such outlier units that are very different from other units either genuinely or because of misscoding, measurement error etc. In such a case, omittance of these outliers can change the mean efficiency and rankings based on efficiency scores. Second, besides affecting mean efficiency and the whole efficiency score distribution, the use of different combinations of inputs and outputs may also change the ranking of individual schools.

5.2.1 Jackknifing with outlier observations

To study whether there were extreme outliers which affected frontier and efficiency scores in each of the models, we ran DEA analyses by dropping out each efficient school one at a time from the analysis. This is a procedure called jackknifing[9] and it tests the robustness of the DEA results. For example, in case of model 1, we ran four additional DEA analyses. We then tested the similarity of efficiency rankings between the model with all the schools included and those based on dropping each efficient unit out at a time by the Spearman rank correlation coefficient. We also calculated the mean efficiency from the iterations and the standard deviation of the mean efficiencies of each iteration. The results of these analyses are summarized in table 3.

Table 3. The stability of DEA results in regard to outlier schools.

	Number of efficient schools	The range in Spearman rank correlation coefficient		Mean efficiency	Iterated mean efficiency	Standard deviation of means
		min	max			
Model 1 (crs)	4	0.98	0.9999	66.26*	67.88*	3.23
Model 1 (vrs)	17	0.96	1.00	72.23	72.45	0.86
Model 2 (crs)	9	0.96	1.00	77.48	77.99	0.87
Model 2 (vrs)	22	0.88	1.00	80.00	80.28	0.75
Model 3 (crs)	29	0.97	1.00	81.89	82.02	0.56
Model 3 (vrs)	48	0.92	1.00	84.12	84.22	0.48
Model 4 (crs)	61	0.96	1.00	91.30	91.33	0.19
Model 4 (vrs)	97	0.94	1.00	93.73	93.74	0.11

*The means are not equal (F-test, 5 % significance level)

The results of rank correlation coefficients show that the rankings[10] are relatively stable in

[9] Also Bonesrønning and Rattsø (1994) and Färe *et al.* (1989) used jackknifing in their analysis. The difference to our analysis is that they dropped each school one at the time as we dropped only the efficient units that construct the frontier.

[10] The value of 1 of Spearman rank correlation coefficient indicates that the rankings are exactly the same. The

regard to outlier schools determining the efficiency frontier. In case of CRS, the variation of rank correlation coefficient was lowest in model 1 ranging from 0.98 to 1.00. In other models the rank correlation varied somewhat more but ranged still at most from 0.96 to 1.00 in models 2 and 4. In case of VRS, more variation occurred. The largest differences were in model 2, where the rank correlation coefficient ranged from 0.88 to 1.00. The smallest differences were also in this case in model 1 ranging from 0.96 to 1.00. Thus, relative to CRS, the VRS frontier and efficiency scores were somewhat more sensitive to outliers.

DEA efficiency scores seem to be the more stabile with respect to outlier schools the greater is the number of inputs and outputs in the model. This is especially so under the assumption of CRS. In model 1 under CRS the iterated mean efficiencies differed significantly (F-test; 5 % level) from the mean efficiency based on DEA with all the schools included. The difference decreased as the number of inputs and outputs increased and it was statistically insignificant in the rest of the CRS models. In case of VRS, the difference between and the mean efficiency with all the schools in analysis was statistically insignificant in all the models. Also the standard deviation of the means of iterated runs decreased as the number of inputs and outputs increased.

5.2.2 Stability of rankings between the models

Besides the changes in the efficiency score distributions the schools may also change their ranking based on efficiency score from one model to another. We studied these changes by looking at the Spearman rank correlation coefficients between all the four models.

In table 4 there are the Spearman rank correlation coefficients between the different models when CRS was assumed. The correlation coefficient was 0,84 between efficiency rankings of models 1 and 2. There were, in other words, rather small differences in the efficiency rankings between these two models as the correlation coefficient of 1 means that the rankings are exactly the same. The rankings between model 2 and 3 where somewhat more similar as the correlation coefficient was 0.86 whereas the rankings from model 3 to model 4 changed quite a lot as the coefficient was only 0.66. The largest differences in the efficiency rankings were

---

value of 0 indicates no relation between the rankings and the value of -1 reverse rankings.

unsurprisingly between models 1 and 4 as the respective correlation coefficient was only 0.44.

Table 4. Spearman rank correlation coefficients between models 1-4 (based on efficiency scores assuming CRS)

	Model 1	Model 2	Model 3	Model 4
Model 1	1.0			
Model 2	0.84	1.0		
Model 3	0.73	0.86	1.0	
Model 4	0.44	0.54	0.66	1.0

In table 5 the Spearman correlation coefficients between the rankings of different models when VRS was assumed are presented. These results are very similar in both pattern and size as those based on CRS assumption.

Table 5. Spearman rank correlation coefficients between models 1-4 (based on efficiency scores assuming VRS)

	Model 1	Model 2	Model 3	Model 4
Model 1	1.0			
Model 2	0.83	1.0		
Model 3	0.71	0.87	1.0	
Model 4	0.45	0.60	0.69	1.0

To get a more intuitively appealing picture of what happens to the rankings we also grouped schools into quartiles according to efficiency scores from each model. This makes it possible to study whether the schools remain in the same quartile or change their position from one model to another. In Tables 6 (CRS) and 7 (VRS) we have cross-tabulated the number of schools in each quartile based on efficiency scores from models 1 and 3. If the rankings would remain the same the off diagonal elements of the table would have no observations. This, however, is not quite true.

