Tanja Kirjavainen

Essays on the Efficiency of Schools and Student Achievement

HELSINGIN KAUPPAKORKEAKOULU
HELSINKI SCHOOL OF ECONOMICS

A-348

Tanja Kirjavainen

Essays on the Efficiency of Schools and Student Achievement

HELSINKI SCHOOL OF ECONOMICS

ACTA UNIVERSITATIS OECONOMICAE HELSINGIENSIS

A-348

© Tanja Kirjavainen and
Helsinki School of Economics

ISSN 1237-556X
ISBN 978-952-488-336-8

E-version:
ISBN 978-952-488-337-5

Helsinki School of Economics -
HSE Print 2009

To my family.

Abstract

This thesis consists of five essays that investigate the efficiency differences of schools, explain the differences in efficiency and student achievement and take a further step by looking inside schools with differing efficiency. In the studies presented in the essays, efficiency differences were estimated with various methods, both parametric and non-parametric, and at different points in time using both cross section and panel data. The use of panel data, in particular, brings new possibilities to the analysis and understanding of efficiency differences. The stability of the efficiency estimations was also tested with several models. Information on this matter is a prerequisite if efficiency differences are to be studied further.

The causes of inefficiency were examined in this thesis with both quantitative and qualitative methods. The use of both of these methods broadened the scope of earlier studies and provided useful information on their advantages and disadvantages in analysing causes of efficiency. The application of different approaches also provided a richer description and understanding of the case schools, their processes and factors affecting efficiency.

Finnish general upper secondary schools provided a good empirical context to analyse efficiency differences between schools and factors affecting student achievement, because the schooling is fairly short, lasting on average for three years, information is available on the earlier school achievement of the students as well as their family background, and most importantly, it concludes with a nation-wide matriculation examination. The empirical analyses in this thesis are based on register data on Finnish general upper secondary schools, their students and students' parents in 1991, 1990-1998 and 2000-2004. The register data are supplemented with survey data from 73 schools in 1995 and 27 semi-structured interviews of principals and teachers in 9 schools in 1999.

The results of the studies show that there were efficiency differences between Finnish general upper secondary schools. The size of the differences depended on the method and model used for calculating the efficiency. The rankings of schools based on efficiency scores were also fairly unstable. Concerning organizational practices, very few of them were related to efficiency. The results of case studies showed that schools with caring views about the students, professional staff relations, participative management and decision making, and curriculum work which was perceived as a way to develop school and own work were more efficient.

In addition to investigating efficiency differences, the studies in this thesis examined the influence of the comprehensive school grade point average, family background, school resources, length of studies and decentralization of test taking on performance in the matriculation examination with different methods and data and at different points in time. The results provided causal information on the effect of changes in school resources on student achievement and they showed that student achievement was not affected by the changes in resources. New evidence was also provided on the effect of lengthening of studies and decentralization of test taking in the matriculation examination. According to results, schools with longer length of studies and higher rate of decentralization performed worse in the matriculation examination.

Key words: Efficiency, Productivity, Upper secondary education

JEL Classification: I21

Acknowledgements

When I started my university studies after taking the matriculation examination, I did not have a clue that it would take me this far and especially this time. Even though I have on many occasions had doubts about this journey, I am very happy to reach this point. There are many people that deserve to be mentioned for their efforts. First, I would like to express my gratitude to my supervisor and co-author, Professor Heikki Loikkanen. His encouragement, broadmindedness and patience have carried me through this process. Moreover, his character has made it easy for me to return to this thesis after many shorter and longer interruptions. I also highly appreciate the thoughtful comments throughout the process, not to mention the help in acquiring funding in the early stages of this work. There is no doubt that this thesis would not exist without Heikki's involvement.

The Unit for Organization and Management at the Helsinki School of Economics (HSE) provided me with an inspiring and tolerant environment to pursue my doctoral studies. I would especially like to mention Professor Keijo Räsänen, who has during the years provided assistance and insightful views about qualitative research. His gentle and determined guidance at crucial moments has been invaluable and kept me on the right track. I have also very much enjoyed our discussions about the nature of academic work. In addition, I would like to thank Professors Risto Tainio, Kari Lilja and Raimo Lovio for their encouragement during the whole process.

For my pre-examiners, Research Director Rita Asplund and Senior Lecturer Jill Johnes, I express my gratitude for their careful work and insightful comments and suggestions that clearly improved the manuscript. My other coauthors, Doctor Iida Häkkinen Skans and Research Director Roope Uusitalo, deserve special thanks for fruitful cooperation.

During the many years spent writing this thesis I have had the privilege to work in different research communities and government institutions. All of them have broadened my views about the multitude of research and deepened my understanding of the subject matter of this thesis.

For that I want to thank my colleagues in the Unit for Management and Organization at HSE, in the Unit for General Upper Secondary Education at the Finnish National Board of Education (OPH) and at the Government Institute for Economic Research (VATT). At VATT, I would especially like to thank Director General Seija Ilmakunnas and Research Director Aki Kangasharju for providing the opportunity to finish this thesis, Nina Intonen for her diligent preparation of the manuscript for the publication and Maija-Liisa Järviö for careful proof-reading of the final version.

Moreover, I would like to express my sincere gratitude to the principals and teachers acting as respondents in this thesis. Their willingness to participate has been of vital importance for this work. Financial support from the Academy of Finland 'Effectiveness of Schooling' programme, the Finnish Cultural Foundation and Gramis (Graduate School of Management and Information Systems) are gratefully acknowledged. Roy Siddall has kindly checked the language of the introductory essay and the third and fourth essays.

Life without good friends and colleagues would be miserable. Helena, Leena, Raimo, Susanna, Kimmo, Leo, Elina, Ria and many others, thank you for sharing the ups and downs of working life and life in general, both here in Finland and abroad. Naturally, I am very indebted to my parents, Ani and Jorma and my sister Katja for all they have done for me. My deepest gratitude is to my dear family, my husband Sami and our two children, Henri and Elina. I'm very fortunate to have you around. With great happiness and relief I can finally tell you, that 'the big book' is now ready.

Herttoniemi, June 2009

Tanja Kirjavainen

Table of Contents

ESSAYS AND ARTICLES

I. Kirjavainen, Tanja – Loikkainen, Heikki A. (1998): Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit-Analysis. *Economics of Education Review* 17, 377-394.*

II. Kirjavainen, Tanja: Efficiency of Finnish Senior Secondary Schools: Explaining High and Low Performance with Resources, Environmental, and Organizational Factors. An earlier version of this paper has been published as Kirjavainen, Tanja (1999): Efficiency of Finnish Senior Secondary Schools: Explaining High and Low Performance with Resources, Environmental, and Organizational Factors. Helsinki School of Economics and Business Administration, *Working papers W-221*, 1999.

III. Kirjavainen, Tanja (2007): Efficiency of Finnish Upper Secondary Schools: An Application of Stochastic Frontier Analysis with Panel Data. *VATT-discussion papers no. 428*. Helsinki: Government Institute for Economic Research.

IV. Kirjavainen, Tanja: Understanding Efficiency Differences of Schools: Practitioners' Views on Students, Staff Relations, School Management and Curriculum. An earlier version of this paper has been published as Kirjavainen, Tanja (2008): Understanding Efficiency Differences of Schools: Practitioners' Views on Students, Staff Relations, School Management and Curriculum. VATT-*discussion papers no. 450*. Helsinki: Government Institute for Economic Research.

V. Häkkinen, Iida – Kirjavainen, Tanja – Uusitalo, Roope (2003): School Resources and Student Achievement Revisited: New Evidence From Panel Data. *Economics of Education Review* 22, 329-335.**

*Reprinted with the permission of Elsevier Limited, Copyright © 1998 Elsevier B.V.

** Reprinted with the permission of Elsevier Limited, Copyright © 2003 Elsevier B.V.

INTRODUCTION

1. Background of the thesis

Education plays a major role in modern society and in peoples' lives. Children spend a considerable part of their early life in schooling, acquiring knowledge and skills that are useful later in their lives. Schooling does not end with the coming of age, but often continues through life, as many people change their occupation or complement their earlier studies. The concept of lifelong learning depicts this development.

Education not only occupies time but provides wealth, health and pleasure for people and nations. It is considered a key element in keeping up in the present day global competition. Therefore, governments invest considerable amounts of money in the education sector. From their point of view and from the point of view of taxpayers, it is not irrelevant how this spending is used. There are both equity and efficiency aspects that have to be taken into account.

Equity of schooling emphasizes the importance of providing equal opportunities for education regardless of students' family background, race or place of residence. From the point of view of efficiency, the focus is on delivering good quality education by minimizing the use of resources. To some extent, these two aspects can be viewed as being two opposite goals. However, taking into consideration the way efficiency is measured, the question of equity is also addressed. Since differences between educational institutions appear in efficiency analysis as efficiency differences, large differences indicate that there also might be problems with equity.

If increased efficiency and productivity is taken as one of the means to improve the functioning of educational sector, information is needed on efficiency measurement and differences as well as factors contributing to higher efficiency and productivity. There is a large body of research on efficiency measurement and efficiency differences in general (see Fried *et al.,* 2008). This also applies to the education sector (Johnes, 2004; Worthington, 1999). The Finnish schooling system has been analysed to a lesser extent (see Kangasharju, 2007).

As for studies examining the causes of efficiency, they are more limited in number, both in general and concerning education. In order to improve efficiency, such information is, however, essential. Not only is research evidence concerning the system level effects such as governance and competition important, but also information concerning the processes at the local and school levels.

As is often the case, there is also a more personal side. As a young researcher conducting my first efficiency analysis, I started wondering if my measurements had anything to do with 'real life'. I had these nice efficiency distributions that I enjoyed calculating and charts and figures that looked nice on paper. But were the figures reliable? What was behind them? What were these schools and what kind of school was actually efficient? These are, I suppose, obvious questions coming to many people's minds when conducting efficiency analysis with statistical data.

The essays of this thesis provide answers to these questions. They investigate the efficiency differences among schools, explain the differences in efficiency and student achievement and take a further step and look inside those schools with differing efficiency. Efficiency differences are estimated with various methods, both parametric and non-parametric, and at different points in time using both cross section and panel data. The use of panel data, in particular, brings new possibilities to the analysis and understanding of the efficiency differences. The stability of the efficiency estimations is also tested with several different models. Information on this matter is a prerequisite if efficiency differences are to be studied further.

Causes of inefficiency are examined in this thesis with both quantitative and qualitative methods. The use of both of these methods broadens the scope of earlier research and provides useful information on their advantages and disadvantages in addressing these questions. The application of different approaches also provides a richer description and understanding of the schools, their processes and factors affecting efficiency. As such, it is useful information for future studies.

In addition to investigating efficiency differences, the studies in this thesis examine the influence of the comprehensive school grade point average, family background, school resources, length of studies and decentralization of test taking on performance in the matriculation examination using different methods and data and at different points in time. The results provide causal information on the effect of changes in school resources on student achievement. New evidence is also provided about the effect of the lengthening of studies and decentralization of test taking in matriculation examination.

Finnish general upper secondary schooling[1], which provides post-compulsory general secondary education for students aged 16-19, offers a good empirical context to examine efficiency differences between schools and factors affecting student achievement because it is fairly short, lasting for three years on average, information is available on the students' earlier school achievements and family background, and most importantly, it concludes with a nationwide matriculation examination.

The empirical analyses in this thesis are based on register data on Finnish general upper secondary schools, their students and students' parents in 1991, 1990-1998 and 2000-2004. The register data are supplemented by one set of survey data from 73 schools in 1995 and 27 semi-structured interviews of principals and teachers in 9 schools in 1999.

[1] Two of the essays in this thesis use the term 'senior secondary school' to refer to general upper secondary schools. Earlier, this term was more common. The present recommendation is the term 'general upper secondary school'. Therefore, this term is used in the introduction and in three of the other essays in this thesis.

During the years covered in this thesis, general upper secondary schooling has undergone major changes. The reforms have increased school-level autonomy in decision making and in the content of instruction. They have also increased students' autonomy and initiative in studying. The effects of these changes have not been very intensely studied, especially from economic and efficiency perspectives. The results of this thesis also provide evidence of some of the consequences of these reforms.

This introduction continues as follows. Chapter 2 presents research related to this thesis. The studies in this thesis are based on two research traditions. First, I discuss studies on the education production function, which concentrate on examining the effect of peers, school resources and class size on student achievement. In the following chapter, the measurement of efficiency and different methods used in this thesis are reviewed. School effectiveness is then introduced and to a lesser extent school improvement studies that concentrate on the processes of schooling and their effects on student achievement. Chapter 3 briefly reviews the institutional context and describes the Finnish general upper secondary schooling system in 1985-2005. The content of the thesis is presented in chapter 4 and the main findings and conclusions in chapter 5.

2. Approaches to studying school efficiency

Interest in studying factors affecting student performance started at the end of the 1950s. The Coleman report (Coleman *et al.,* 1966) is usually mentioned as the first systematic work in the field trying to uncover factors behind successful student performance in public elementary and secondary schooling. Following the Coleman report, abundant research has been carried out on the effect of family background, peers and school resources on student achievement.

This thesis relies on two research traditions that both have their origins in the Coleman report. Economists have used education production functions when studying the effect of family background, peers and school resources on student achievement. Educationalists have concentrated more on the processes of schooling and their effect on student achievement in school effectiveness research. Literature on the education production function provides the general framework for efficiency measurement, since it deals with the inputs and outputs of schooling and their relationship. The focus in studies on the efficiency measurement of schooling is usually more limited and technical in nature, concentrating on developing and testing the results of different methods. The relationship of inputs and outputs in many cases plays minor part.

In the following, I present the main developments in studies on the education production function, discuss the measurement of efficiency and studies measuring the efficiency of schooling, as well as school effectiveness studies. Since the education production function and school effectiveness studies, in particular, have addressed a wide range of issues, I will be concentrating on those factors that are relevant to the studies in this thesis.

2.1 The education production function

The education production function relates observed student outcomes to characteristics of the students, their families and other students in the school, as well as other school characteristics. The basic structure of the education production function can be written as

$$O_g = f(F^{(g)}, P^{(g)}, C^{(g)}, T^{(g)}, S^{(g)}, \alpha), \tag{1}$$

where O_g is the achievement for a student in grade *g; F, P, C, T* and *S* represent vectors of family, peer, community, teacher and school inputs, respectively; α depicts ability; and the superscript *g* indicates that all of the inputs are cumulative from birth to grade *g*. In this model, achievement is cumulatively dependent on the various past inputs. As such, it creates many complications for empirical work, the most important being omitted variable bias. To avoid the problems, the following value-added form is usually used by adding prior achievement in grade *g** to the equation.

$$O_g = f'(O_{g^*}, F^{(g)}, P^{(g)}, C^{(g)}, T^{(g)}, S^{(g)}, \alpha). \tag{2}$$

The education production function is an extension of the theory of production and implicitly assumes that schools know the form of the production function and maximize outputs, i.e. student achievement, subject to a resource constraint or budget. Thus, schools are engaged in multi-product or joint production in a range of cognitive, affective and social outcomes. The resources designated for each of the students may be separable (as with time spent on particular curricular activities) or shared (e.g., capital expenditure on school infrastructure). (Levačić and Vignoles, 2002) Education production function studies are very empirical in nature and there is very little theoretical research in this area. Recent exceptions are the studies of Lazear (2001), Akerlof and Kranton (2002) and Todd and Wolpin (2003).

Determining the *outcomes of schooling* is not straightforward, since schooling has several tasks in society. Economists view schooling as investment or consumption. Schooling as an investment is related to its various outputs contributing to individuals' or society's productive skills and future well-being (Schultz, 1961; Becker, 1975). When schooling is considered as consumption it refers to the joy, pleasure and similar benefits derived from studying (Schultz, 1967).

Schooling also functions as a socialization or screening device (Spence, 1973). Socialization refers to the formal role of education influencing the later lives of individuals. Schools take much of the work of parents as educators and socialize children into this society. Studies concentrating on the socialization task of schooling analyze the effect of schooling on occupational choice, mobility, earnings and the relationship between schooling and personal and family characteristics. The screening task of schools refers to schooling as a device for sorting students for labor market purposes.

Studies on the education production function have usually a narrower point of view and take the skills and knowledge acquired through the schooling process as the output of schooling. In the early stages of schooling, the acquisition of basic skills, such as mathematical and verbal skills, provides a measure of the success of the process, and these skills are usually measured with standardized tests. In later years, examination results also provide information on the knowledge of students. There are at least two pragmatic arguments (Hanushek, 1979) in favour of achievement tests. First, test scores seem to be valued in and of themselves, and parents, educators and employers think that they are important indicators of individual skills. Second, test scores seem to be important in selecting individuals into further schooling.

The disadvantage of using more easily measurable test scores as the output measure is that they crowd out noncognitive skills, i.e. personal and interpersonal skills, behaviour and social attitudes, which are also highly important outputs in the process of schooling. There is recent empirical evidence that these skills are significant determinants of educational attainment and labour market success (for a review see e.g. Cunha *et al.,* 2006). Studies have also shown that investments in noncognitive skills in early childhood improve the acquisition of cognitive skills later on. The importance of noncognitive skills is also stressed in school curricula, where they are usually mentioned as major goals of primary and secondary education.