According to table 6 (CRS) out of 291 schools some 56 per cent (163 schools) stayed in the same quartile (diagonal elements) and the rest 128 schools (off-diagonal) changed quartile when results of model 3 were used instead of model 1. The ranking remains more stable in the tails in the sense that some 70 per cent of those schools that were either in the lowest or in the highest quartile in model 1 remained there also in model 3. Less than half (43 per cent) of those schools that were either in the second or third quartile in the model 1 were in the same quartile in model 3. The rest of the schools changed their quartile and there were also schools which moved from one end of the distribution to the other end. This is the case especially for six schools which according to model 1 were in the first quartile but moved to the fourth quartile in model 3. This kind of result indicates that at least for some schools it may make a big difference what is included as inputs and outputs in evaluating their efficiency. Note, however, that the risk of missevaluating seems to be asymmetric as no school in the lowest quartile according to model 1 moved to first or second quartiles when more variables in model 3 were used.

Table 6. Frequencies of schools in quartiles according to their efficiency scores in model 1 and model 3 assuming CRS (the range of efficiency scores in each quartile in parentheses).

Model 3/ Model 1	**I quartile** (38.14-70.96)	**II quartile** (71.07-77.37)	**III quartile** (77.49-83.45)	**IV quartile** (83.47-100)	**Total**
I quartile (16.99-59.56)	51 (17.53)	11 (3.78)	5 (1.72)	6 (2.06)	73 (25.09)
II quartile (59.58-65.10)	18 (6.19)	33 (11.34)	16 (5.50)	6 (2.06)	73 (25.09)
III quartile (65.14-73.57)	4 (1.37)	28 (9.62)	30 (10.31)	11 (3.78)	73 (25.09)
IV quartile (73.58-100)	0 (0.00)	1 (0.34)	22 (7.56)	49 (16.84)	72 (24.74)
Total	73 (25.09)	73 (25.09)	73 (25.09)	72 (24.74)	291 (100.00)

Table 7. Frequencies of schools in quartiles according to their efficiency scores in model 1 and model 3 assuming VRS (the range of efficiency scores in each quartile in parentheses).

Model 3/ Model 1	**I quartile** (58.40-75.78)	**II quartile** (75.83-82.57)	**III quartile** (82.64-93.54)	**IV quartile** (93.95-100)	**Total**
I quartile (40.61-64.09)	45 (15.46)	16 (5.50)	4 (1.37)	8 (2.75)	73 (25.09)
II quartile (64.18-69.84)	25 (8.59)	27 (9.28)	12 (4.12)	9 (3.09)	73 (25.09)
III quartile (69.91-80.26)	3 (1.03)	30 (10.31)	32 (11.00)	8 (2.75)	73 (25.09)
IV quartile (80.62-100)	0 (0.00)	0 (0.00)	25 (8.59)	47 (16.15)	72 (24.74)
Total	73 (25.09)	73 (25.09)	73 (25.09)	72 (24.74)	291 (100.00)

Under the assumption of VRS there seems to be somewhat more instability as compared to the above CRS results. According table 7 (VRS) out of 291 schools some 52 per cent (151 schools) stayed in the same quartile (diagonal elements) and the rest 140 schools (off-diagonal) changed quartile when results of model 3 were used instead of model 1. The pattern of changes in rankings is very similar to those obtained assuming CRS.

Similar comparisons could be made between any two pairs of models. Instead of presenting and commenting them all, we feel that Tables 6 and 7 comparing model 1 with a limited list of outputs and inputs, and a more extensive model 3 are sufficient to illustrate how the ranking of schools depends on variables available.

5.3 On the connections of efficiency scores to some other variables

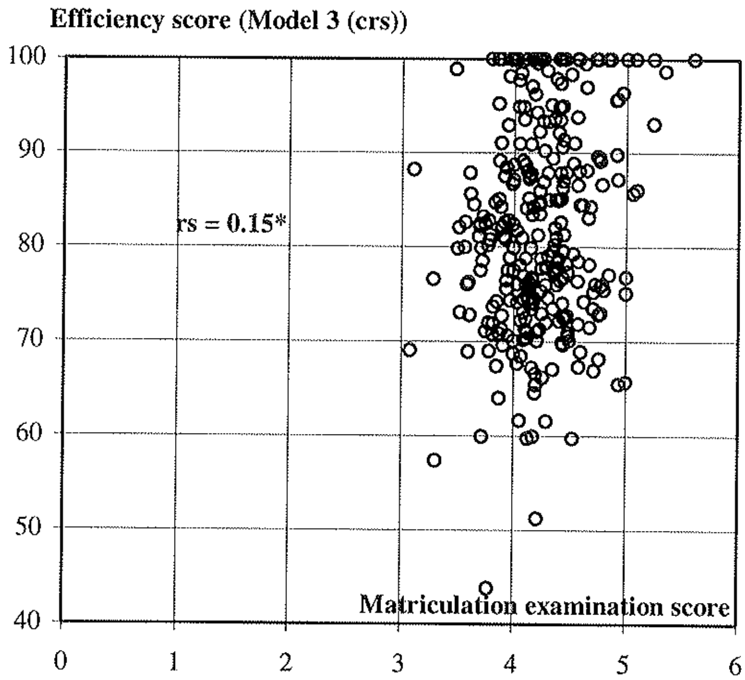
Above we presented results on the efficiency score distributions of the alternative models and studied their stability. Here, we consider how efficiency is related (or unrelated) to some of variables of interest. In section 5.3.1, to illustrate that efficiency as a ratio of weighted outputs and weighted inputs is a different concept than e.g. any typical output indicator, we present scatter diagrams where efficiency scores and matriculation examination scores are plotted. Thereafter, we present scatter diagrams in which school size and class size is plotted with efficiency scores. In section 5.3.2 we present Tobit models in which efficiency scores derived from the most extensive models are explained by variables not included in the DEA models.

5.3.1 Scatter diagrams

In the introduction we noted that schools can be evaluated from various points of view paying attention to outputs (students' achievements etc.) or inputs (expenditures etc.) only. Here, we have been interested in efficiency which has to do with the relation between outputs and inputs. Obviously, if the view-point differs also the evaluation and rankings may change drastically. To see that this is the case, we have plotted (figure 4) the efficiency scores of Model 3 (assuming CRS) and one output indicator, namely average matriculation examination score which is most often regarded as an indicator of the goodness of senior secondary schools.

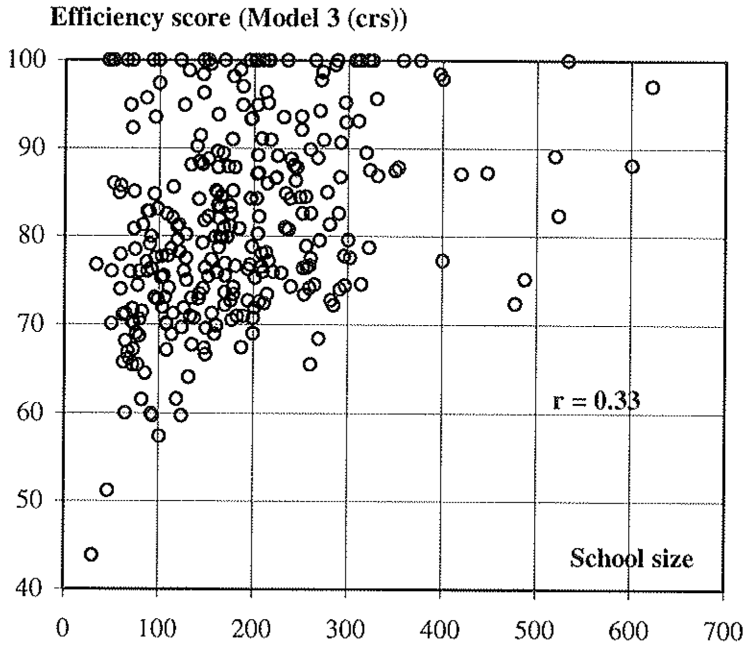
We see from figure 4 that efficiency scores and the average grades in matriculation examination are almost unrelated (Spearman rank correlation coefficient 0.15). A school with poor matriculation examination results can be efficient because it may also have a small amount of inputs. On the other hand, good performing schools by matriculation examination scores can be inefficient since their use of inputs is large.

Figure 4. Scatter diagram for efficiency scores (from model 3 under CRS) and matriculation examination scores.



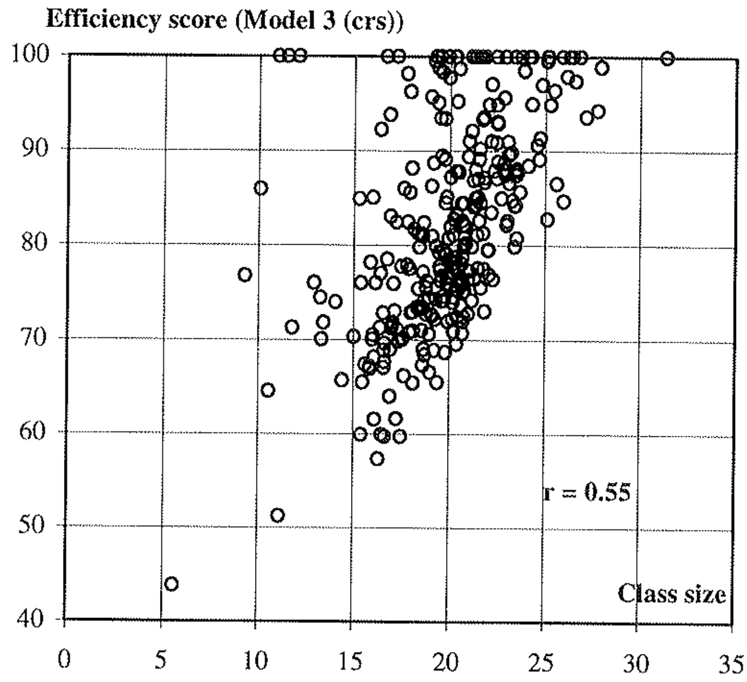
In figure 5 we have plotted efficiency scores and school size measured by the number of pupils. Their correlation is 0.33. The least one can say about the figure is that inefficiency seems to be more related to small schools than big schools although the efficiency varies practically in all school sizes. There are small schools with less than 100 pupils which are fully efficient (score 100) and large schools with over 400 students that are quite inefficient.

Figure 5. Scatter diagram for efficiency scores (from model 3 under CRS) and school size (measured by number of pupils).



The relation between class size and efficiency score is shown in figure 6. The average class size in the picture is a proxy and it is calculated by first dividing the number of teaching hours in the school by 20 (our estimate of the average teaching hours of each teacher per week) and then dividing the obtained number with the number of pupils. The correlation between the class size and efficiency scores is 0.55 suggesting that schools with bigger classes are more efficient. Despite this general relation, one must note again that there are some fully efficient schools with small class sizes.