Other outcomes of the schooling process are students' attitudes towards learning, the attempts of schools to influence student lifestyles, including career aspirations, health habits, and sex and family education, which are even less frequently used as outputs in studies on the education production function. In all the essays of this thesis, scores in national matriculation examinations are used as the outcome measure of schooling, i.e. this thesis concentrates on measuring the ability of schools to promote the cognitive skills of students.

Family background is one of the key *inputs* in the educational process and one of the strongest predictors of student performance. Its importance was already stressed in the Coleman report (Coleman *et al.,* 1966). Family background can be measured, for instance, in terms of parental education, income or occupational status, students' eligibility for free school meals (e.g. in UK and US studies), and the number of books at home. These variables are used to depict the influence parents have on their offspring and they capture both the inherited characteristics and the influence of the social environment.

Differences in children's abilities across different income levels and racial groups are already clear in early childhood. These differences are significantly reduced after controlling for maternal education, cognitive ability, and family structure (see e.g. Cunha *et al.,* 2006). Since skill formation is a cumulative process, there is also evidence that investments in cognitive and noncognitive skills, especially in early childhood, have high returns. In particular, many noncognitive abilities are more difficult to influence in later years. Hence, children with parents having a high income and education are better off.

Only a few studies have explicitly concentrated on the effect of family background and parental effort on student achievement. Recently, Houtenville and Smith Conway (2008) examined the effect of parental effort on student achievement. They measured parental effort in terms of how often parents discuss important events, activities, and different school matters with the child and how often parents attend school meetings and events. The results showed that parental effort is strongly related to student achievement and is not captured by the usual family characteristics included in education production function models. The exclusion of parental effort, however, did not bias the results. Their results also showed that parental effort acted as a substitute for school resources, i.e. parents reduced their effort in response to increases in school resources.

Family structure as one component of family background also influences the educational process and success in the labour market. Recently, Black *et al.* (2005) investigated the effect of family size and birth order on children's education based on Norwegian data. According to their results, birth order had a significant and large negative effect on children's education. It also strongly affected adult earnings, employment and child bearing, especially among women.

These findings are congruent with numerous earlier studies (Black *et al.,* 2005; Pekkala, 2003). In Finland, Pekkala (2003) found that family size and the birth order interval affected schooling but not later earnings.

In addition to the direct effect on achievement, e.g. through parents' efforts to contribute to their offspring's education, family background may also have an indirect effect so that it affects, for instance, grouping decisions in the school. In some studies it is treated as a non-school factor that is outside the school's control but nevertheless affects school work.

Alongside the background of parents, *students' innate abilities* influence their achievement. There is a lack of clarity in defining these factors (see Hanushek, 1979), but most presumably they refer to students' learning capacity, motivation to learn or effort. Studies on the education production function do not usually include measures of innate ability of students, because such measures are not available. Todd and Wolpin (2003) have considered modelling the education production function so that it captures the theoretical notions that a child's development is a cumulative process depending on family history and school inputs as well as on innate ability, and have tackled the problem of missing information on innate ability.

In the value-added formulation of the education production function, previous achievement is used to control the cumulative nature of past inputs. As mentioned earlier, they have a positive impact on educational attainment and subsequent earnings in the labour market. There are different ways to measure previous achievement, but the most frequently used measures are certain standardized tests taken in earlier grades. In the studies of this thesis, the grade point average (GPA) of students in the comprehensive school report is used as a measure of earlier achievement.

Community characteristics in the education production function refer to socioeconomic characteristics of the whole community and they are usually measured in terms of income level, occupational structure, racial distribution, education level, poverty rates and other similar indicators.

Peer influences refer to the social impact of other students in the school. They are usually measured with factors related to general characteristics of the student body such as earlier achievement, family income, and educational level of the mother and father. There has been very little theoretical discussion on peer effects, but Akerlof and Kranton (2002) recently developed a theoretical model to examine how social categories, and the prescriptions of those categories, affect academic achievement, as well as how schools can influence students' choices of peers. They based their models on sociological research on peer behaviour in schools.

Interest in studying *peer effects* has been growing in recent years among economists. There have been numerous studies, which have usually found positive effects of school- or classmates on student achievement (Summers and Wolfe, 1977; Evans *et al.,* 1992; Hoxby, 2000; Levin, 2001; Fertig, 2003; Hanushek *et al.,* 2003; Robertson and Symons, 2003; Angrist and Lang, 2004; Schneeweis and Winter-Ebmer, 2007). Educationalists have also been investigating peer effects or context and contextual effects (Willms, 1986; Willms and Raudenbush, 1989).

Teachers play a central role at school. Their skills and personal qualities mostly determine the schooling experience and learning of students. Teachers' salaries also constitute the largest budgetary element in schools. Despite their evident role in the schooling process, research has had difficulties in verifying their importance. To date, the characteristics of good teachers have remained somewhat undetermined, even though their importance has been indirectly shown (Rivkin *et al.,* 2005).

Measurement of the qualities of teachers has been one of the major problems. The usual measures include formal education, experience and earnings. It is not, however, clear how far these measures capture the actual variation in quality. In addition, the measures may have only limited variation. For example, the variable for teachers' education may only contain a small level of variance because of the formal requirements set for teachers.[2]

Different selection mechanisms also complicate the study of teacher effects. First, there is some evidence that teachers tend to be selective in choosing their school. Teachers with longer experience prefer working in schools with higher student achievement (Hanushek *et al.,* 2004). Selection also takes place inside schools, where students are assigned to certain teachers or teachers are assigned to classes based on student ability.

Partly because of the above-mentioned problems, results concerning the effect of teachers' traits on student achievement are mixed. Hanushek (2003) reported that in high quality studies using value-added estimates and student-level data, teachers' education was statistically insignificant in 91% of the 34 estimates. Teacher experience seems to be clearly more important, since in 41% of the 33 estimates the relationship was positive and statistically significant. The effect of teacher salary is less convincing, as 82% of the 17 estimates were statistically insignificant. Studies using teachers' test scores as a measure of teacher quality are fewer in number compared to studies using the above-mentioned measures, but they do not provide any more robust results. Six out of nine estimates were statistically insignificant in explaining student achievement (Hanushek, 2003).

Recent studies using data linking students and teachers and employing school and teacher fixed effects to control for the selection and omitted variable bias have revealed that observable teachers' characteristics matter. The test results of students with more educated teachers were higher (Goldhaber and Brewer, 1997; Ehrenberg and Brewer, 1994). Rivkin *et al.* (2005) concluded that the observable effect was small, even though the contribution of teachers on the whole was large. Goldhaber and Anthony (2007) have also found that students of certified teachers had a higher level of achievement. Dee (2007)

[2] In Finland a qualified upper secondary school teacher has to have a Masters' level university degree in most subjects.

found in his study that after controlling for subject and student fixed effects constant across subjects, students assigned to a teacher of the same gender achieved better results than those assigned to teachers of the opposite gender.[3]

From the policy making perspective, the *effect of school expenditure* on student achievement and subsequent labour market success is very important. The usual assumption is that additional resources enhance learning. The effect of school expenditures and class size has been widely studied since the Coleman report (1966) and the results have been mixed. Hanushek's (2003) famous review showed that according to the majority of estimates, expenditures per student do not affect student achievement, whereas Hedges and Greenwald (1996) came to the opposite conclusion in their meta-analysis.

Traditionally, educational production functions have been estimated assuming that school resources or class size are exogenous, i.e. independent of other regressors. There is, however, selection taking place both between and within schools that causes endogeneity. Parents with a higher socioeconomic status may choose a school based on its better resources or teachers. Inside the school, weaker students may be placed in smaller groups or students may be grouped based on their knowledge. The problem of endogeneity has been resolved in empirical work in various ways.

The instrumental variables approach has most often been applied to overcome the endogeneity problem, and the results of such studies have usually shown a small positive effect of school resources on student achievement (Hægeland *et al.,* 2007; Levačić *et al.,* 2005; Jenkins *et al.,* 2006). However, there is also evidence that resources do not matter, at least not for women, but that they have an effect on the wages of black women (Betts, 2001). Studies using cross-country variation have often shown a positive relationship between school resources and student achievement (see Fuchs and Wössmann, 2007).

[3] The results of earlier studies on the effect of the teacher's gender have been mixed (see e.g. Dee, 2007).

The problem with these studies is the level of aggregation, which produces uncertainty in the interpretation of the results.

The fifth article in this thesis examines the effect of teaching expenditures on student achievement. Instead of the instrumental variables approach it uses relatively long panel data from the years 1990-1998 with considerable reduction in teaching expenditures during the period and two-way fixed effects models to control for heterogeneity in the unobserved time constant and omitted variable bias. The problem of selection is controlled for by the comprehensive school grade point average. According to the results, teaching expenditures did not affect student achievement in Finnish upper secondary schools. The third essay of this thesis uses school level panel data from the years 2000-2004. It controls for inefficiency and school fixed effects and finds some evidence for the negative effect of teaching expenditures on matriculation examination results.

Research concerning the effect *of class size* on student achievement has been quite active in recent years. It has mostly been empirical in nature, while theoretical discussions have been more limited. A recent exception was Lazear's (2001) disruption model for class size effects based on the idea of disruptions disturbing teaching, requiring the teachers' attention and causing students to learn less. Hence, smaller classes are advantageous for disadvantaged students and the optimal class size is larger for better-behaved students. Lazear's model is close to an earlier model by Brown and Saks (1975) claiming that teachers take into account the distribution of abilities in the classroom and allocate their time accordingly.[4] Correa's (1993) model is based on the time allocation of rational teachers and predicts that teachers will focus on the overall performance of the class, which is more likely to harm individual students as the class size increases.

According to other possible explanations, teachers may be more effective in smaller classes. They have more time to concentrate on individual students and the curriculum may vary more. They may

[4] Empirical results by Brown and Saks (1987) showed that teachers use compensative strategies according to which they pay more attention to low performing students.

also be better able to vary their instructional methods according to the needs of individual students. As a consequence of more individual instruction, they may also be able to give feedback more often to individual students. Hence, students learn more in smaller classes.

The empirical evidence giving support to smaller classes is mixed (Hanushek, 2003; Krueger, 2003; Hedges and Greenwald, 1996), partly because studying the effect of class size is complicated. Researchers are faced with same kinds of difficulties as in studying the effect of teachers and school resources. A major problem is caused by selection, when parents with a higher education and income choose their place of residence and hence schools. Schools with higher spending and smaller class sizes attract parents with higher incomes, causing student background factors and class size to be endogenous in statistical models. In addition, the most often used measure for class size, the student-teacher ratio, is imprecise and the actual class size naturally varies considerably. Few studies have been able to use a precise measure for class size.

The problem of selection has been solved in different ways in class size research. Studies using natural experiments are perhaps the most convincing, even though they also suffer from some weaknesses. They are also rare because they are laborious to execute. The Tennessee STAR experiment is one of the few studies using an experimental design (see Mosteller, 1995). Krueger (1999) analyzed the Tennessee STAR experiment conducted on elementary school children in the U.S. at the end of the 1980s and found fairly large effects of smaller classes for the first years of schooling. These effects were also quite stable in the sense that they persisted throughout the schooling career, so that students in smaller classes were more likely to take the college entrance test and also performed somewhat better in it (Krueger and Whitmore, 2001). Particularly minority students benefited from smaller classes.

Quasi-experiments and the instrumental variables approach are more often used methods to overcome the selection problem. Many recent studies have exploited some administrative rule determining the maximum class size as their instrument for class size. Most of these studies have provided evidence that learning is improved in smaller classes (Angrist and Lavy, 1999; Bonesrønning, 2003), but there have also been contradictory results (Hoxby, 2000; Dobbelsteen *et al.,* 2002). Browning and Heinesen (2007) and Bingley *et al.* (2007) found that smaller classes increased the number of years of schooling. Some studies have based their identification strategy on average class size or student enrolment in the school (Akerhielm, 1995; Levačić *et al.,* 2005; Jenkins *et al.,* 2006; Wössmann and West, 2006) and have found small negative effects.

In this thesis, the effect of the student-teacher ratio is examined in two essays. In the first essay assessing the efficiency of Finnish general upper secondary schools with DEA, the effect of the student-teacher ratio on inefficiency (1-efficiency) is tested with Tobit models. According to results, the relationship is non-linear, so that inefficiency initially increases and after reaching a class size of eleven it starts to decrease.

In the third essay of this thesis, the effect of the student-teacher ratio on matriculation examination results is tested with panel data using stochastic frontier models that control for inefficiency and variation that is constant through time and not captured by variation in the explanatory variables. In most models, the effect is positive, indicating that student performance is better in schools with a higher student-teacher ratio.

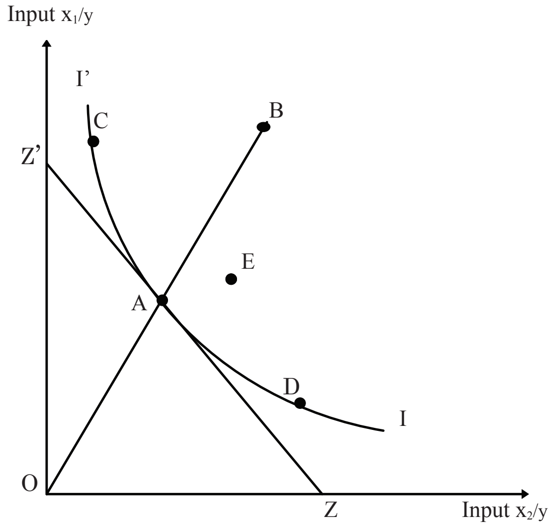
2.2 Measurement of efficiency

The concept of efficiency may have different meanings depending on the context. Sometimes, the concepts of productivity and efficiency are used as counterparts. In efficiency measurement they usually have separate meanings. The productivity of a production unit, in this thesis a school, refers to the ratio of its outputs to inputs, whereas efficiency refers to a comparison between observed and optimal values of a school's output and input, i.e. a comparison between observed and optimal values of a school's productivity.

Economic theory recognizes several types of efficiency. *Technical efficiency* is determined either as the ratio of observed to maximum potential outputs obtainable from the given inputs or as the ratio of the minimum potential to observed inputs required to produce the given outputs. In the former case, a school is viewed as maximizing its outputs with the given inputs, and in the latter case as minimizing the use of inputs when the outputs are fixed. *Cost efficiency* takes the prices of inputs into account. In order to be cost-efficient, a school must be technically efficient and produce its outputs by minimizing the costs of inputs. *Allocative efficiency* considers revenues and requires that schools are technically efficient and produce their outputs by maximizing their revenues.

The first two efficiency concepts are illustrated in Figure 1 when constant returns to scale (CRS) are assumed, i.e. an increase in the input increases the output in the same proportion. Let us consider an educational sector that is using two inputs, x_1 and x_2, to produce one output, y. These two inputs are depicted in the vertical and horizontal axes, respectively. The curve I'I represents the isoquant determining the best possible combinations of x_1 and x_2 to produce the given level of y. As this curve determines the input-minimizing efficient production technology, it is also called the efficiency frontier. Points A, B, C, D, and E depict schools that use different combinations of inputs x_1 and x_2 to produce the same output y.

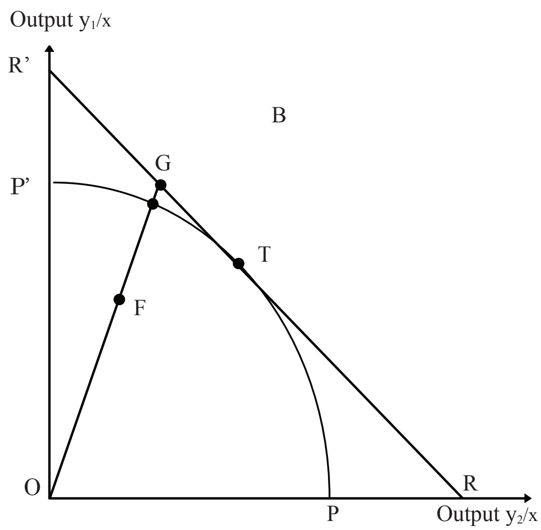
Figure 1. *The concepts of technical and cost efficiency when inputs are minimized*



Schools C, A, and D are at the production isoquant I'I and are therefore technically efficient. All the other schools above the isoquant are technically inefficient, since they use more of at least one input to produce the same amount of output. Furthermore, if the line Z'Z depicts the iso-cost line determining the cost-minimizing production technology, school A is also cost efficient because it operates at the point where the iso-cost line passes through the isoquant I'I, i.e. it produces its output with cost-minimizing inputs.

School B is technically inefficient because it uses more of both of the inputs relative to efficient technology described by isoquant I'I. An input-oriented measure for technical efficiency TE proposed by Farrell (1957) is obtained for school B as the ratio of OA/OB. It obtains a value between 0 and 1 depending on the school's distance from the efficiency frontier. In case of technical efficiency, it obtains a value of 1. Inefficiency is obtained as 1 – TE and it gives the proportion by which the use of school's inputs could be reduced without reducing the output y. As for school E, it is both technically inefficient and cost inefficient, because it uses inputs in cost-inefficient proportions.

Figure 2. The concepts of technical and allocative efficiency when outputs are maximized



Alternatively, technical efficiency is possible to determine so that schools are assumed to maximize the output while holding the use of inputs constant (see e.g. Johnes, 2004; Fried *et al.,* 2008). This is the output-oriented case and technical inefficiency gives the proportion of outputs that could be increased with the given inputs. This is illustrated in Figure 2. In this case, schools are producing two outputs y_1 and y_2 with one input x and face the production possibility curve PP'. Technical efficiency TE for school F is determined as the ratio of the observed output to the maximum possible output OF/OG. Inefficiency is obtained as 1-TE and it depicts the proportion by which outputs of the school could have been increased without increasing the level of input used. Line RR' depicts the isorevenue line. School T is both technically and allocatively efficient. Since G has the same revenue as T, allocative efficiency can be measured by the ratio of OF/OG.