Figure 6. Scatter diagram of efficiency scores (from model 3 under CRS) and class size.



5.3.2 A Tobit model

Here, we shall present results of a statistical exercise where efficiency differences are explained by some variables not directly included in the DEA analysis. As efficiency scores are limited to range to the [0,1] interval one must use methods that take this fact into account.

In a standard Tobit model the dependent variable is either zero or some positive number (c.f. e.g. Maddala 1983) so that only negative values are excluded. By choosing the dependent variable Y_i for school i to be the inefficiency score (1-efficiency score), the model can be written as

$$\begin{aligned} Y_i^* &= X_i\beta + \mu_i, \\ Y_i &= Y_i^*, \textit{if } Y_i^* > 0, \\ Y_i &= 0, \textit{otherwise.} \end{aligned} \qquad (11)$$

In (11) X_i is a vector of explanatory variables and β is a vector of parameters to be estimated. Y_i* is a latent variable which can be viewed as a threshold beyond which the explanatory variables must affect in order for Y_i to "jump" from 0 to some positive value. In our case, the inefficiency score is meaningful to view as a continuous variable limited to a minimum value of 0. Thus, the threshold has no special interpretation in our case, but the model specification makes it possible to estimate the model by the maximum likelihood method assuming normally distributed errors μ_i.

This approach with Tobit model as a second step after deriving efficiency scores by DEA has been used e.g. by McCarthy and Yaisawarng (1993) in a school context as pointed out earlier, and also by Luoma *et al.* (1995) in studying the efficiency of health centers in Finland. In our case we simply want to test whether some variables related to the schools or their environment have explanatory power for the efficiency differences. We shall not use the results here e.g. to calculate efficiency scores corrected by variables related to school environment as McCarthy and Yaisawarng did.

The inefficiency scores which are the dependent variables in the subsequent Tobit models will be based on models 3 and 4 assuming CRS. Recall that in model 3 we did not have parents' education (PARENTS' EDU) as an input so it is interesting to test weather the inefficiency scores of model 3 can be explained by this factor (in tobit models it is an average educational level of student's parents multiplied by 100) using a Tobit model.

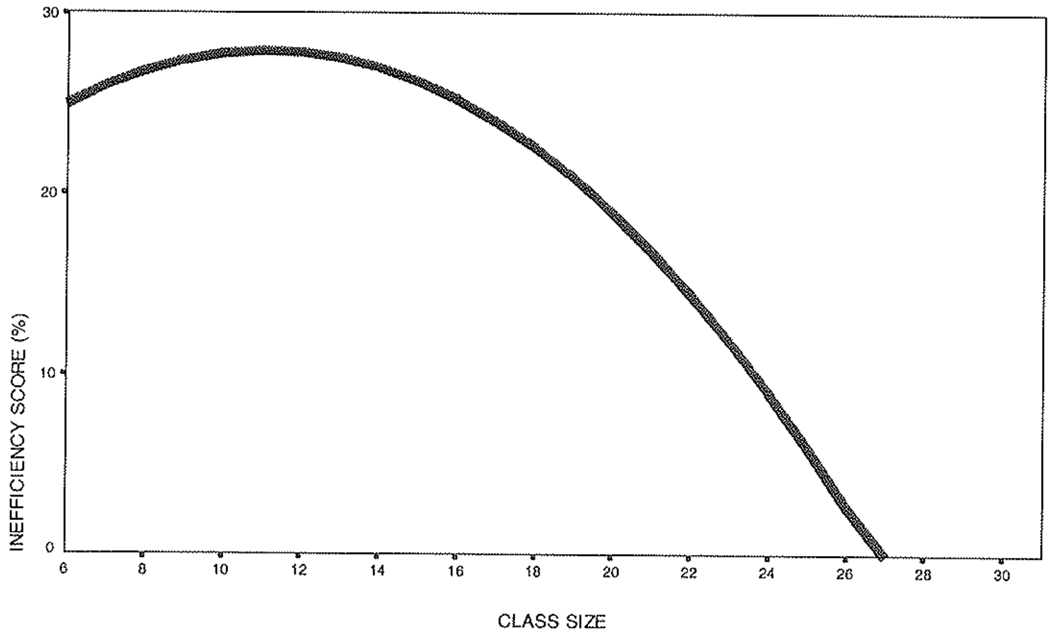
The differences between the results from CRS and VRS models are an indication of scale inefficiency. Instead of reporting results on measures of scale inefficiency (see section 2) e.g. by decomposing VRS efficiency scores into technical and scale efficiency, we use school size (SCHOOL SIZE) and class size (CLASS SIZE) in our Tobit model to explain inefficiency scores from CRS models. If there is a school size or a class size (measured by number of pupils) that is efficient, inefficient schools have smaller and larger class size than this optimal size. To test this, in addition to first order terms, we also included second order terms (SCHOOL SIZE SQ. AND CLASS SIZE SQ.) of these variables.

We also have information on sector specific state grants to municipalities. The grant system in 1989-91 was based on a "per cent of administratively determined acceptable educational expenditures" type formula where the per cent (GRANT RATIO1) varied according to economic, demographic and regional characteristics of municipalities ranging from 50 to 86 per cent. Together with this variable we use the ratio of actual to acceptable expenditures (ACT/ACC EXP.). Instead of these two variables we also use the ratio of grants to actual expenditures in 1992 (GRANT RATIO2) to explain inefficiency. The range of this ratio was from 0.408 to 1.006 in 1992. If the grant ratios are high, one would expect that the incentive for efficient use of resources is low leading to inefficiency.

There are 16 private schools (getting grants as public schools) in our data and we have a dummy variable (PRIVATE) to test its relation to (in)efficiency. As explanatory variables we shall also use the share of female students in the school (FEMALE). To study the role of heterogeneity of students, we have calculated for each school the standard deviation of students' mean grades (grades range 4-10) given by their own teachers at the end of senior secondary school (HETEROGENEITY.). We also have two dummy variables which are related to the type of municipality (URBAN; intermediate densely populated area = DENSE; and sparsely populated country-side as the reference case). As the generousness of the grant system is related to the type of municipality, the use of these dummy variables is an alternative to our GRANT RATIO variables. The summary statistics of the above variables are in Appendix 2.

The results of three Tobit models explaining the inefficiency scores of CRS model 3 are in table 7. As for the effects of individual explanatory variables, school size, share of female students, heterogeneity of students in the school, type of municipality (urban, dense, rural) and variables related to the grant system remain clearly insignificant in explaining inefficiency. Class size is related to inefficiency nonlinearly as both the first and second order terms become significant implying (according to Tobit model 3A) a relation depicted in Figure 7 in the range of class size in our data. According to these results inefficiency first increases as the class size increases. After the class size is on average 11 students inefficiency starts to decrease as class size increases and when the class size is on average 27 students inefficiency is minimized.

Figure 7. The relation between inefficiency and class size (from Tobit model 3A).



Somewhat surprisingly private schools are less efficient than public schools. Parents' educational level has a clear positive effect on efficiency.

Table 7. Parameter estimates (normalized coefficients) of Tobit models explaining school inefficiency (model 3 assuming CRS)

	Tobit 3A		Tobit 3B		Tobit 3C	
Variable	**Coeff.**	**t-ratio**	**Coeff.**	**t-ratio**	**Coeff.**	**t-ratio**
Constant	2.34	1.59	2.49	1.53	0.92	0.47
SCHOOL SIZE	0.0006	0.22	0.0016	0.68	0.0016	0.69
SCHOOL SIZE SQ.	-0.000003	-0.71	-0.000004	-1.00	-0.000004	-0.97
CLASS SIZE	0.26	2.04	0.25	1.95	0.25	1.96
CLASS SIZE SQ.	-0.012	-3.58	-0.012	-3.48	-0.011	-3.44
PRIVATE	0.63	2.03	0.66	2.15	0.72	2.30
FEMALE	0.0008	0.12	0.0005	0.07	-0.00004	0.01
HETEROGENEITY	0.66	1.24	0.66	1.25	0.70	1.32
PARENTS' EDU	-0.0056	-3.19	-0.0051	-2.81	-0.0047	-2.55
URBAN	0.26	1.33				
DENSE	0.11	0.58				
GRANT RATIO1					0.0023	0.27
ACT/ACC. EXP					0.99	1.05
GRANT RATIO2			-0.0027	-0.37		
Proportion of efficient schools	29/291					
R^2	0.361		0.358		0.362	
Log-likelihood	211.82		210.98		211.51	

Next, we shall take the inefficiency scores from model 4 (assuming CRS), i.e. a model where parents' education level is treated as an input, and explain the differences using Tobit models. The results are presented in table 9. Note that the number of schools that are fully efficient (inefficiency score 0) is greater and the average inefficiency is lower here as compared to results of model 3 (assuming CRS). Thus it is not surprising that the R^2 values are lower here than in the models in Table 8.

When parents' educational level has been included already in DEA-model and it is not an explanatory variable in Tobit models, there are some changes in the effects of variables. School size remains insignificant and class size significant as before. Type of municipality becomes significant such that schools in urban areas are inefficient relative to those in rural and other less populated areas. Also student composition begins to matter. Although the share of female students does not quite get a significant coefficient, heterogeneity in grades is positively related to inefficiency. When grant variables are used instead of municipality type variables they, surprisingly, begin to get negative coefficients which are close to being significant. When the insignificant school size variables are excluded from the model, their (asymptotic) t-statistics increase in absolute value (1.94 for GRANT RATIO1 and 1.99 for GRANT RATIO2). The same happens to the dummy variable related to private schools making the dummy significant also in Tobit model 4A. These results do not support the hypothesis that high grant ratios and being a public school lead to inefficiency. Rather, they somewhat unexpectedly support, if anything, the opposite conclusion.

Table 9. Parameter estimates (normalized coefficients) of Tobit models explaining school inefficiency (from model 4 assuming CRS).

	Tobit 4A		Tobit 4B		Tobit 4C	
Variable	**Coeff.**	**t-ratio**	**Coeff.**	**t-ratio**	**Coeff.**	**t-ratio**
Constant	-4.90	-2.94	-3.51	-2.08	-3.40	-1.68
SCHOOL SIZE	0.00009	0.04	0.0022	0.95	0.0023	0.97
SCHOOL SIZE SQ.	-0.000002	-0.57	-0.000004	-1.07	-0.000005	-1.12
CLASS SIZE	0.66	4.10	0.61	3.86	0.61	3.84
CLASS SIZE SQ.	-0.021	-5.09	-0.020	-4.86	-0.020	-4.85
PRIVATE	0.55	1.85	0.65	2.14	0.62	1.99
FEMALE	-0.0097	-1.44	-0.011	-1.58	-0.011	-1.53
HETEROGENEITY	1.77	3.27	1.73	3.18	1.72	3.15
URBAN	0.63	3.32				
DENSE	0.16	0.86				
GRANT RATIO1					-0.014	-1.80
ACT/ACC. EXP					0.063	0.07
GRANT RATIO2			-0.012	-1.81		
Proportion of efficient schools	61/291					
R^2	0.253		0.229		0.229	
Log-likelihood	214.34		209.69		209.66	