In this thesis, technical efficiency is used as the efficiency concept throughout all the essays. The concept of cost efficiency is not used, because the data requirements are usually more demanding.

In addition, because the focus is on schools providing general upper secondary education and with limited possibilities to influence the cost structure, it is more relevant to concentrate on technical efficiency.

Both input- and output-oriented technical efficiencies are used in this thesis. In essays 1 and 2, input-oriented technical efficiency is used when efficiency is measured with DEA. Output-oriented technical efficiency is used in essays 3 and 4 measuring efficiency differences with stochastic frontier analysis. In this case, the choice of orientation is constrained by the method.

2.3 Methods for efficiency measurement

Methods for efficiency measurement are either statistical or non-parametric. In the latter case they are usually based on linear programming. In addition, methods can be divided into deterministic and stochastic ones. Deterministic methods interpret the whole deviation from the efficiency frontier as inefficiency, whereas stochastic methods allow for random shocks in production. In such a case, only part of the deviation from the efficiency frontier is interpreted as inefficiency. A recent description of both parametric and non-parametric methods for efficiency measurement is provided by Fried *et al.* (2008). Stochastic frontier analysis is presented in Kumbhakar and Lovell (2000). There have also been some excellent surveys concerning the measurement of efficiency in educational institutions (see e.g. Worthington, 2001 and especially Johnes, 2004). Recent developments have narrowed the gap between statistical methods and linear programming. Kuosmanen (2006) has presented a stochastic non-parametric envelopment method that attempts to combine the advantages of data envelopment analysis (DEA) and stochastic frontier analysis.

Statistical methods

In statistical methods, it is assumed that efficiency follows a certain distribution. In addition, these methods are usually parametric. Hence, one has to assume some functional form for the efficiency frontier.

As a consequence of this assumption, statistical methods provide a description of the relationship between inputs and output. In case of the education production function, they provide information on the relationship between the outcome measure and, for instance, family background and school resources.

The possibility for description of the production function is one of the advantages of statistical methods compared to non-parametric methods. Other advantages include the possibility to use statistical significance testing when choosing the inputs for the analysis and the separation of statistical noise from inefficiency. This latter characteristic is especially important in practically oriented efficiency measurement, because it reduces the probability of overestimating the magnitude of inefficiency.

The disadvantages of statistical methods are that the misspecification of the functional form affects the efficiency and there is no theoretical grounding for the choice of functional form. In addition, in case of multiple inputs and multiple outputs, they become more complicated to apply. Some critics have also suggested that the confidence intervals for the inefficiency term are too wide to gain credibility in practice (Johnes, 2004, 628). In addition, since there is no theoretical grounding for the distributional assumption of the inefficiency term, a misspecification of the assumption is also reflected in the inefficiency. This is not a minor issue, since the assumption affects the magnitude of inefficiency and the ranking of units. According to Kumbhakar and Lovell (2000, 90), the sample means are sensitive to the distribution assumption, whereas the rankings and top and bottom deciles of inefficiency scores are more likely to remain unaffected.

Statistical parametric methods for efficiency measurement can be deterministic or stochastic. In deterministic methods, the residual term of the regression equation is interpreted as depicting efficiency. There are essentially three methods (COLS, MOLS, MLE) to estimate the production frontier and they differ in how they define the location of the efficiency frontier (see e.g. Greene 2008; Kumbhakar and Lovell, 2000; Johnes, 2004). In addition, in COLS there is no assumption about the distribution of the error term. Deterministic approaches are easy to use, but their drawback is that they interpret the whole deviation from the frontier as inefficiency, and they do not allow for random shocks, which might also contribute to variation in the output. In that sense they are similar to conventional OLS. A more realistic approach is to allow for random shocks.

In stochastic methods, random shocks are separated from inefficiency. Stochastic production frontier models were first introduced in the late 1970s (Aigner *et al.,* 1977 and Meeusen and van den Broeck, 1977). These methods are referred to as stochastic frontier analysis (SFA). Stochastic frontier analysis allows for technical inefficiency and also takes into account random shocks and fluctuations in the production process. The separation of random fluctuations from technical efficiency is performed by dividing the residual component into two parts. The first part is the normally distributed error term capturing the measurement error and random fluctuations. The other part is the one-sided inefficiency term, which is usually assumed to have a half-normal, exponential or truncated normal distribution.

Stochastic frontier models have been developed both for cross section and panel data (see Kumbhakar and Lovell, 2000). Cross section models use only variation across production units, whereas panel data models also take into account the variation within production units. There are several variations for panel data models. The model by Pitt and Lee (1981) is an extension of the random effects model, treating the time constant random effect as inefficiency. A fixed effects counterpart was introduced by Cornwell *et al.* (1990).[5] In both of these models, inefficiency is treated as constant through time. Recent panel data frontier models further separate inefficiency that varies through time from the unmeasured heterogeneity that is constant through time (Greene, 2005a, b). Latent variable models have also been developed (Greene, 2005a, b).

[5] This model is actually deterministic, since no assumption for the inefficiency term is imposed but the whole fixed effect is interpreted as inefficiency.

In more advanced cross section and panel data models, it is possible to account for heterogeneity in the inefficiency term (Battese and Coelli, 1995 and Coelli *et al.,* 1999). In these models, the inefficiency term is a function of some explanatory factors that characterize the environment or the organization of the producer.

Empirical studies using stochastic frontier analysis in the field of education are quite small in number. In addition to measuring efficiency differences across schools or school districts, they have investigated the effect of school resources (Deller and Rudnicki, 1993), class size (Cooper and Cohn, 1997; Mizala *et al.,* 2002) and teachers' merit pay (Cooper and Cohn, 1997) on student achievement.

The question of what factors affect inefficiency differences has only recently been addressed in some studies. Kang and Greene (2002) investigated the effect of parental choice and competition on inefficiency using the model by Battese and Coelli (1995). Conroy and Arguea (2008) studied the effect of the crime rate, suspension and expulsion rate, the proportion of disadvantaged students, the parent-teacher organization and location on inefficiency using Tobit models in the second stage.

Stochastic frontier analysis has more often been applied to study the cost efficiency of schools and school districts. One of the early studies was that by Barrow (1991), which investigated the cost efficiency of British local education authorities (LEAs) with panel data. Heshmati and Kumbhakar (1997) used a model by Battese and Coelli (1995) in their study of Swedish primary schools. In their model, cost inefficiency was dependent on socioeconomic characteristics, teacher characteristics and the wealth of the municipality.

Only a few studies have made use of panel data. Barrow (1991) studied the cost efficiency of local education authorities using both cross-section and panel data with stochastic and deterministic methods. Johnes and Johnes (2009) analysed the cost efficiency of British universities using SFA, allowing heterogeneity between universities with random parameters (Greene, 2005a, b). Aaltonen *et al.* (2006) studied the cost efficiency of Finnish comprehensive schools with pooled panel data and true fixed effects models (Greene, 2005a; b), assuming a half-normal and truncated normal distribution for the inefficiency term.

In this thesis, I apply stochastic frontier methods in measuring the efficiency of Finnish general upper secondary schools in two essays. In the third essay, I compare the results of different stochastic frontier models for panel data, emphasizing the properties of recent models by Greene (2005a,b). In the fourth essay, efficiency measurement of case schools is based on random (Pitt and Lee, 1981) and fixed effects (Cornwell *et al.,* 1991) stochastic frontier models, and the results of the estimations are connected to qualitative analysis and findings concerning the organizational practices of the schools.

Non-parametric methods

Non-parametric methods for efficiency measurement are based on linear programming. Data envelopment analysis (DEA) is the most common non-parametric method for efficiency measurement. In DEA, the efficiency frontier is estimated as a linear combination of inputs and outputs and the efficiency of each unit is assessed relative to the frontier. All the deviation from the efficiency frontier is interpreted as inefficiency. Because of the method's non-parametric nature, no assumption is made about the functional form of the production technology. Hence, misspecification of the production function is not a problem in DEA. For the same reason, the method does not provide any estimates or significance tests of the parameters and efficiency scores.[6] DEA is easy to use in a setting with multiple inputs and multiple outputs. The weakness of the method is its sensitivity to errors in the data. This is a serious problem, especially if there are only a few observations determining the efficiency frontier.

[6] In recent years, to overcome some of the shortcomings, bootstrapping methods have been developed to calculate the confidence intervals for DEA efficiency scores (see Simar and Wilson, 2000; 2007).

Different DEA models have been developed over the years. The basic model, assuming constant returns to scale in production, was developed by Charnes *et al*. (1978). A model assuming variable returns to scale was presented by Banker *et al.* (1984). These two models are used in efficiency measurement in the first and second essays of this thesis. Since the publication of these models, various extensions have been developed. For a recent review of DEA and its various extensions, see Thanassoulis *et al*. (2008) and Charnes *et al*. (1994).

DEA has been widely used in efficiency evaluation in education. One of the first applications of DEA was applied to schooling (Charnes *et al.,* 1981). Since this study, there have been number of DEA applications measuring the efficiency differences between schools and school districts and testing the applicability of the method in number of countries (Bessent and Bessent, 1980; Bessent *et al*., 1982; Ludwin and Guthrie, 1989; Färe *et al*., 1989; Jesson *et al*., 1987; Smith and Mayston, 1987; Bonesrønning and Rattsø, 1994; Kirjavainen and Loikkanen, 1993).

The choice of inputs is somewhat problematic in DEA because of its non-parametric nature. One strategy is to include all inputs into the DEA and pay no attention to whether or not they are controllable by the school. This approach does not make a distinction between the environments schools are facing, and a school with a harsh environment may be judged inefficient even though its performance is due to the environment. This approach, in other words, may overestimate inefficiency (Johnes, 2004, 657). Two main strategies have been adopted to overcome the problem related to environmental factors. Some studies (e.g. Banker and Morey, 1986; Ruggiero, 1996; Thanassoulis, 1996) have developed DEA models that take into account differences in the environment.

Alternatively, the majority of studies have adopted a two-stage approach in which the efficiency differences are determined using DEA with inputs controllable by the school. In the second stage, efficiency scores are statistically explained with environmental and other variables that may be the cause of inefficiency. This latter approach is adopted in the first and second essays of this thesis.[7]

There are several problems associated with this approach (see Johnes, 2004). First, there are no clear guidelines for the selection of variables in first and second stage. The most common approach is to include variables controlled by the institutions in the first stage and those not controllable by the institution in the second stage. The specification of the functional form in the second stage also affects the results and is hence important to avoid misspecification errors.

The first essay of this thesis presents one of the first studies applying a two-stage approach to study the efficiency of schools. It concentrated on the effects of the student-teacher ratio, school size, the financial situation of the municipality and location of the school on efficiency. Subsequently, a two-stage approach has been applied to investigate the effect of competition (Bradley *et al*., 2001; Duncombe *et al*., 1997; Waldo, 2003), voter monitoring (Duncombe *et al*., 1997) and the political context of the municipality (Waldo, 2003) on efficiency.

A few recent studies have used qualitative research methods in the second stage. Mancebón and Bandrés (1999) examined Spanish secondary schools by interviewing the headteachers of five efficient schools, asking their view of the reasons for the school being efficient. Dodd (2006) studied 38 efficient UK secondary schools with surveys and interviews. Portela and Camanho (2007) analyzed a few Portuguese secondary schools by visiting the school site. The fourth essay of this thesis uses a similar approach and assesses efficient and inefficient schools with case study data based on interviews.

To date, most DEA applications have used school, institution, district or municipal level data. Applications using individual-level data have only been used in some recent studies, even though it has been a standard in other education production function studies for some

[7] Ray (1991) and McCarty and Yaisawarng (1993) have compared the results of DEA and the two-stage approach.

time. Portela and Thanassoulis (2001) and Thanassoulis and Portela (2002) analysed the efficiency of English secondary school students and schools by decomposing the efficiency into one attributable to the individual and one attributable to school. The results of individual-level DEA have been compared to those of multilevel modelling (Johnes, 2006).

Comparisons of results between different methods have been quite common. Some studies have compared the results of SFA and DEA. An early study by Sengupta and Sfeir (1986) compared the results of regression methods and linear programming, concluding that the results differed between these two methods. Mayston and Jesson (1988) also came to the same conclusion when comparing the performance of local education authorities in the UK using both regression techniques and DEA.

Mizala *et al*. (2002) assessed the efficiency of Chilean schools with DEA and SFA. According to their results, the efficiency rankings of SFA assuming a half-normal distribution for inefficiency were similar to DEA efficiency scores. Chakraborty *et al*. (2001) and Aaltonen *et al*. (2006) compared the results of SFA with Tobit-model residuals that relate the DEA efficiency scores to uncontrollable environmental factors. Both studies demonstrated that the efficiency scores of SFA models assuming a half-normal distribution and Tobit model residuals are very similar.

2.4 School effectiveness studies

School effectiveness studies or school effectiveness research (SER) have the same origin as education production function studies in economics. School effectiveness studies are conducted by educationalists and have concentrated on school effects and processes and their effect on student achievement. School effectiveness studies have largely ignored the effect of school resources on student achievement, which has been the main interest of economists. The education production function may also be viewed as a simplification of the school effectiveness framework, since it does not consider the effect of process variables (Levačić and Vignoles, 2002).

The study of school processes is methodologically based on viewing schools and school processes as a nested structure (Bidwell and Kasarda, 1980; Burstein, 1980), where students are nested within classes, classes within schools and schools within communities. School effectiveness studies have emphasized the importance of examining these different layers and their inter-relationships. These studies are usually statistical and they apply multilevel modelling (Goldstein, 2003; Raudenbush and Bryk, 2002) to analyse school effects.

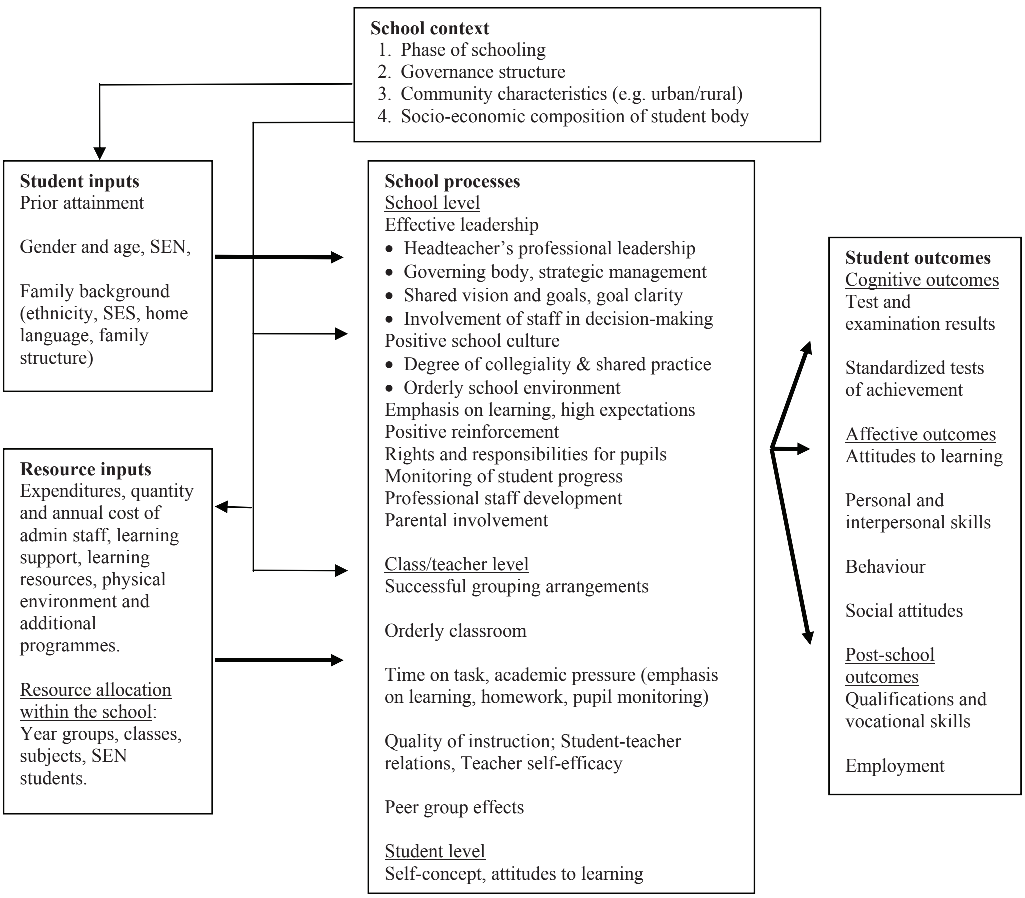
Main interest in statistical school effectiveness studies has been in studying the magnitude of school effects, their consistency across different subjects and stability through time, differential school effects on different characteristics of students, and the continuity of school effects measured at different points in time during the school career (Teddlie and Reynolds, 2000; Sammons, 1999).

A different approach has been taken in studies that first identify schools with differential effectiveness and then take a closer look at some of these schools, usually the ones with a high or low performance. The focus has been on both school- and teacher-level processes. At the school level, the focus has naturally varied across studies, but the following topics have been investigated: school climate, culture, leadership, instructional arrangements, teacher effectiveness, student monitoring, staff development and parental involvement, among others.

Data on school processes in statistical studies have usually been collected from surveys or direct observation and transformed into quantitative variables. Studies on effective and ineffective schools have also relied on interviews and other sources of qualitative data that are used to understand the processes of these schools.

Figure 3 summarizes the school effectiveness framework and presents the main groups of variables (see Teddlie and Reynolds, 2000; Levačić and Vignoles, 2002). The outcome of the schooling process is measured at the individual level and depicts cognitive, affective and/or post school performance. In the model, it is assumed that these outcomes are a product of interaction between the school context, student inputs and resource inputs. The assumed causal links are from the context and inputs to school processes, which in turn determine student outcomes. Processes act at school, teacher and student levels.

Figure 3. A context-input-process-outcome framework for school effectiveness (Levačić and Vignoles, 2002)



School improvement research as a research tradition close to school effectiveness research has focused on studying schools that are implementing school change efforts and aiming at improving their performance (Teddlie and Reynolds, 2000). Such research has mainly been conducted in the U.S. and U.K. While school effectiveness studies have often been statistical, studies of school improvement are more often qualitative. They also concentrate on thoroughly describing the school practices and how the change has evolved in the case schools. As effective schools research has mainly been cross-sectional, examining schools at one point in time, some school improvement studies have also applied longitudinal approaches by following case schools over a longer period of time (see e.g. Gray *et al*., 1999). In doing so, they attempt to reveal the dynamic processes behind the change.

Processes of effective schools

The results of various school effectiveness studies have been reviewed several times over the years. Recent excellent and thorough reviews include those by Teddlie and Reynolds (2000) and Sammons (1999). Other older reviews include Purkey and Smith (1983), Levine and Lezotte (1990) and Scheerens and Bosker (1997). The review of Lee *et al*. (1993) concentrated on studies of secondary schools. These reviews have produced lists of processes in effective schools. There is a small amount of variation between the lists, but the school processes in Figure 3 are mentioned in almost all of them. Some of the processes have consistently been found in effective schools in every study where they have been investigated. However, the list also includes processes that have not been present in every study. In the following, the main findings related to school processes at the school level are discussed. The presentation is limited to school-level factors, since they are the main focus of this thesis.[8]

[8] Some of the other factors were already discussed earlier in connection with studies on the education production function.

Effective leadership

Strong school-site leadership fostered by clear goals has been found to characterize effective schools in almost every study (Reynolds and Teddlie, 2000). This involves a head teacher who is firm and purposeful and takes the initiative and a leading role in development processes of the school. Effective school leaders also mediate and buffer the school from outside disturbances and provide autonomy and continuity for school work.

A participative approach is also one feature of effective leadership, so that in decision making the views of all teachers are taken into account and the teachers feel they can influence school matters even though the primary source of teachers' power comes from within the classroom, i.e. from the ability to control the curriculum and classroom practices (Lee *et al*., 1993). Teachers' involvement in curriculum planning is also important in participative leadership (Mortimore *et al*., 1988). Rutter *et al*. (1979) also pointed out that the participative approach involved the motivation and use of the management team. All this is related to another characteristic of an effective school, namely the extent to which the school culture is a collaborative one.

Instructional leadership is a third component of effective leadership. It is related to the role of principals as leading professionals (Sammons, 1999) who are aware of what goes on in the classroom, including the curriculum, teaching strategies and the monitoring of student progress. The role also requires supporting, encouraging and giving practical assistance to teachers if needed.

The fourth and fifth components in effective leadership include frequent, personal monitoring of staff performance and the proactive selection and replacement of staff. Personal monitoring of staff has turned out to be important in every study in which it has been included (Reynolds and Teddlie, 2000). It involves visiting classrooms and giving advice and support to teachers in decision making. Proactive selection and replacement of staff involves head hunting for good recruits and pressuring less competent staff to improve their performance or to move to another school.

Focusing on learning

Focusing upon the importance of academic goals and processes and an academic emphasis have been shown to be correlates of effective schools. These factors have been followed using measures such as a high entry rate for public examinations, the setting of homework with regular tests and a wide curriculum coverage. In addition, maximizing the available learning time at school is important (Reynolds and Teddlie, 2000).

Generating a positive school culture

A positive school culture in terms if a shared vision or mission among staff members that is collaboratively put into practice is an important characteristic of effective schools pointed out in many studies and reviews. It also involves consistency in approaching the school curriculum and in rules regarding rewards and sanctions.

Collegiality and collaboration are important prerequisites for the unity of purpose, and it has also described earlier how teachers' power to influence decision making is another a vital part of a positive school culture. There is also research evidence that changes are more successful when they are planned and implemented in collaboration with teachers and administrators. This kind of engagement brakes barriers between teachers and administrators and increases the feeling of unity and communality among staff (Lee *et al*., 1993).

Other characteristics involved in a positive school culture include order and discipline within the school. An orderly atmosphere enables students to concentrate on their school work without unnecessary interruptions. Mortimore *et al*. (1988) pointed out that although an orderly environment is a condition for effective learning to take place, it does not necessarily mean that when schools become more orderly they also become more effective. A positive climate for students is also important. It relates to the research results that a school culture with harsh punishments and overly strict control creates tense and negative attitudes towards teachers (Reynolds and Teddlie, 2000). A school culture that rewards good behaviour, achievement, effort and attributes is more productive in promoting a positive climate in the school.

High expectations of achievement and behaviour

High expectations that are clearly communicated to the students are also an important characteristic of effective schools. This is also a factor that has been consistently related to effective schooling in various studies and reviews. According to Reynolds and Teddlie (2000), it is most likely related to other factors and characteristics mentioned earlier such as time management, principal leadership and a positive school climate. High expectations may also be a result of other teacher qualities such as activism, a sense of the internal locus of control or the belief that schools can counter the effects of a disadvantaged student background. And it is quite likely that high expectations occur together with staff that place high expectations on themselves.

A school culture that recognizes and honours academic success officially and school-wide was also one characteristic of successful schools. Academic goals were stressed through the use of symbols and ceremonies that encourage students to adopt similar norms and values.

Emphasizing student responsibilities and rights

There is evidence that students' involvement in different school-related activities and school decision making is present in more effective schools. Giving students more responsibilities and a chance to influence school matters is expected to enhance their commitment and internalization of school values. There is also some evidence that giving students more responsibility for their own work and control over learning situations enhances their performance (Reynolds and Teddlie, 2000).

Monitoring progress at all levels

The monitoring of student progress is a characteristic of many effective schools. Sammons (1999) points out that it may be less directly related to student achievement but is an important ingredient of effective schools. It helps to determine how well the school goals have been achieved and it focuses the attention of staff, students and parents on these goals. It also informs planning, teaching methods and assessment and it gives a clear message to students that teachers are interested in their progress. There are also opinions that monitoring has been spuriously defined, providing little guidance for practical work. Some schools also waste effort in too frequent monitoring. Hence, some reviews emphasize that monitoring has to be 'appropriate' (Levine and Lezotte, 1990).

Staff development

School-based staff development is usually mentioned as one of the characteristics of effective schools. It seems to be important that it is practical, school-wide and closely related to the instructional program of the school. It should also be based on the expressed needs of teachers rather than on outside suggestions for remediation for teachers' deficiencies in certain skills. In some studies (Mortimore *et al*., 1988), teachers in ineffective schools were attending courses that were unrelated to the core mission of their school.

Parental involvement

In some studies, parental involvement and support has been found to be a major factor in student achievement, although contrary results have also been reported (Reynolds and Teddlie, 2000). Purkey and Smith (1993) suggested that parental involvement alone is not sufficient, but is only one of the factors that could positively influence achievement. Above, results were discussed according to which parental expectations enhanced student learning. However, other forms of parental behaviour, such as monitoring homework, tutoring, minimizing distractions from schoolwork and engaging in active choices concerning their children's educational programme also positively affect student learning (Lee *et al*., 1993).

The scope of two essays (essays 2 and 4) that concentrate on the causes of efficiency differences is based on school effectiveness studies, the findings of which guided the data-gathering process in both studies. The survey questionnaire used in the second essay and interviews analyzed in the fourth essay were constructed based on this research.

3. Institutional context: Finnish general upper secondary schooling in 1985-2005

Finnish general upper secondary schools provide post-comprehensive education for students aged 16 to 19.[9] The general upper secondary school certificate together with the matriculation examination certificate provides eligibility for university or tertiary level vocational education. Roughly 55% of students in each age cohort continue their studies in general upper secondary schools. Basic vocational education is pursued by some 40% of the age cohort.

The statute of general upper secondary education sets the objectives for schools, emphasizing the stable mental and intellectual development of individuals. The objective of the Finnish general upper secondary education is to support the growth of students into good, balanced and educated individuals and members of society as well as to provide them with the skills and knowledge needed in further studies, working life, hobbies, and all-round personal development. Moreover, the instruction should support the students in lifelong learning and self-development throughout their lives.

[9] General upper secondary education and the matriculation examination can also be completed in general upper secondary schools for adults or in the adult study line in general upper secondary schools. Students have to be over 18 years of age to qualify for adult education. In this thesis, the education for youngsters is studied and adult education is therefore not further discussed.

Since 1982, general upper secondary education has been divided into courses consisting of about 38 lessons. The school year is usually divided into five or six periods. A timetable is devised for each period, focusing on certain subjects. Until 1994, students followed year classes and mostly progressed with the same pace, completing their studies in three years. After 1994 the year classes were abolished and general upper secondary schools had a two-year transition period to start following a non-graded system. In this system, students prepare their individual study plans and devise a schedule for each period, focusing on certain subjects.[10] Student progress and the composition of teaching groups depend on the students' choice of courses. Students are able to carry out the studies over 2 to 4 years.

Within the present system, the general upper secondary education syllabus consists of a minimum of 75 courses. There are compulsory, specialization and applied courses. All students must complete the compulsory courses, the number of which varies from 47 to 52 depending on the course choices. Schools must provide specialization courses for students to choose from. Each student is responsible for completing a sufficient number of courses. Applied courses may be integrative courses with elements from various subjects, methodological courses, or vocational or other studies organized by the same or another education provider. Applied courses may also include courses taken as part of the general upper secondary school diplomas completed in art or physical education.

Institutional setting[11]

The Ministry of Education grants the authorization to provide general upper secondary education. Most general upper secondary schools follow a generalized curriculum. There are some 50 schools that have been permitted by the Ministry of Education to have a specialized

[10] There were also some 20 general upper secondary schools that experimented with the non-graded system before 1994.

[11] The main sources are the Internet pages of the Finnish National Board of Education (www.oph.fi), the Ministry of Education (www.minedu.fi) and Takala (1994).

curriculum. Schools can specialize in music and arts, sports, mathematics and sciences, and languages.[12] Specialized schools have their own entrance requirements that are related to their area of specialization. Some schools, especially in rural areas, also offer some of their courses as distance education jointly with some other general upper secondary schools.

General upper secondary schools are mostly maintained by municipalities. In addition, there are some private schools, schools maintained by the joint organization of municipalities and some state-owned schools. No municipal general upper secondary schools charge school fees. Municipalities cover school expenditures from their general revenue services, which consist of local income tax, property tax and non-earmarked grants[13]. Private schools receive a state grant on a per student basis, and they also obtain funding from municipalities. Private schools may charge minor fees.

The period of 1985-2005 can roughly be divided into two institutional settings. From 1985 until 1993 the regulations at the state level concerning the curriculum and funding of general upper secondary schools were quite restrictive, leaving only limited discretion for the schools and municipalities that largely maintained the schools. After 1993, state control was considerably reduced and decision making was decentralized to the local level.

In 1985-1993, schools followed a national curriculum that stated in some detail the content of instruction in each subject. The role of schools was to put the curriculum into action. The provision of general upper secondary schooling was also regulated so that the Ministry of Education made the decision concerning study places. Since 1993 the number of study places has followed local needs.

In 1994, state level control was considerably reduced and the autonomy of schools increased as decision-making power was transferred to the local level. Both the new state grant system and the national core

[12] These schools also have classes that follow the generalized curriculum.

[13] Until 1993 the grants were earmarked (see discussion below).

curriculum supported this trend. The national core curriculum acted as frame on which each school based its own curriculum.

The broad national objectives and the allocation of time to instruction in different subjects and subject groups as well as student counselling are decided on by the Government. By devising the National Core Curriculum, the Finnish National Board of Education determines the objectives and core contents of the various subjects, subject groups, cross-curricular themes and other instruction. It also sets the main principles for cooperation between the home and school and the key principles and goals in student welfare services. Based on these, each education provider draws up its own local curriculum. The curriculum must provide students with individual choices concerning studies, including utilizing instruction given by other education providers, if necessary.

At the same time as the transfer of decision-making power to the local level, the inspection of schools was also abolished in 1994. To ensure the quality of instruction, a national level evaluation system was established. The Finnish National Board of Education was responsible for evaluation at the national level. The first extensive evaluation concerning general upper secondary schools was published by the Finnish National Board of Education in 1994 (Jakku-Sihvonen and Blom, 1994). Since then there have been smaller evaluations or follow-ups concentrating mainly on the operation of the non-graded system (see e.g. Opetushallitus, 2000 and Opetushallitus, 2005).

Matriculation examination[14]

General upper secondary schooling concludes with a matriculation examination, which is a compulsory nationwide set of tests. The purpose of the examination is to determine whether students have assimilated the knowledge and skills required by the general upper secondary school curriculum. The examination is arranged in general

[14] The main sources are the Internet pages of the Matriculation Examination Board (www.ylioppilastutkinto.fi), the Finnish National Board of Education (www.oph.fi) and the Ministry of Education (www.minedu.fi).

upper secondary schools around the country. The Matriculation Examination Board is responsible for administering the examination, for preparing the tests and for the final assessment of the answer papers[15]. The results of each individual test are normalized to be comparable each year.

Matriculation examinations are arranged in the autumn and spring during a two-week examination period. In 1985-1995, all the tests had to be taken during one period. Since general upper secondary schooling was usually completed in three years, the majority of students took their tests during the spring term of the third year. In 1996 the matriculation examination underwent some changes following the structural changes taking place in general upper secondary schooling. Students were able to decentralize their test taking into three consecutive examination periods. The requirement for the test taking is the completion of all the compulsory courses in that subject.

Students can take individual tests in up to three consecutive examination periods. The examination consists of at least four tests. One of these, the test of the candidate's mother tongue, is compulsory for all candidates. The candidate then chooses three other compulsory tests from among the following four: the test in the second national language, a foreign language test, the mathematics test and the general studies test. As part of the examination, the candidate may additionally include one or more optional tests (a foreign language test, mathematics test or general studies test).

The intial grading of the tests is carried by the teachers of the school according to the guidelines of the Matriculation Examination Board. Associate members of the Matriculation Examination Board (external to the school) are responsible for the final grading of the tests. During 1985-1995 the grading followed a seven-step scale: improbatur, approbatur, lubenter approbatur, cum laude approbatur, magna cum laude approbatur and laudatur. In 1996 the scale was changed to eight

[15] The preliminary assessment takes place by the teachers in each school.

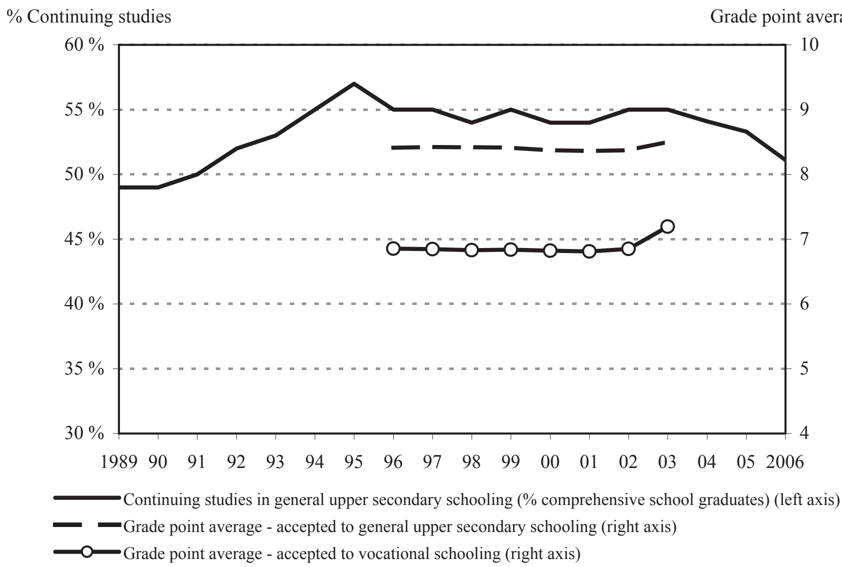
steps so that the highest grade, laudatur, was split into eximia cum laude approbatur and laudatur.

Finnish senior secondary schooling in numbers

The institutional context of Finnish general upper secondary schools was described above. In the following some general background statistics of schooling are briefly summarised in order to provide a somewhat broader view of the system than is given in each individual essay of this thesis.

The size of the age cohort, the labour market situation and attractiveness of vocational schooling affect the number of students in general upper secondary schooling. Some 55% of graduates in each age cohort continue their studies in general upper secondary schooling after comprehensive schooling. The share was slightly lower at the end of the 1980s, but gradually increased to 55% in the first half of the 1990s. Thereafter, it remained at this level until 2003. Due to increased attractiveness of vocational education, the proportion of students continuing their studies in general upper secondary schooling has decreased some five percentage points since 2003.