6 CONCLUSIONS

In this paper we studied the efficiency differences among 291 Finnish senior secondary schools. The relations between outputs and inputs were analyzed by Data Envelopment Analysis (DEA). Four different models with school level data were used. In the simplest one both input and output variables were quantitative, in the most extensive ones also quantified qualitative variables were included. They were related to students who were in the (typically) three year school in 1989-91 and participated in the national matriculation examinations in the spring of 1991. Output variables in our cross-section data included number of graduates who passed their grade during 1989-91, number of students who passed the matriculation examinations in Spring 1991 and related scores in compulsory and additional subjects. Input variables included teaching and other hours, experience and education of teachers, and quality of pupils when entering the school. As an input outside the schools control we had education of students' parents.

DEA is a non-parametric method which derives from input-output data a frontier production function (efficiency frontier) by linear programming. Efficient schools which are part of the frontier get an efficiency score of 100 per cent and inefficient ones located under the efficiency frontier scores below 100 per cent indicating the share of resources with which the current outputs could be produced if they were efficient. Average efficiencies in the most extensive models were 82-84 per cent, ranging from 44 to 100 per cent. When in addition to inputs related to the schools, parents' educational level was treated as an input, average efficiency increased to 91 per cent. The results depend somewhat on whether one assumes that constant returns to scale (CRS) or variable returns to scale (VRS) applies to the frontier.

In addition to finding out efficiency differences among senior secondary schools, one purpose of the study was to analyze how efficiency rankings of schools depend on inputs and outputs included in DEA. Will a ranking based on only a few readily available variables produce similar rankings as one in which there are more input and output variables. By ranking schools according to their efficiency scores, 60-70 per cent of schools in the topmost and the lowest efficiency quartiles tended to maintain their ranking from the simplest model 1 to model 3 where all of our input and output variables (except parents' education) were included. The schools in the centrally located quartiles 2 and 3 were more mobile as only 40 - 45 per cent of them stayed in the initial quartile. Stability of the rankings from one model to another was also studied by calculating Spearman rank correlation coefficients. These results showed that the largest differences in rankings in subsequent models were between the models 3 and 4 where the correlation was 0.66 assuming CRS and 0.69 assuming VRS.

As another test of stability of our results we applied so called jackknifing. We run DEA analysis by dropping one fully efficient (i.e. frontier) observation at a time to see how the omittance of these eventual outlier observations would affect the results. Changes in efficiency scores and rankings (according to Spearman rank correlation coefficients) turned out to be minor.

The results also showed that e.g. the rankings of schools by matriculation examination scores differed markedly from their rankings by efficiency. This is not surprising as efficiency measures the relation between outputs and inputs whereas matriculation examination scores are output indicators. Efficiency scores were, however, positively related to class size (correlation 0.55) and to a lesser extent to school size (0.33).

In addition to presenting scatter diagrams and correlations between efficiency scores and some individual variables of interest, we also did some statistical modeling. The degree of inefficiency (100-efficiency score) of schools was explained by a Tobit model. Here, the explanatory variables included factors which were related to the student body, scale of operation, and financing of the school but were not included in the DEA models. As for the effects of individual explanatory variables, school size, share of female students, heterogeneity of students in the school, type of municipality (urban, dense, rural) and variables related to the grant system remained clearly insignificant in explaining inefficiency. On the other hand, inefficiency decreased as class size, and parents' educational level increased. Somewhat surprisingly, private schools were inefficient relative to public schools.

When parents' education level was included as an input in the DEA analysis and the inefficiency scores were then explained in a Tobit model, some results changed. School size remained insignificant and class size significant as before. Municipality type becomes significant such that schools in urban areas are inefficient relative to those in rural and other less populated areas. Also student body composition begins to matter. Although the share of female students does not quite get significant coefficient, heterogeneity in grades is positively related to inefficiency. When grant variables are used instead of municipality type variables they, surprisingly, begin to get negative coefficients which are close to being significant. The same happens to the dummy variable related to private schools making the dummy significant. These results do not support the hypothesis that high grant ratios and being a public school lead to inefficiency. Rather, they somewhat unexpectedly support, if anything, the opposite conclusion.

REFERENCES

Banker, R. D., Charnes, A. and Cooper, W. W. (1984) Some Models For Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science* **30**, 1078-1092.

Barrow, M. M. (1991) Measuring Local Education Authority Performance: A Frontier Approach. *Economics of Education Review* **10**, 19-27.

Bessent, A. and Bessent, W. (1980) Determining the Comparative Efficiency of Schools through Data Envelopment Analysis. *Educational Administration Quarterly* **16**, 57-75.

Bessent, A., Bessent. W., Charnes, A., Cooper, W. and Thorogood, N. (1983) Evaluation of Educational Program Proposals by Means of DEA. *Educational Administration Quarterly* **19**, 82-107.

Bessent, A., Bessent, W., Elam, J. and Long, D. (1984) Educational Productivity Council Employs Management Science Methods to Improve Educational Quality. *Interfaces* **14**, 1-8.

Bessent, A., Bessent, W., Kennington, J. and Reagan, B. (1982) An Application of Mathematical Programming to Assess Productivity in the Houston Independent School District. *Management Science* **28**, 1355-1367.