Figure 4. Percentage share of comprehensive school graduates continuing their studies in general upper secondary schooling in 1989–2006 (left axis) and the average grade point average of comprehensive school report of students accepted to general upper secondary schooling (right axis) and vocational schooling (right axis)



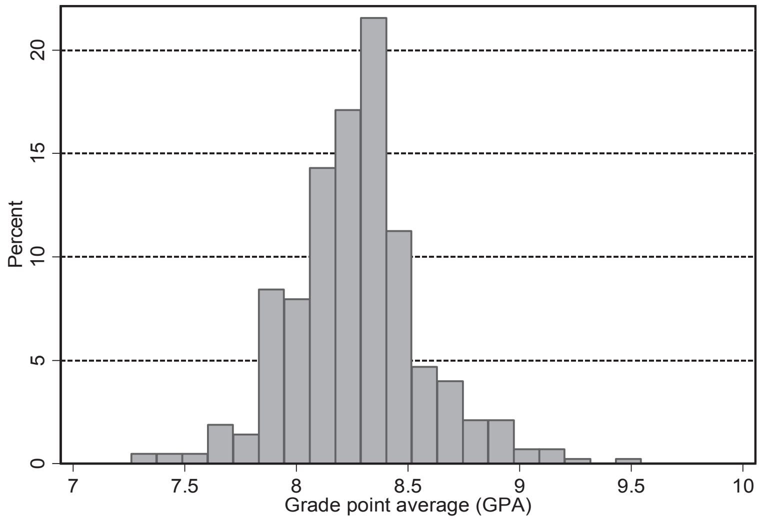
Students apply to general upper secondary and vocational education via a joint application system. Acceptance to a specific school is based on the student's comprehensive school grade point average. Competition for places in some of the most popular schools is especially fierce in general upper secondary education. Unfortunately, no national statistics are published on the results of the joint application system. Thus, no systematic information is available on the comprehensive school success of students accepted to general upper secondary and vocational schooling. Accordingly, there is no systematically published information on the competition between general upper secondary schools.

In two of the studies in this thesis, a joint application register containing information on students' grade point averages in comprehensive schooling[16] was available for the years 1996-2003. In Figure 4, this information is used to calculate a rough estimate of the average grade point average for students accepted to general upper secondary or vocational schooling directly after completing comprehensive schooling. As shown in Figure 4, there is a clear selection between these two schooling institutions so that students applying to general upper secondary schooling perform clearly better in comprehensive school than students applying to vocational schooling. The difference in grade point average (GPA) is some 1.6 points. The average grade point average in both schoolings remained quite stable during the eight year period. Only in 2003 was there a slight increase in both of them.

There is also clear selection between general upper secondary schools. In Figure 5 the distribution of average grade point averages of schools is depicted for those students that completed their studies in 2004. In majority of the schools the average GPA varied between 7.8 and 8.6, but in some schools the average GPA of accepted students exceeded 9. In the lower tail of the distribution there are schools that have an average GPA of 7.3.

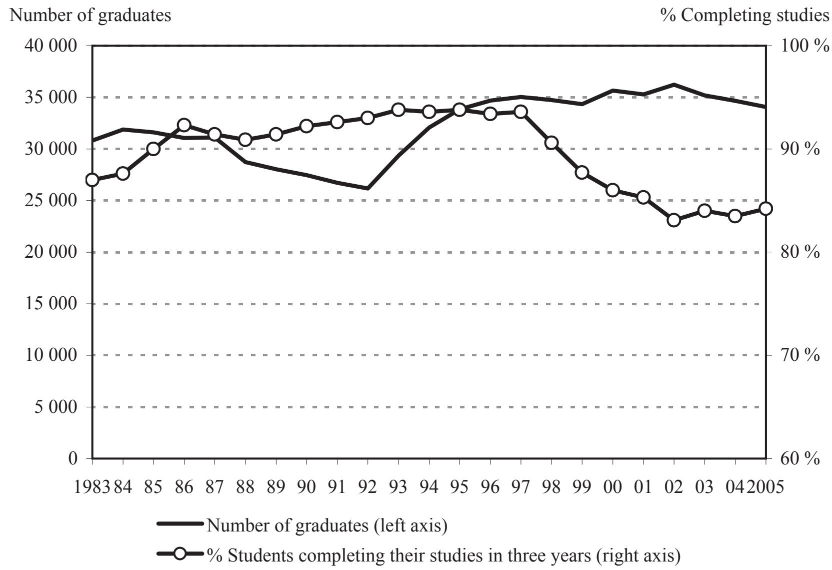
[16] Grades for each subject are given by teachers and range from 4 to 10.

Figure 5. The distribution of average GPA in general upper secondary schools based on information for students completing their studies in 2004



In 1985 there were some 115 000 students in general upper secondary schools (see Figure 7; Statistics Finland). Towards the end of the 1980s the number of students decreased, but the severe recession at the beginning of the 1990s increased the total to almost 130 000 students. The peak was reached in 2000 and thereafter there has been a moderate decrease. General upper secondary education has been more popular among female students, since over half of students are female. The proportion of the female students has remained quite stable, and in 2004 it was 57%.

Figure 6. *Number of graduates (left axis) and proportion of students completing their studies in three years (right axis) in 1983–2005*



The number of matriculating students follows the number of students in general upper secondary schools, since the proportion of students failing in separate tests is kept stable across years and the dropout rate is low. In 1990, some 27 500 students passed the matriculation examination, while in 2005 this figure was 34 000 students (see Figure 6).[17] The proportion of students discontinuing their studies is quite low in general upper secondary schooling compared to vocational and tertiary level education in Finland. It has also remained quite stable over time. During the 1980s it increased from some 5% up to 7% in the school year 1989-90 (Statistics Finland; Ojala, 1994). Thereafter it decreased and during the first half of the 2000s it remained at approximately 4% (Statistics Finland). Male students discontinue their studies more often than female students (Statistics Finland).

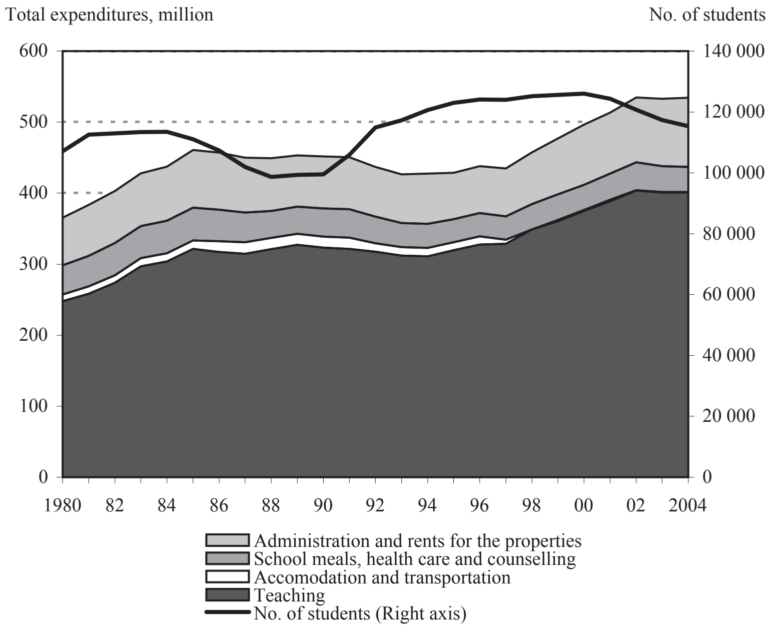
[17] These numbers include adult students, who account for approximately 5% of the student population. Unfortunately, it was not possible to obtain figures for young students only.

Teaching arrangements have clearly influenced the length of studies, even though the rate of discontinuing studies has remained unaffected. During the period of fixed classes, over 90% of students completed their studies in three years (see Figure 6). In the school year 1985-86 this figure was 92%, and in the school year 1991-92 some 93% (Statistics Finland; Ojala, 1994). Since year classes were abolished in 1996 and students started following individual study plans the increased freedom and responsibility has caused a clear increase in the length of studies. The peak was reach in 2002, when some 17% of graduating students studied for over 3 three years.

The number of general upper secondary schools has only recently been decreasing. Before that the situation was quite stable. In 1982 the number of schools was 464, while in 1993 there was only one school less (Ojala, 1994). In 2004 the number of schools was 436 and by 2005 it had decreased to 428 (Statistics Finland). The average size of schools is quite small by international standards (Figure 8). It has fluctuated between 225 (in 1989) and 275 (in 1998) students. In 2004 there were on average some 250 students per general upper secondary school. However, the variation in size is quite considerable: in 2004 it varied from 30 to 900.

There was an increasing trend in total expenditures in general upper secondary schooling over the period of 1980-2004 (see Figure 7), even though expenditures slightly decreased during the recession as a consequence of the cutting of state subsidies to organizers of general upper secondary schooling. The highest increase in spending took place after the recession in the second half of the 1990s. Teaching expenditures constitute the largest part of the total expenditures, with a share of some 75%. Administration costs and rents for school properties are the second largest element, with a share of about 20%. Costs of school meals (students are served a warm school meal during the school day), health care and counselling are the smallest elements in expenditures.

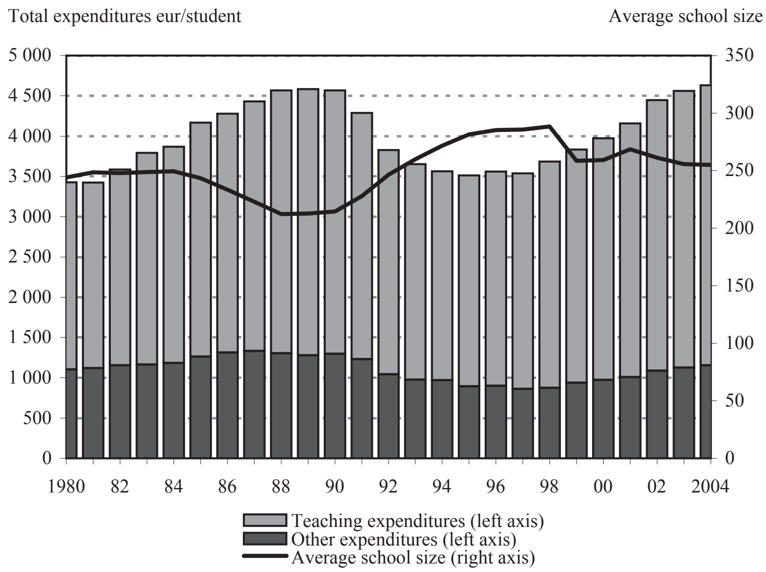
Figure 7. *Total expenditures and number of students in general upper secondary schooling in 1980–2004*



Expenditures deflated to year 2003 prices.

Expenditures per student have considerably fluctuated following the economic cycles and the development of number of students (see Figure 8). In the 1980s they constantly increased, reaching a peak at the end of the decade. As a consequence of the depression and spending cuts in 1991-1995, expenditures fell to almost the same level as at the beginning of the 1980s. After the recession, expenditures per student started growing again and by 2004 they were at the same level as before the recession. The falling number of students, especially in rural areas, was one cause of the rising expenditures per student after 1995 (Kirjavainen, 2005).

Figure 8. *Teaching expenditures and other expenditures per student in general upper secondary schooling in 1980–2004*



Expenditures deflated to year 2003 prices.

4. Content of the thesis

This thesis is comprised of five essays that investigate the efficiency of schools, factors affecting efficiency and the effect of resources on student achievement. More specifically, this thesis addresses the following questions:

How large are the efficiency differences between schools?

How stable are the efficiency differences across different input and output variables?

How stable are the efficiency differences across different model specifications?

Do some factors that are related to resources, students and organizational practices of schools explain the efficiency differences statistically?

What organizational practices characterize efficient and inefficient schools?

What is the impact of school resources and the student-teacher ratio on student achievement?

In the study described in the *first essay*, the efficiency differences between Finnish general upper secondary schools were assessed with Data Envelopment Analysis (DEA) and the efficiency differences were explained by the scale of operation, location of the school and financial situation of the municipality using statistical Tobit models. The study aimed at testing the stability of the results of DEA models by changing the scale assumption and input and output variables, and by dropping schools at the frontier (jackknifing). It also tested the stability of the results in models explaining efficiency differences by varying the dependent and independent variables.

The empirical data comprised a cross-section of Finnish general upper secondary schools (291 schools) for students completing their studies in 1991. The data contained information on student background, school resources, teachers' education and experience, the number of students, state grants, location, and grades in the matriculation examination. The resources were measured as a three-year average when this was available.

The results of the study showed that the efficiency frontier was relatively stable with regard to outlier observations. Especially in the models assuming constant returns to scale (Charnes *et al*., 1978), the efficiency frontier and the rankings of schools measured using the Spearman's rank correlation coefficient remained almost unchanged.

In models assuming variable returns to scale (Banker *et al*., 1984), there was more variation.

The change in input and output variables affected the average efficiency and the rankings based on efficiency scores. The classification of schools into quartiles according to their efficiency scores, assuming constant returns to scale, revealed that schools in the upper and lower tail remained unchanged, whereas schools in the middle of the distribution often changed their ranking. When the assumption of variable returns to scale was applied, the rankings were more unstable.

As a second stage, the inefficiency scores (1 - efficiency) assuming constant returns to scale were explained using Tobit models with variables related to the school student body, scale of operation, state grants and school location. Two sets of models were estimated. In the first, parental educational level was not included in DEA model but was used as an explanatory variable, while in the second set it was included in DEA models and the models were estimated with the other independent variables.

Schools with a higher level of parental education and larger class sizes had higher efficiency in all models. School size, the proportion of female students, heterogeneity of the student body, school location and variables related to the grant system were statistically insignificant in explaining inefficiency differences in models using parental educational level as an explanatory variable in Tobit models. The inclusion of the parents' educational level as an input in the DEA model somewhat changed the results. Schools with a less heterogeneous student body were more efficient and those in urban areas were less efficient than schools in densely populated and rural areas. Private schools turned out to be less efficient than public schools.

Tobit-model analyses were also carried out using the efficiency scores of variable returns to scale as the dependent variable and dropping out the variables controlling for the scale of operation (i.e.

class and school size). According to results, efficiency decreased as the heterogeneity of the student body increased. Contrary to earlier results, public schools were less efficient than private schools.

The results of this essay show the importance of selection of input and output variables in DEA efficiency analysis. Average efficiency and the rankings of the schools vary depending on the input and output variables used. The instability of the results of Tobit models related to the inclusion of variables in DEA models as compared to using them as explanatory factors raises questions about the reliability of the results of DEA. Parental educational level seems to be important for efficiency. As for other factors, their effect is less clear and is dependent on the DEA model used. Thus, there is a clear need for further research with alternative methods and data.

In the *second essay* of this thesis, I further analyze the schools at the both ends of the efficiency distribution, concentrating on their organizational characteristics, with survey data that were collected for the study. The study presented in this essay aimed at identifying factors related to organizational practices that would characterize efficient and inefficient schools.

The efficiency measurement using DEA was based on cross section data on Finnish general upper secondary schools (309 schools) with students completing their studies in 1991. The data contained information on student background, school resources, teachers' education and experience, and grades in the matriculation examination. Organizational factors affecting efficiency were studied with a survey data. The survey consisted of responses of principals of 73 schools, of which 40 were among the most efficient and 33 among the least efficient ones in at least one of three DEA models. The survey was conducted in the spring of 1995. It gathered information on the students, parent-school relations, management, decision making, staff development, student monitoring, climate and cooperation, development of instruction, and school facilities. Efficiency differences were examined with descriptive statistics and statistical probit models. Tobit models were used to test the stability of the results.

The results of this study demonstrated that very few organizational practices differentiated schools with high and low efficiency. Schools with a management group or other similar body were more likely to be efficient. In addition, a school's probability of being efficient increased the fewer teachers participated in training, the more there were clubs in the school, and the less there were joint teaching projects across the subject. For the most part, these results were not robust to model specifications. As for factors related to students, the admittance level and heterogeneity of the student body, which were not taken into account in efficiency measurement, were the main factors contributing to efficiency. The results of this study revealed the uncertainties that are related to DEA efficiency scores based on cross section data.

The *third essay* examines the efficiency differences of Finnish general upper secondary schools with stochastic frontier models. In the study, five-year panel data were used and advantages of the data exploited by taking into account the school-specific random or fixed effects that are constant through time to separate the unobserved heterogeneity (Greene 2005a, b) from inefficiency. The objective of the study was to test the stability of average efficiency and efficiency rankings between the different stochastic frontier models.

The study also aimed at testing the effect of expenditures, the student teacher ratio, study length and decentralization in test taking on student achievement with models accounting for inefficiency. These two latter factors have not previously been examined in the Finnish context, and international evidence is also limited. Both the length of studies and decentralization of test taking have increased following the reform of general upper secondary education in the mid-1990s which abolished fixed classes and allowed the possibility to take tests in the matriculation examination over up to three consecutive test-taking periods.

Five-year panel data on 436 Finnish general upper secondary schools for the years 2000-2004 were used in the study. Besides earlier school achievement, family background, teaching expenditures and the student-teacher ratio, the data included information on the average length of studies and decentralization of test taking in the matriculation examination.