Bonesrønning, H. and Rattsø J. (1994) Efficiency Variation Among the Norwegian High Schools: Consequences of Equalization Policy. *Economics of Education Review* **13**, 289-304.

Bonesrønning, H. and Rattsø J. (1992) Effektivitetsforskjeller i Videregående Skole: Analyse av Almenfaglig Studieretning i 34 Skoler. *Norsk Økonomisk Tidskrift* (NØT) **106**, 211-242.

Charnes, A., Cooper, W. W. and Rhodes, E. (1978) Measuring the Efficiency of Decision Making Units. *European Journal of Operations Research* **2**, 429-444.

Charnes, A., Cooper, W. W. and Rhodes, E. (1981) Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through. *Management Science* **27**, 668-697.

Fried, Harold, Knox C. A. Lovell and Shelton S. Schmidt, eds. (1993) *The Measurement of Productive Efficiency. Techniques and Applications.* New York: Oxford University Press.

Färe, R., Grosskopf, S. and Weber, W. (1989) Measuring School District Performance. *Public Finance Quarterly* **17**, 409-428.

Hanushek, E. A. (1986) The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature* **26**, 1141-1177.

Hanushek, E. A. (1979) Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources* **14**, 351-388.

Hanushek, E. A. and Taylor, L. L. (1990) Alternative Assessment of the Performance of Schools. *The Journal of Human Resources* **25,** 179-201.

Jesson, D., Mayston, D. and Smith, P. (1987) Performance Assessment in the Education Sector: Educational and Economic Perspectives. *Oxford Review of Education* **13**, 249-266.

Johnes, J. and Johnes, G. (1995) Research Funding and Performance in U.K. University Departments of Economics: A Frontier Analysis. *Economics of Education Review* **14**, 301-314.

Kirjavainen, T. and Loikkanen, H. A. (1993) On Efficiency Differences of Finnish Senior Secondary Schools - An Application of DEA (in Finnish). *VATT-research reports* no. 16. Helsinki: Government Institute for Economic Research.

Ludwin, W. and Guthrie, T. (1989) Assessing Productivity with Data Envelopment Analysis. *Public Productivity Review* **12**, 361-372.

Luoma, K., Järviö, M.-L., Suoniemi, I. and Hjerppe R. (1995) Productive Efficiency and Determinants of Inefficiency at Health Centres in Finland. *VATT-discussion papers* no. 94. Helsinki: Government Institute for Economic Research.

Maddala, G.S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics.* Econometric Society Monographs No. 3. Cambridge University Press. Cambridge.

Mayston, D. and Jesson, D. (1988) Developing Models of Educational Accountability. *Oxford*

Review of Education **14**, 321-339.

McCarty, T. A. and Yaisawarng, S. (1993) Technical Efficiency in New Jersey School Districts. In *The Measurement of Productive Efficiency. Techniques and Applications.* (Edited by Fried, H. and Lovell, K. C. A. and Schmidt, S. S.). New York: Oxford University Press.

Ray, S. C.. Resource-use Efficiency in Public Schools: A Study of Connecticut Data. *Management Science* **37**, 1620-1628.

Sengupta, J. K. and Sfeir, R. (1986) Production Frontier Estimates of Scale in Public Schools in California. *Economics of Education Review* **5**, 297-307.

Smith, P. and Mayston, D. Measuring Efficiency in the Public Sector. *OMEGA International Journal of Management Science* **15**, 181-189.

Appendix 1. Summary statistics of the variables used in DEA analyses

Inputs:	**mean**	**std. dev.**	**min.**	**max.**
Teaching hours per week	182.2	89.2	43.8	567.0
Other than teaching hours per week	21.0	8.8	3.50	67.6
Experience of teachers	54.4	26.5	2.7	165.3
Education of teachers	10.8	4.3	3.0	32.7
Admission level	441.9	249.9	60.95	1548.5
Education of students' parents	166.9	112.7	18.0	894.6
Outputs:				
Number of students who passed their grade (were moved up)	178.3	95.8	27.7	601.7
Number of graduates	50.0	28.4	7.0	177.0
Score in compulsory subjects in matriculation examination	891.4	520.4	140.0	3530.0
Score in other subjects in matriculation examination	276.7	173.2	42.0	1258.0

Appendix 2. Summary statistics of the variables used in Tobit analyses

	mean	**std. dev.**	**min.**	**max.**
PARENTS' EDU	301.5	48.6	200	531.6
SCHOOL SIZE	186	99.3	31	623
CLASS SIZE	19.9	3.32	5.6	31.4
PRIVATE	0.05	0.22	0	1
FEMALE	59.1	9.59	29.4	87.5
HETEROGENEITY	0.90	0.11	0.52	1.30
URBAN	0.43	0.50	0	1
DENSE	0.19	0.39	0	1
GRANT RATIO1	71.0	11.1	40.8	100.6
ACT/ACC. EXP.	1.025	0.06	0.75	1.26
GRANT RATIO2	72.5	10.9	50.0	86.0

ELINKEINOELÄMÄN TUTKIMUSLAITOS (ETLA)
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
LÖNNROTINKATU 4 B, FIN-00120 HELSINKI

Puh./Tel. (09) 609 900
Int. 358-9-609 900

Telefax (09) 601753
Int. 358-9-601 753

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

No 541 RITA ASPLUND, The Gender Wage Gap in Finnish Industry in 1980-1994. An empirical analysis of non-manual workers. 11.10.1995. 28 p.

No 542 TOR ERIKSSON - SUSANNA FELLMAN, Determinants of Firms' Operating Times - Some Evidence from Firm-Level Data. 23.10.1995. 19 p.