The study demonstrated that average inefficiency clearly diminishes as the school-specific random or fixed effects capturing the uncontrolled heterogeneity are separated from inefficiency. The study also showed that random effects in the conventional random effects model are very similar to those in models separating random effects from inefficiency (i.e. true random effect model). Only the magnitude is smaller. Hence, inefficiency in true random effects models depicts a small effect that is due to yearly fluctuations in output and input variables. The same pattern emerged in fixed and true fixed effects models.

Inefficiency has been suggested as one reason for the mixed results concerning the effect of school resources on student achievement. In this study, despite taking inefficiency into account, the results did not differ from those in earlier research. The effect of teaching expenditures per student did not have a statistically significant effect on matriculation examination scores. Models taking into account efficiency and correcting for omitted variable bias (true random and fixed effects models) produced a negative correlation between student achievement and teaching expenditures per student. Models replacing teaching expenditures per student with the student-teacher ratio gave similar results, since the larger the student-teacher ratio is, the higher are the grades in the matriculation examination.

A longer average length of school studies was related to lower achievement. Equally, the more students decentralized their test taking into different test-taking periods in the matriculation examination, the worse they performed. These results were somewhat surprising, but it might be the case that students studying for longer are also working or have other activities that occupy their time and reduce the time used for studying. The negative impact of decentralization of test taking may be the result of students going to the test less well prepared than is the case when there is no possibility to retake the test. These results were obtained with school-level data and they therefore ignore the between-student variation. To gain further certainty and information on this matter, student-level data should be used.

In the *fourth essay*, the results of efficiency analysis are further investigated by taking a different methodological approach. The organizational practices of nine case schools that are differentially efficient were assessed. The focus of the study was on practitioners' views, i.e. how principals and teachers perceived and described students and practices in their school. The aim of the study was to examine whether there were differences in organizational practices and characteristics of schools and whether these differences were related to differences in efficiency.

The study was based on interview data for principals and teachers in nine case schools. In each school, the principal and two teachers were interviewed in 1999. The respondents were unaware of their school's efficiency. The interviews covered themes concerning the students, staff relations, school goals, school management and decision making, curriculum, teacher training, the evaluation of students and parent-school relations. The identification of efficiency differences was based on the results of the third essay of this thesis.

The schools differed with respect to four themes, namely the views on the students, staff relations, school governance and the curriculum. With the remaining themes the schools were quite similar. Within each theme, schools were categorized into groups and each group was described and characterized in the study. Efficiency was also considered, and it emerged that the categories were also related to efficiency.

Teachers' and principals' views about their students have more seldom been investigated. The results of this study showed that in efficient schools, respondents' views about the students were attentive or respecting. Attentive views emphasized the importance of taking care of all the students and especially those with low performance. Their success was accepted with pride. The respecting view involved seeing students as independent and grown-up decision makers responsible for themselves. In inefficient schools, views were more often educated or frustrated. Educated views were somewhat indifferent towards the students. The low performing students were seen as those whose place is not in a general upper secondary school. There was less talk about taking care of all students. Schools with frustrated views were somewhat hopeless concerning the low skills of the students. There was some caring involved, but the problems in some cases were found so severe that the teachers also felt exhausted.

Other results of this study were very similar to many school effectiveness studies. In efficient schools staff relations were good and they were characterized as professional. Interaction was uncomplicated and there was a sense of collegiality so that teachers were sharing matters related to instruction and students. Staff members were happy with the situation. Schools with problems in staff relations were more often inefficient in this study. These problems were openly admitted during the interviews and they somewhat complicated the interaction between staff members. Teachers also felt to some extent dissatisfied with the situation.

Management and decision making were characterized in most of the schools as participative. The roles of the school head and teachers were clear and accepted. The school head was assisted by a management group or corresponding working groups. Teachers were very autonomous with their classroom work. Decision making was carried out jointly with the head of the school and the teachers. Teachers felt that they were able to influence school matters and that their views were listened to and taken into account. In one inefficient school, management and decision making was described as hierarchical. In that school, the principal had centralized most of the decision making to herself. The role of the vice principal was limited and there were no permanent working groups assisting the principal. Teachers were independent in their classroom work but they felt unable to influence school-level decision making.

In efficient schools, the curriculum and curriculum work were more often seen as a way to develop the school and the school work. In such a case the curriculum also developed so that teachers added new courses and removed old ones. The teachers had realized that the school curriculum was actually a way to develop their own work. Updating the curriculum had also opened up a possibility for productive discussions. In some schools, this development was still ongoing and the curriculum was seen to some extent as an administrative tool without practical relevance. In inefficient schools, the school curriculum was usually seen as an administrative tool. In schools with problems in staff relations there were also problems in curriculum work and teachers expressed their discontent with the work.

The *fifth essay* focuses on the effect of school resources on student achievement. The study aimed at examining how a considerable decrease in schooling expenditures due to the severe recession of the Finnish economy during the research period affected student achievement. Nine-year panel data were used covering the years 1990-1998 and matching students with schools. With the use of two-way fixed effects models with school and time fixed effects, the unobserved time constant heterogeneity could be controlled for and the omitted variable bias reduced.

The data comprised a sample of 20 505 students in 444 schools and included information on the students' earlier school achievements, gender, parental education and working while at school. School site information consisted of teaching expenditures, school size, regional unemployment and the location of the school.

The results showed that changes in school expenditures per student did not affect student performance in the matriculation examination. The effect was also tested with other outcome variables and panel data models, but the results were not sensitive to these changes. Hence, the results were consistent with those of many previous studies. Earlier achievements and parental educational level were the major determinants of student performance. Males performed slightly better than females according to results of some of the models, and working during the school year decreased performance.

5. Main findings and conclusions

The results of this thesis demonstrate that there are efficiency differences between Finnish general upper secondary schools. Students in some schools learn more after controlling for their schooling and family background as well as the resources of the school. At the beginning of the 1990s the average potential for efficiency improvement was some 6%. The estimate for the beginning of the 2000s was somewhat lower. There was variation across schools and in some schools the potential for efficiency improvement was at least15% in the latter period. This means that in some schools, average grades in the matriculation examination could have been 15% higher given their intake of students and school resources.

The identification of efficiency differences is only the beginning. They raise the question of why some schools are more efficient than others. Policy makers are also interested in knowing how to increase efficiency. As in earlier studies, factors related to student characteristics seem to be important. Earlier school achievement and family background are strong predictors of student achievement, and when they are taken into account in efficiency analysis, efficiency differences clearly diminish. Heterogeneity in the skills of students decreases efficiency and student achievement. In one of the essays, evidence was found that heterogeneous classes complicate the instruction and the work of teachers. Hence, in improving the efficiency of general upper secondary schooling, homogeneity of skills and knowledge of admitted students would be beneficial. This issue should already be addressed at the primary school level.

Homogeneity in the student body can also be achieved by increasing the segregation across schools. There is some evidence of this kind of development taking place in general upper secondary schooling in recent decades. Segregation across schools is, however, a somewhat more complicated issue. Evidence was found that teachers perceived high performing students to act as role models and have a positive effect on the performance of low achieving students. In schools with a lower proportion of high performing students, these role models were mostly lacking and this affected the motivation and learning of students. The work of teachers also became more demanding and the teachers sometimes felt frustrated.

Reducing the resources of schools and increasing group sizes is a conventional way of improving efficiency. In this thesis, the results of the effect of teaching expenditures and class size on grades in the matriculation examination were mixed. No evidence was found that the exceptionally large fall in teaching expenditures following the recession at the beginning of the 1990s affected student achievement. During the period of rising teaching expenditures in 2000-2004, schools with lower teaching expenditures per teacher performed better. In line with these results, the student-teacher ratio had a positive effect on grades in the matriculation examination. The latter results were somewhat dependent on the model specification.

Concerning the consequences of reforms in general upper secondary schooling that took place a decade ago, it seems that the increased freedom in studies introduced along with them should be reconsidered, at least from an efficiency point of view. According to the results of this thesis, performance was lower in schools in which students studied for longer or decentralized their test taking. Hence, instructional arrangements affect student achievement and therefore also efficiency. Tighter control of studies could improve efficiency.

As for organizational characteristics, numerous factors were statistically tested but very few of them distinguished efficient and inefficient schools. There are several possible explanations for this result. It could partly be a consequence of the method used in efficiency analysis, and may partly reflect a more general finding of many school effectiveness studies that organizational factors account for only a small part of the variance in student achievement. Behind these results could also be the fact that Finnish general upper secondary schools were quite tightly regulated during the period of the analysis, creating homogeneous practices.

With the case study approach it was possible to go slightly deeper into the organizational practices of certain schools. This approach revealed that the studied Finnish general upper secondary schools were, indeed, fairly similar in many respects, but that there were also some differences. These were related to staff members' views about the students, staff relations, school management and curriculum work. There was also some indication that these differences were related to school efficiency. According to these results, efficiency improvements are possible if schools consider all students as important, have professional staff relations and participative governance and management, and use the school curriculum as a development tool.

Deepening of the case study analysis is one way to confirm the results based on interviews with the principal and some teachers in each case school. The approach adopted in this study was a good starting point, but during the research process it became evident that it was not sufficient. To draw a more complete picture of the schools, more intense research approaches are needed. This includes a comparative research setting and ethnographic research methods with participant observation, longer follow-up periods and the collection of historical data. In addition to using staff members of the school as informants, the views of students and parents should be taken into account. With such data we could try to understand how different people and their interaction affect the operation of the school. We could also examine the paths that lead a school to perform efficiently or inefficiently.

As for the reliability of the efficiency measurement, the level of efficiency was sensitive to the method of analysis and model specification. The rankings of schools based on efficiency scores also varied depending on the method and model specification. This complicates the second stage analysis. The problem is likely to be more severe when data envelopment analysis is used. In this method, the discriminating power decreases as the number of inputs and outputs increases. This probably explains why the results of the second stage analysis conducted with statistical models in the second essay were unstable. The problem applied both to models taking the whole efficiency distribution as an explanatory variable and those examining schools in the upper and lower tails of the distributions.

However, the problem was more severe when the ends of the distribution were considered.

The findings of this thesis cast some doubts about the applicability of results of DEA. Some earlier studies have only examined the organizational practices of efficient schools and found some common characteristics shared by these schools. If the same characteristics are nevertheless also present in inefficient schools, improving efficiency becomes a rather complicated exercise. Hence, one has to be cautious when basing policy recommendations on such results.

In stochastic frontier analysis, the results also vary depending on the model. The variation is partly related to the concept of inefficiency, which differs between models. In more advanced models based on panel data, time constant heterogeneity is separated from inefficiency that fluctuates over time. This difference is very important to take into account when interpreting the results. It also influences the second stage analysis seeking to explain the efficiency differences. Inefficiency that depicts small yearly fluctuations in inputs and output is not perhaps easily explained with factors that remain constant through time, such as many organizational characteristics. Therefore, the use of a time-constant effect to measure inefficiency is more relevant in such a context. An inefficiency term capturing yearly fluctuation is more suitable as a short-term goal to improve efficiency.

Time gaps in the various sources of information used in efficiency analysis are one problem that is difficult to avoid when investigating the causes of efficiency differences between schools. The identification of efficiency differences is usually based on register data and there is often a time lag before the data are available for research. Only seldom are good quality data available that can be used in the second stage to explain efficiency differences. Instead, additional data has to be gathered. This retrospective research setting brings additional uncertainty to the interpretation of the second stage results. This is a fact that one should already be aware of when planning the research. The gathering of extensive and multifaceted data is one way to reduce this problem.

To conclude, the separate studies of this thesis highlight the importance of examining school efficiency and student achievement with different approaches. Only in this way it is possible to deepen our understanding of the methods and their applicability to different situations as well as processes and operation of schools.

References

Aaltonen, J. – Kirjavainen, T. – Moisio, A. (2006): Efficiency and Productivity of Finnish Comprehensive Schooling 1998-2004. *VATT-research reports* 127. Helsinki: Government Institute for Economic Research.

Aigner, D. J. – Lovell, C. A. K. – Schmidt, P. (1977): Formulation and Estimation of Stochastic Frontier Production Function Models. J*ournal of Econometrics* 6, 21–37.

Akerhielm, K. (1995): Does Class Size Matter? *Economics of Education Review* 14, 229 –241.

Akerlof, G. A. – Kranton, R. E. (2002): Indentity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature* 40, 1167–1201.

Angrist, J. D. – Lang, K. (2004): Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review* 94, 1613 –1634.

Angrist, J. D. – Lavy, V. (1999): Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics* 114, 533–575.

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

Banker, R. D. – Morey, R. (1986): The Use of Categorical Variables in Data Envelopment Analysis. *Management Science* 32, 1613–1627.

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

Battese, G. – Coelli, T. (1992): Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis* 3, 153–169.

Battese, G. – Coelli, T. (1995): A Model for Technical Efficiency Effects in a Stochastic Frontier Production Function for Panel Data. *Empirical Economics* 20, 325–332.

Becker, G. S. (1975): *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, 2nd ed. New York: National Bureau of Economic Research.

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

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

Bidwell, C. E. – Kasarda, J. D. (1980): Conceptualizing and Measuring the Effects of School and Schooling. *American Journal of Education* 88, 401–430.

Bingley, P. – Myrup Jensen V. – Walker, I. (2007): The Effects of School Resources on Participation in Post-Compulsory Education: Danish Quasi-Experimental Evidence, and Evidence That Controls for Family, School, and Neighbourhood Effects. Http://www2.warwick.ac.uk/fac/soc/economics/staff/faculty/walker/current_research

Betts, J. (2001): The Impact of School Resources on Women's Earnings and Educational Attainment: Findings from the National Longitudinal Survey of Young Women. *Journal of Labour Economics* 19, 635–657.

Black, S. E. – Devereux, P. J. – Salvanes, K. G. (2005): The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *The Quarterly Journal of Economics* 119, 669–700.

Bonesrønning, H. (2003): Class Size Effects on Student Achievement in Norway: Patterns and Explanations. *Southern Economic Journal* 69, 952–965.

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

Bradley, S. – Johnes, G. – Millington, J. (2001): The Effect of Competition on the Efficiency of Secondary Schools in England. *European Journal of Operational Research* 135, 545–568.

Brookover, W. B. – Beady, C. – Flood, P. - Schweitzer, J. – Wisenbaker, J. (1979): *Schools, Social Systems and Student Achievement: Schools Can Make a Difference.* New York: Praeger.

Brown, B. W. – Saks, D. H. (1987): The Microeconomics of the Allocation of Teachers' Time and Student Learning. *Economics of Education Review* 6, 319-332.

Brown, B. W. – Saks D. H. (1975): The Production and Distribution of Cognitive Skills within Schools. *Journal of Political Economy* 83, 571–593.

Browning, M. – Heinesen, E. (2007): Class Size, Teacher Hours and Educational Attainment. *Scandinavian Journal of Economics* 109, 415–438.

Burstein, L. (1980): The Analysis of Multilevel Data in *Educational Research and Evaluation in Review of Research in Education vol. 8*. Washington DC.: American Educational Research Association.

Chakraborty, K. – Biswas, K. – Lewis, C. (2001): Measurement of Technical Efficiency in Public Education: A Stochastic and Nonstochastic Production Function Approach. *Southern Economic Journal* 67, 889–905.

Charnes, A. – Cooper, W. W. – Lewin, A. Y. – Seiford, L. M. (1994): *Data Envelopment Analysis: Theory, Methodology, and Application.* Boston: Kluwer Academic Publishers.

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

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

Coe, R. – Fitz-Gibbon, C.T. (1998): School Effectiveness Research: Criticism and Recommendations. *Oxford Review of Education* 24, 421–438.

Coelli, T. (1995): Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis. *Journal of Productivity Analysis* 6, 247–268.

Coelli, T. – Rao, P.D.S. – Battese, G. E. (1999): *An Introduction to Efficiency and Productivity Analyses*. Kluwer Academic Publishers. Boston.

Coleman, J. S. – Campbell, E. Q. – Hobson, C. F. – McPartland, J. – Mood, A. M. – Weifeld, F. D. – York, R. L. (1966): *Equality of Educational Opportunity.* Washington: U.S. Government Printing Office.

Cooper, S. T. – Cohn, E. (1997): Estimation of a Frontier Production Function for the South Carolina Educational Process. *Economics of Education Review* 16, 313–327.

Conroy, S. J. – Arguea, N. M. (2008): An Estimation of Technical Efficiency for Florida Public Elementary Schools. *Economics of Education Review* 27, 655 –663.

Cornwell, C. – Schmidt, P. – Sickles, R. (1990): Production Frontiers with Cross Sectional and Time Series Variation in Efficiency Levels. *Journal of Econometrics* 46, 185–200.

Correa, H. (1993): An Economic Analysis of Class Size and Achievement in Education. *Education Economics* 1, 129–135.

Cunha, F. – Heckman, J. J. (2008): Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *The Journal of Human Resources* 43, 738–782.

Cunha, F. – Heckman, J. J. – Lochner, L. – Masterov, D. V. (2006): Interpreting the Evidence on Life Cycle Skill Formation. In Hanushek, E. A. – Welch, F. (eds.) *Handbook of the Economics of Education*. Amsterdam: Elsevier.

Dee, T. S. (2007): Teachers and the Gender Gaps in Student Achievement. *The Journal of Human Resources* 42, 528–554.

Deller, S. C. – Rudnicki, E. (1993): Production Efficiency in Elementary Education: The Case of Maine Public Schools. *Economics of Education Review* 12, 45–57.

Dobbelsteen, S. – Levin, J. – Oosterbeek, H. (2002): The Causal Effect of Class Size on Scholastic Achievement: Distinguishing the Pure Class Size Effect from the Effect of Changes in Class Size Composition. *Oxford Bulletin of Economics and Statistics* 64, 17–38.

Dodd, A. (2006): Investigating the Effective Use of Resources in Secondary Schools. *Research Report to DfES*. Reserch Report No 779. Hedra plc, London.

Duncombe, W. – Miner, J. – Ruggiero, J. (1997): Empirical Evaluation of Bureaucratic Models of Inefficiency. *Public Choice* 93, 1–18.

Ehrenberg, R. G. – Brewer, D. J. (1994): Do School and Teacher Characteristics Matter? Evidence from High School and Beyond. *Economics of Education Review* 13, 1–17.

Eren, O. – Millimet, D. L. (2007): Time to Learn? The Organizational Structure of Schools and Student Achievement. *Empirical Economics* 32, 301–332.

Evans, W. – Wallace, E. O. – Schwab, R. M. (1992): Measuring Peer Group Effects: A Study of Teenage Behavior. *Journal of Political Economy* 100, 966–991.

Farrell, M. J. (1957): The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society* 120, 253–281.

Fertig, M. (2003): Educational Production, Endogenous Peer Group Formation and Class Composition. Evidence from the PISA 2000 Study. Institute for the Study of Labor (IZA), Bonn, *Discussion Paper* no 714.

Fried, H. O. – Lovell, C. A. K. – Schmidt, S. S. (Eds.) (2008): *The Measurement of Productive Efficiency and Productivity Growth*. New York: Oxford University Press.

Fuchs, T. – Wössmann, L. (2007): What Accounts for International Differences in Student Performance? A Re-examination Using PISA Data. *Empirical Economics* 32, 433–464.

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

Goldstein, H. (2003): *Multilevel Statistical Models*. 3rd ed. London: Edward Arnold.

Grosskopf, S. – Hayes, K. – Taylor, L. – Weber, W. (2001): On the Determinants of School District Efficiency: Competition and Monitoring. *Journal of Urban Economics* 49, 453–478.

Goldhaber, D. – Anthony, E. (2007): Can Teacher Quality Be Effectively Assessed? National Board Certification as a Signal of Effective Teaching. *The Review of Economics and Statistics* 89, 134–150.

Goldhaber, D. D. – Brewer, D. J. (1997): Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity. *The Journal of Human Resources* 32, 505–523.

Gray, J. – Hopkins, D. – Reynolds, D. – Wilcox, B. – Farrell, S. – Jesson, D. (1999): *Improving Schools. Performance and Potential*. Buckingham: Open University Press.

Greene, W. (2005a): Fixed and Random Effects in Stochastic Frontier Models. *Journal of Productivity Analysis* 23, 7–32.

Greene, W. (2005b): Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model. *Journal of Econometrics* 126, 269–303.

Hægeland, T. – Raaum, O. – Salvanes, K. J. (2007): Peannis from Haeven: Using Exogenous Tax Variation to Identify Effects of School Resources on Pupil Achievement. *Discussion Papers* No. 508. Statistics Norway, Research Department.

Hanushek, E. A. (2003): The Failure of Input-Based Schooling Policies. *The Economic Journal* 113, F64–F98.

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. – Kain, J. F. – Markman, J. M. – Rivkin, S. G. (2003): Does Peer Ability Affect Student Achievement? *Journal of Applied Economics* 18, 527–544.

Hanushek, E. A. – Kain, J. F. – Rivkin, S. G. (2004): Why Public Schools Loose Teachers. *Journal of Human Resources* 39, 326–254.

Hanushek, E. – Rivkin, S. (2006): Teacher Quality. In Hanushek, E. – Welch, F. (Eds.) *Handbook of the Economics of Education*, Vol. 2. Amsterdam: Elsevier.

Hedges L. – Greenwald, R. (1996): Have Times Changed? The Relation between School Resources and Pupil Performance. In Burtless, G. (ed.) *Does Money Matter? The Effect of School Resources on Pupil Achievement and Adult Success*. Washington DC: Brookings Institution Press.

Heshmati, A. – Kumbhakar, S. C. (1997): Efficiency of the Primary and Secondary Schools in Sweden. *Scandinavian Journal of Educational Research* 41, 33–52.

Houtenville, A. J. – Conway, K. S. (2008): Parental Effort, School Resources, and Student Achievement. *The Journal of Human Resources* 43, 437–453.

Hoxby, C. M. (2000): The Effects of Class Size and Composition of Student Achievement: New Evidence from Natural Population Variation. *Quarterly Journal of Economics* 115, 1239–1285.

Häkkinen, I. – Kirjavainen, T. – Uusitalo, R. (2003): School Resources and Student Achievement Revisited: New Evidence from Panel Data. *Economics of Education Review* 22, 329–335.

Jenkins, A. – Vignoles, A. – Levačić, R. – Allen, R. (2006): The Effect of Resources on Pupil Attainment in English Secondary Schools. Paper presented at the European Conference on Educational Research (ECER), Geneve.

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

Johnes, G. – Johnes, J. (2009): Higher Education Institutions' Costs and Efficiency: Taking the Decomposition a Further Step. *Economics of Education Review* 28, 107–113.

Johnes, J. (2006): Measuring Efficiency: A Comparison of Multilevel Modelling and Data Envelopment Analysis in the Context of Higher Education. *Bulletin of Economic Research* 58, 75–104.

Johnes, J. (2004): Efficiency Measurement. In Johnes, G. and Johnes, J. (eds.) *International Handbook on the Economics of Education*. Northampton: Edward Elgar Publishing.

Kang, B-G. – Greene, K. V. (2002): The Effects of Monitoring and Competition on Public Education Outputs: A Stochastic Frontier Approach. *Public Finance Review* 30, 3–26.

Kangasharju, A. (Ed.) (2007): The Productivity of Welfare Services: Lessons Learned (in Finnish). *VATT-Julkaisuja* 46. Helsinki: Valtion taloudellinen tutkimuskeskus..

Kirjavainen, T. (2007a): The Efficiency of Finnish Upper Secondary Schools in 2000–2004 (in Finnish). *VATT-Tutkimuksia* 131. Helsinki: Valtion taloudellinen tutkimuskeskus.

Kirjavainen, T. (2007b): Efficiency of Finnish Upper Secondary Schools: An Application of Stochastic Frontier Analysis with Panel Data. *VATT-Discussion Papers* 428. Helsinki: Government Institute for Economic Research.

Kirjavainen, T. (2005): On Cost Effectiveness of Upper Secondary Schools in 1980–2005 (in Finnish). *Moniste* 23/2005. Helsinki: Opetushallitus.

Kirjavainen, T. (1999): Efficiency of Finnish Senior Secondary Schools: Explaining High and Low Performance with Resources, Environmental, and Organizational Factors. *Helsinki School of Economics and Business Administration Working Papers* W-221. Helsinki: Helsinki School of Economics and Business Administration.

Kirjavainen, T. – Loikkanen, H. A. (1998): Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit-Analysis. *Economics of Education Review* 17, 377–394.

Kirjavainen, T. - Loikkanen, H. A. (1993): On Efficiency Differences of Finnish Upper Secondary Schools (in Finnish). *VATT-Tutkimuksia* no. 16. Helsinki: Valtion taloudellinen tutkimuskeskus.

Kumbhakar, S. C. – Lovell, K. C. A. (2000): *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

Kuosmanen, T. (2006): Stochastic Nonparametric Envelopment of Data: Combining Virtues of SFA and DEA in a Unified Framework. *MTT-discussion Paper* No. 3/2006.

Krueger, A. (2003): Economic Considerations and Class Size. *The Economic Journal* 113, F34–F63.

Krueger, A. (1999): Experimental Estimates of Education Production Functions. *The Quarterly Journal of Economics* 114, 497–532.

Krueger, A. – Whitmore, D. M. (2001): The Effect of Attending Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence From Project Star. *The Economic Journal* 111, 1–28.

Kumbhakar, S. C. (1990): Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency. *Journal of Econometrics* 46, 201–212.

Kumbhakar, S. C. – Lovell, K. C. A. (2000): *Stochastic Frontier Analysis.* Cambridge: Cambridge University Press.

Lazear, E. (2001): Educational Production. *Quarterly Journal of Economics* 113, 777–803.

Lee, V. E. – Bryk, A. S. – Smith, J. B. (1993): The Organization of Effective Secondary Schools in L. Dardling Hammond (ed.) *Review of Research in Education* 19. Washington DC.: American Educational Research Association.

Lee, Y. H. – Schmidt, P. (1993): A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency. In Fried H. O. – Lovell, C. A. K. – Schmidt, S. S., eds. *The Measurement of Productive Efficiency: Techniques and Applications.* New York: Oxford University Press.

Levačić, R. – Jenkins, A. – Vignoles, A. – Steele, F. – Allen, R. (2005): Estimating the Relationship Between School Resources and Pupil Attainment at Key Stage 3. Report to DfES, January 2005. Institute of Education. University of London.

Levačić, R. – Vignoles, A. (2002): Researching the Links between School Resources and Student Outcomes in the UK: A Review of Issues and Evidence. *Education Economics* 10, 313–331.

Levin, J. (2001): For Whom the Reduction Count? A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement. *Empirical Economics* 26, 221–246.

Levine, D. U. and Lezotte, L. W. (1990): *Unusually Effective Schools: A Review and Analysis of Research and Practice*. Madison, WI: National Center for Effective Schools Research and Development.

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

Lupton, R. (2005): Social Justice and School Improvement: Improving the Quality of Schooling in the Poorest Neighbourhoods. *British Educational Research Journal* 31, 589–604.

Luyten, H. – Visscher, A. – Witziers B. (2005): School Effectiveness Research: From a Review of the Criticism to Recommendations for Further Development. *School Effectiveness and School Improvement* 16, 249–279.

Mancebon, M-J. – Bandres, E. (1999): Efficiency Evaluation of Secondary Schools: The Key Role of Model Specification and Ex Post Analysis of Results. *Education Economics* 7, 121–152.

Mayston, D. – Jesson, D. (1988): Developing Models of Educational Accountability. *Oxford Review of Education* 14, 321–339.

Meeusen, W. – van den Broeck, J. (1977): Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review* 18, 435–444.

Mizala, A. – Romaguera, P. – Farren, D. (2002): The Technical Efficiency of Schools in Chile. *Applied Economics* 34, 1533–1552.

Mortimore, P. – Sammons, P. – Stoll, L. – Lewis, D. – Ecob, R. (1988): *School Matters: The Junior Years.* Wells, Somerset: Open Books.

Mosteller, F. (1995): The Tennessee Study of Class Size in the Early Grades. *The Future of Children* 5, 113–127.

Nicolaidou, M. – Ainscow, M. (2005): Understanding Failing Schools: Perspectives from the Inside. *School Effectiveness and School Improvement* 16, 229–248.

Ojala, M-L. (1994): The Supply and Teachers in General Upper Secondary Schools (in Finnish). In Jakku-Sihvonen and Blom (eds.) General Upper Secondary Schooling in Finland 1994. *Arviointi ja seuranta* 5. Helsinki: Opetushallitus

Opetushallitus (2005): Functioning of the Non Graded General Upper Secondary Schooling 2005 (in Finnish). *Moniste* 20/2005. Helsinki: Opetushallitus.

Opetushallitus (2000): Functioning of the Non Graded General Upper Secondary Schooling (in Finnish). *Moniste* 6/2000. Helsinki: Opetushallitus.

Pekkala, S. (2003): Is Little Brother Nothing But a Trouble? Educational Attainment, Returns to Schooling and Sibling Structure. *VATT-Discussion papers* no. 302. Helsinki: Government Institute for Economic Research.

Pitt, M. – Lee, L. F. (1981): The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics* 9, 43–64.

Portela M. C. – Camanho, A. S. (2007): Performance Assessment of Portuguese Secondary Schools. Working paper, economics 07/2007. Universidade catolica Portuguesa (Porto). http://dspace.feg.porto.ucp.pt:8080/dspace/bitstream/2386/100/1/072007+Portela+e+Camanho+-+Performance+assessment+of+portuguese+secondary+schools.pdf.

Portela, M. C. A. S. – Thanassoulis, E. (2001): Decomposing School and School-Type Efficiency. *European Journal of Operational Research* 132, 357–373.

Purkey, S. C. – Smith, M. S. (1983): Effective Schools: A Review. *The Elementary School Journal* 83, 427–452.

Raudenbush, S. – Bryk, A. (2002): *Hierarchical Linear Models: Application and Data Analysis Methods*. 2nd ed. Newbury Park: Sage Publications.

Reynolds, D. – Packer, A. (1992): School Effectiveness and School Improvement in the 1990s. In Reynolds, D. – Cuttance, P. (eds.) *School Effectiveness: Research, Policy and Practice.* London: Cassell.

Reynolds, D. – Teddlie, C. (2001): Reflections on the Critics and Beyond Them. *School Effectiveness and School Improvement* 12, 24–31.

Rivkin, S. – Hanushek, E. – Kain, J. (2005): Teachers, Schools, and Academic Achievement. *Econometrica* 73, 417–458.

Robertson, D. – Symons, J. (2003): Do Peer Groups Matter? Peer Group Versus Schooling Effects on Academic Achievement. *Economica* 70, 31–53.

Ruggiero, J. (1996): On the Measurement of Technical Efficiency in the Public Sector. *European Journal of Operations Research* 90, 553–565.

Ruggiero, J. – Vitaliano, D. F. (1999): Assessing the Efficiency of Public Schools Using Data Envelopment Analysis and Frontier Regression. *Contemporary Economic Policy* 17, 321–331.

Rutter, M. – Maughan, B. – Mortimore, P. – Ouston, J. (1979): *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children.* Boston: Harvard University Press.

Sammons, P. (1999): *School Effectiveness. Coming of Age in the Twenty-First Century.* Lisse, Swets & Zeitlinger.

Sammons, P. – Thomas, S. – Mortimore, P. – Walker, A. – Cairns, R. – Bausor, J. (1998): Understanding Differences in Academic Effectiveness: Practioners' Views. *School Effectiveness and School Improvement* 9, 286–309.

Scheerens, J. – Bosker, R. (1997): *The Foundations of Educational Effectiveness.* Oxford/New York/Tokyo: Pergamon.

Schultz, T. W. (1975): The Value of the Ability to Deal with Disequilibria. *Journal of Economic Literature* 13, 827–846.

Schultz, T. W. (1967): The Rate of Return in Allocating Investment Resources to Education. *Journal of Human Resources* 2, 293–309.

Schultz, T. W. (1961): Investment in Human Capital. *American Economic Review* 51, 1–17.

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

Simar, L. – Wilson, P. W. (2007): Estimation and Interference in Two-Stage, Semi-Parametric Models of Production Processes. *Journal of Econometrics* 136, 31–64.

Simar, L. – Wilson, P. W. (2000): A General Methodology for Bootstrapping in Nonparametric Frontier Models. *Journal of Applied Statistics* 27, 779–802.

Silverman, D. (2000): *Doing Qualitative Research. A Practical Handbook.* London: Sage Publications Ltd.

Smith, P. – Mayston, D. (1987): Measuring Efficiency in the Public Sector. *OMEGA International Journal of Management Science* 15, 181–189.

Spence, M. A. (1973): Job Market Signaling. *Quarterly Journal of Economics* 87, 355–374.

Summers, A. - Wolfe, B. L. (1977): Do Schools Make a Difference? *American Economic Review* 67, 639–652.

Takala, P. (1994): The Development of General Upper Secondary Schools in 1970-1994 (in Finnish). In Jakku-Sihvonen and Blom (eds.) General Upper Secondary Schooling in Finland 1994. *Arviointi ja seuranta* 5. Helsinki: Opetushallitus

Teddlie, C. – Reynolds, D. (2000): *The International Handbook of School Effectiveness Research.* London/New York: Routledge.

Teddlie, C. – Stringfield, S. (1993): *Schools Make a Difference: Lessons Learned From a 10-Year Study of School Effects.* New York: Teachers College Press.

Thanassoulis, E. (1996): Altering the Bias in Differential School Effectiveness Using Data Envelopment Analysis. *Journal of the Operational Research Society* 47, 882–894.

Thanassoulis, E. – Portela, M. C. A. S. (2002): School Outcomes: Sharing the Responsibility Between Pupil and School. *Education Economics* 10, 183–207.

Thanassoulis, E. – Portela, M. C. S. – Despic, O. (2008): Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis. In Fried, H. O. – Lovell, C. A. K. – Schmidt, S. S. (Eds.): *The Measurement of Productive Efficiency and Productivity Growth.* New York: Oxford University Press.

Todd, P. – Wolpin, K. (2003): On the Specification and Estimation of the Production Function for Congnitive Achievement. *Economic Journal* 113, F3–F33.

Waldo, S. (2003): Efficiency in Education. Ph. D. Thesis. Department of Economics, Lund University.

Willms, J. D. (1986): Social Class Segregation and Its Relationship to Pupils' Examination Results in Scotland. *American Sociological Review* 51, 224–241.

Willms, J. D. – Raudenbush, S. W. (1989): A Longitudinal Hierarchical Linear Model for Estimating School Effects and Their Stability. *Journal of Educational Measurement* 26, 209–232.

Wössmann, L. – West, M. (2006): Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS. *European Economic Review* 50, 695–736.