No 543 TARMO VALKONEN, Corporate and Capital Income Tax Reform in a Numerical Overlapping Generations Model: The Case of Finland. 12.12.1995. 28 p.

No 544 REIJA LILJA, Career Mobility in Finnish Industry. 14.12.1995. 22 p.

No 545 JUKKA LASSILA - TARMO VALKONEN, Policy Credibility in Numerical Overlapping Generations Models. 28.12.1995. 28 p.

No 546 EIJA KAUPPI - JUKKA LASSILA - TIMO TERÄSVIRTA, Short-Term Forecasting of Industrial Production with Business Survey Data: Experience from Finland's Great Depression. 26.01.1996. 20 p.

No 547 JIAN-GUANG SHEN, FDI, Knowledge Spillover and Economic Growth in East Asia. 31.01.1996. 70 p.

No 548 MAARIT SÄYNEVIRTA - PEKKA YLÄ-ANTTILA, Integraation ja yritysten kansainvälistyminen - vaikutuksia kotimaan talouteen. 01.02.1996. 58 s.

No 549 LAURA PAIJA - PEKKA YLÄ-ANTTILA, The Impact of Structure and Competition on Employment in the Telecommunications Cluster Case Finland. 12.02.1996. 34 p.

No 550 PENTTI VARTIA - PEKKA YLÄ-ANTTILA, Technology Policy and Industrial Clusters in a Small Open Economy - The Case of Finland. 16.02.1996. 15 p.

No 551 PONTUS BRAUNERHJELM - PER HEUM - PEKKA YLÄ-ANTTILA, Internationalization of Industrial Firms. Implications for growth and industrial structure in the Nordic countries. 16.02.1996. 33 p.

No 552 REIJO MANKINEN, Alkoholiveron alentamisen kansantaloudellisia vaikutuksia. 26.02.1996. 42 s.

No 553 RITA ASPLUND, Koulutus, työura ja palkkaerot. 22.03.1996. 13 s.

No 554 MARIANNE PAASI, The Absorptive Capacities of Estonian Firms. - Can a Technology-based Industrial Strategy Succeed? 22.03.1996. 17 p.

No 555 HANNU HERNESNIEMI, Barriers to Economic Cooperation of Baltic Rim Countries. 10.04.1996. 46 p.

No 556 ANNICK LARUELLE - MIKA WIDGRÉN, Is the Allocation of Voting Power among the EU States Fair? 17.04.1996. 19 p.

No 557 JARI HYVÄRINEN, A Survey of Corporate Governance - Which Model for Transition Countries? 13.05.1996. 32 p.

No 558 PASI KUOPPAMÄKI, Joint Implementation ilmastopolitiikan välineenä: Suomi ja lähialueiden kasvihuonekaasujen päästöjen rajoittaminen. 12.06.1996. 35 s.

No 559 MIKA PAJARINEN, Työnantajan kansaneläkemaksun porrastuksen toimivuus vuosina 1993 ja 1994. 20.6.1996. 34 s.

No 560 OKKO-PEKKA SALMIMIES, EU:n pakolais- ja maahanmuuttoasioita koskevan yhteistyön kehittäminen ja taloudellinen ulottuvuus. 10.07.1996. 32 s.

No 561 AIJA LEIPONEN, Education, Tenure and Innovation in Manufacturing firms. 16.08.1996. 26 p.

No 562 AIJA LEIPONEN, Education and Innovative Capabilities. 16.08.1996. 20 p.

No 563 AIJA LEIPONEN, Competences, Innovation and Profitability of Firms. 16.08.1996. 16 p.

No 564 JYRKI RUUTU, Suomen valuuttakurssijärjestelmä osana eurooppalaista järjestelmää. Historiallinen katsaus keskiajalta autonomian ajan loppuun. 29.08.1996. 22 s.

No 565 HEIKKI PALM, Eläkeuudistuksen vaikutukset Suomen kansantalouden numeerisessa limittäisten sukupolvien mallissa. 02.09.1996. 34 s.

No 566 JYRKI ALI-YRKKÖ, Teknologiaintensiivisten yritysten kansainvälistyminen - vaikutuksia Suomen talouteen. 05.09.1996. 53 s.

No 567 VENLA SIPILÄ, Suomen ja Venäjän välinen kauppa 1986-1995. 06.09.1996. 100 s.

No 568 EUGEN KOEV, Palkkadiskriminaatio teollisuuden toimihenkilöillä. 11.09.1996. 73 s.

No 569 MARIANNE PAASI, The Inherited and Emerging Absorptive Capacities of Firms - Results of a firms survey in the Estonian electronics industry. 24.09.1996. 26 p.

No 570 TANJA KIRJAVAINEN - HEIKKI A. LOIKKANEN, Efficiency Differences of Finnish Senior Secondary Schools: An Application of Dea and Tobit-analysis. 02.10.1996. 40 p.

Elinkeinoelämän Tutkimuslaitoksen julkaisemat "Keskusteluaiheet" ovat raportteja alustavista tutkimustuloksista ja väliraportteja tekeillä olevista tutkimuksista. Tässä sarjassa julkaistuja monisteita on mahdollista ostaa Taloustieto Oy:stä kopiointi- ja toimituskuluja vastaavaan hintaan.
Papers in this series are reports on preliminary research results and on studies in progress. They are sold by Taloustieto Oy for a nominal fee covering copying and postage costs.

d:\ratapalo\DP-julk.sam/02.10.1996