Article I

Reprinted with the permission of Elsevier Limited, Copyright © 1998 Elsevier B.V.

Pergamon

Economics of Education Review, Vol. 17, No. 4, pp. 377–394, 1998
© 1998 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0272-7757/98 $19.00+0.00

PII: S0272-7757(97)00048-4

Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit Analysis[1]

TANJA KIRJAVAINEN* and HEIKKI A. LOIKKANEN†§

*University of Tampere and Helsinki School of Economics and Business Administration, P.O. Box 1210, 00101 Helsinki, Finland
†Government Institute for Economic Research (VATT), P.O. Box 269, 00531 Helsinki, Finland

Abstract—We studied efficiency differences among Finnish senior secondary schools by Data Envelopment Analysis (DEA). Four model variants were used. Average efficiencies in the most extensive models were 82–84 per cent. When parents' educational level was treated as an additional input, average efficiency increased to 91 per cent. The efficiency rankings of schools changed to some extent when simplest quantitative inputs and outputs were augmented by measures of teacher quality and national matriculation examination results. 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 and heterogenous student bodies were inefficient whereas school size did not affect efficiency. Surprisingly, private schools were inefficient relative to public schools. When parents' educational level was only included in the Tobit model, it affected efficiency positively. [*JEL* I21] © 1998 Elsevier Science Ltd. All rights reserved

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 such as 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 and 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 as well as 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 to 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 are 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 meas-

§Corresponding author.
[Manuscript received 1 November 1996; revision accepted for publication 22 July 1997]

ures 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 such as analyses of courts, health centres, ferries, etc. As we will point out in our brief summary of previous studies, education was one of the earliest areas in which 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 the efficiency score rankings of schools appear to be. If the rankings change appreciably with the addition of a priori meaningful input and output variables, the simplest indicators of efficiency ("partial productivities") are questionable. In addition, we also investigate whether the efficiency in two of the most extensive models is related to some factors that are not included in the DEA models, such as state grant variables, heterogeneity of student body and private/public dichotomy.

This article 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 using Tobit-models with variables not included in the DEA analysis. Section 6 offers some conclusions.

2. PREVIOUS STUDIES

Given the amount of inputs, a 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 the 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. The most common way is to apply statistical methods by either estimating the frontier production function using regression analysis or using 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 Fried *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 and Cooper and Cohn (1997) in the U.S.[2]

The other possibility of defining the efficiency frontier is to use non-parametric methods that are based on linear programming. 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.

One of the first studies applying DEA studied the program and managerial efficiency of a federally sponsored program called Program Follow Through (PFT) being in charge of 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 the method. Examples of such studies are Bessent and Bessent (1980), Bessent *et al.* (1982, 1983, 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 the U.K. Other European studies are those of Bonesrønning and Rattsø (1992, 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 with efficiency scores obtained by using more conventional regression analysis, in which the efficiency scores are calculated from the residuals. Examples of these studies are those of Mayston and Jesson (1988) and Sengupta and Sfeir (1986). In both the writers concluded that the efficiency rankings were different depending on the method used 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 has been that students' socioeconomic status affects 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 has 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 (McCarty and Yaisawarng, 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 modelling alternatives produced similar results in the sense that the rankings of efficiency scores in both cases were positively and significantly correlated.

As we pointed out in the introduction, it is not quite obvious what the inputs and outputs of educational process are and at what stage (timing) they should be measured. Because of this one should pay attention to the robustness of the results with respect to the choice of input and output variables. In most of the studies, the selection of variables seems to be based more on data availability than any other reason and only the final results are reported.[3] Our analysis is done by using quite a 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 Tobit models explaining the efficiency differences with variables related to the scale, state grants, private/public dichotomy, and student body of the school and test whether these results are related to the choice of variables in the DEA models.

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 a mathematical formulation of the 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 the efficiency measure when 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, it is 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 the efficient use of input X to the actual use of input X, i.e. X_N/X_H, keeping output constant.[4]

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 (see 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_i ($i = 1,...,m$) and

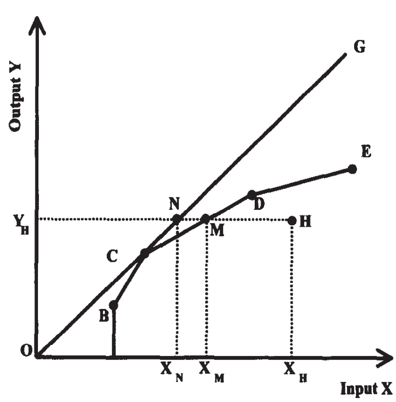


Figure 1. Determination of efficiency frontier when constant and variable returns to scale hold.

output weights by μ_r ($r = 1,\ldots,s$) the optimization problem can be formulated as follows, assuming constant returns to scale:

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

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

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

$$\mu_r, \nu_i \geq \epsilon \quad r = 1,\ldots,s;\ i = 1,\ldots,m \tag{4}$$

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 (Equation (1)) so that the sum of its weighted inputs equals one (Equation (2)) and so that the weighted outputs of all schools minus the weighted inputs of all schools is less than or equals zero (Equation (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 (Equation (1)) as well as to the second restriction (Equation (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 ($\omega > 0$), constant ($\omega = 0$) or increasing returns to scale ($\omega < 0$).

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

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

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

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

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

Above, s_r^+ and e_i^- are so called slack variables measuring the excess of inputs and outputs. The small positive constant ϵ guarantees 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 (Equation (5)) so that the outputs of school 0 equal to the sum of weighted outputs of other schools (Equation (6)). In addition, the weighted inputs of school 0 must equal the weighted inputs of other schools (Equation (7)). The optimal value of parameter θ in Equation (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

$$\sum_{j=1}^{n} \lambda_j = 1 \tag{9}$$

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 per cent 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 examinations, the final grading is done by a national matriculation examination board.

Our data consist of 291 senior secondary schools from all over 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 are 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 cover the years 1988–91. Whenever possible, we have measured the input variables as averages over the whole period. One of our output variables, the number of students who passed their grade, is also an average over three years.

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 measuring the share of state grants are, however, for 1992. This should not be a problem since the state grant system has not changed during the years 1988–92. The information for 1991 is used for other variables e.g. those depicting the environment and student body composition of the schools.

4.2. DEA Models and Variables

In studies of educational production functions 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 a certain time point or one measures these factors during the whole 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

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

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 for the ith student over the period t^* to t, I_i is the vector of innate abilities of the ith student, and $A_i^{t^*}$ is the vector of 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 a 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 the 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 can 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 Equation (10) in a form typically used also in the educational production function literature. Because of the lack of a 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 a 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 A (Table 9).

Our model 1 consisted of the simplest quantitative input and output variables. According to this model, schools were depicted as producers of students who pass their grade or matriculation examination 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 admitted 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 of 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] For each school it is measured by grade of last admitted student. 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 admitted pupils would have been a better measure.

Model 3 consisted of two additional input and output variables. Even though teacher characteristics are rarely found to have an 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 a master's degree. The experience of teachers was measured by giving one point for each five professional year periods and then adding these points 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 the

Table 1. Variables used in the DEA efficiency measurement

	Model 1	Model 2	Model 3	Model 4
		Inputs		
Teaching hours per week	x	x	x	x
Non-teaching hours per week	x	x	x	x
Experience of teachers			x	x
Education of teachers			x	x
Admission level		x	x	x
Educational level 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 additional subjects in matriculation examination			x	x

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 a single composite score we allowed schools to have a different emphasis on their provision of courses.

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. Parents' educational level is measured by giving points for degrees.[8]

Finally, we shall comment on 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.

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 the deviation of scores from one (or 100 per cent) indicates savings possibilities in the use of inputs. As pointed out in Section 2, under the assumption of constant returns to scale (CRS) the efficiency scores are the same as if we had chosen to keep inputs constant and measured efficiency in the output increasing direction. However, under the assumption of variable returns to scale (VRS) this choice matters.

In addition to having four models with different inputs and outputs, we shall present results assuming CRS and VRS. In this connection, it is worth noting that the efficiency scores for each observation under VRS are typically greater than (and at a minimum equal to) those under CRS (see Figure 1). Also, as the number of variables in a DEA model increases, the efficiency scores either increase or remain the same and tend to increase average efficiency scores. Thus average efficiency scores of different models cannot be compared with 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 when there is limited information on educational processes. Then, we would know that efficient (and inefficient) schools remain similarly assessed irrespective of how detailed is the 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, such as class size and school size from the input side, and 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. Towards the end of this section we also present Tobit models in which efficiency differences derived from the most extensive models are explained by variables related to the schools' scale of operation, student body composition, private/public dichotomy and state grants.

5.1. Efficiency Distributions

Table 2 offers basic information on the distribution of efficiency scores obtained from our four models 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 consider-

Table 2. The average efficiency, minimum and maximum 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	66.3	72.2	77.5	80.0	81.9	84.1	91.3	93.7
Minimum	17.0	40.6	38.1	49.0	43.8	58.4	59.7	59.8
Maximum	100	100	100	100	100	100	100	100
Percentage share of efficient schools	1.4	5.8	3.1	7.6	10.0	16.5	21.0	33.3

able 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 model 3 (admission level, teachers' experience, and education as inputs 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.

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 the addition of a variable measuring student quality (admission level in model 2), average efficiency 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 efficiency frontier, increased from only 1.4 per cent in model 1 to 10 per cent in model 3 and 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 they become rather small in the 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 the CRS efficiency frontier OG, the VRS efficiency frontier must also be close to OG implying 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 on the vertical axis and the efficiency scores on the horizontal axis. The efficiency distributions in Figure 2 are based on the results of models 1–4 assuming CRS.

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) again shifted 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.

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 outlier units that are very different from other units, either genuinely or because of miscoding, measurement error, etc. In such a case, omitting 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 the frontier and efficiency scores in each of the models, we ran DEA analyses that dropped 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 in regard to outlier schools. For example, in the case of model 1 assuming CRS, we ran four additional DEA analyses. We then tested the similarity of efficiency rankings

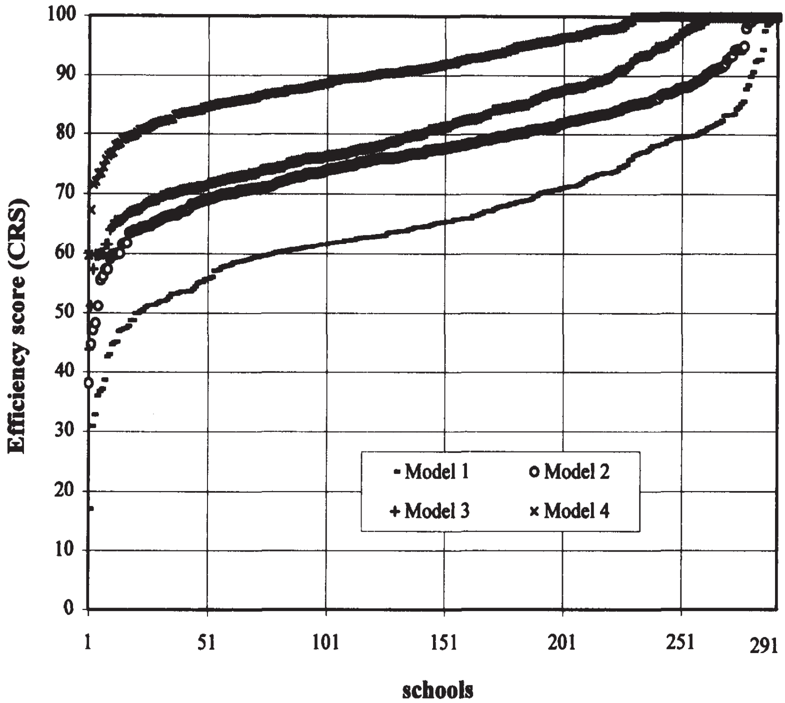


Figure 2. The efficiency distributions of models 1–4 when CRS is assumed.

between the model with all the schools included and those based on dropping out each efficient unit one 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.

The high rank correlation coefficients show that the rankings[10] are relatively stable in 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 still ranged at most from 0.96 to 1.00 in models 2 and 4. In the 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 stable in regard to outlier schools the greater is the number of inputs and outputs in the model. This is especially when CRS is assumed. 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 remainder of the CRS models. In case of VRS, the difference between the mean of iterations and the mean efficiency with all the schools in the 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 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 since a correlation coefficient of 1 means that the rankings are exactly the same. The rankings between model 2 and 3 were again similar as the correlation coefficient was 0.86 whereas the rankings between model 3 and model 4 changed appreciably as the coefficient was 0.66. The largest differences in the efficiency rankings were unsurprisingly between models 1 and 4 as

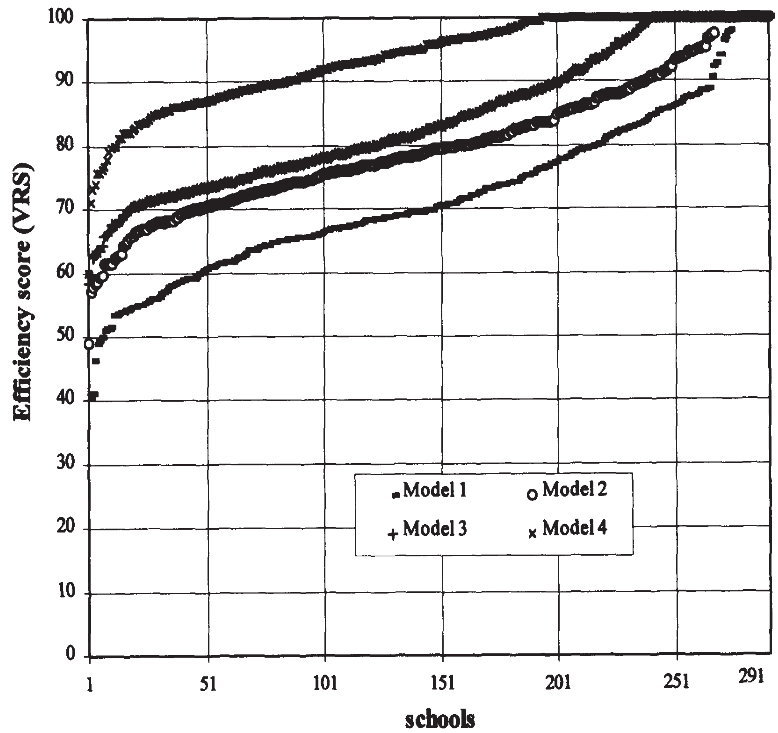


Figure 3. The efficiency distributions of models 1–4 when VRS is assumed.

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

		Number of efficient schools	The range in Spearman rank correlation coefficient Minimum	Maximum	Mean efficiency	Iterated mean efficiency	Standard deviation of means
Model 1	CRS	4	0.98	0.99	66.3*	67.9*	3.2
	VRS	17	0.96	1.00	72.2	72.5	0.9
Model 2	CRS	9	0.96	1.00	77.5	80.0	0.9
	VRS	22	0.88	1.00	80.0	80.3	0.8
Model 3	CRS	29	0.97	1.00	81.9	82.0	0.6
	VRS	48	0.92	1.00	84.1	84.2	0.5
Model 4	CRS	61	0.96	1.00	91.3	91.3	0.2
	VRS	97	0.94	1.00	93.7	93.7	0.1

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

Table 4. Spearman rank correlation coefficients between models 1–4

	Model 1	Model 2	Model 3
Model 2	0.84 (CRS)		
	0.83 (VRS)		
Model 3	0.73 (CRS)	0.86 (CRS)	
	0.71 (VRS)	0.87 (VRS)	
Model 4	0.44 (CRS)	0.54 (CRS)	0.66 (CRS)
	0.45 (VRS)	0.60 (VRS)	0.69 (VRS)

Table 5. Frequencies (and range of efficiency scores) of schools in quartiles according to their efficiency scores in model 1 and model 3 assuming CRS

	Model 3				
	I Quartile (38.1–71.0)	II Quartile (71.1–77.4)	III Quartile (77.5–83.5)	IV Quartile (83.5–100)	Total
Model 1					
I Quartile (17.0–59.6)	51 (17.5)	11 (3.8)	5 (1.7)	6 (2.1)	73 (25.1)
II Quartile (59.6–65.1)	18 (6.2)	33 (11.3)	16 (5.5)	6 (2.1)	73 (25.1)
III Quartile (65.1–73.6)	4 (1.4)	28 (9.6)	30 (10.3)	11 (3.8)	73 (25.1)
IV Quartile (73.6–100)	0 (0.0)	1 (0.3)	22 (7.6)	49 (16.8)	72 (24.7)
Total	73 (25.1)	73 (25.1)	73 (25.1)	72 (24.7)	291 (100.0)

the respective correlation coefficient was only 0.44. When VRS was assumed, the results were very similar in both pattern and size to those based on CRS assumption.

To obtain 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 Table 5 (CRS) and Table 6 (VRS) we have cross-tabulated the number of schools in each quartile on the basis of efficiency scores from models 1 and 3. If the rankings remained the same, the off-diagonal elements of the table would have no observations. This, however, is not quite true.

According to Table 5 (CRS), out of 291 schools some 56 per cent (163 schools) stayed in the same quartile (diagonal elements) and the remaining 128 schools (off-diagonal elements) 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 a 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 remainder 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, what is included as inputs and outputs may make a big difference in evaluating their efficiency. Note, however, that the risk of mis-evaluation seems to be asymmetric as no school in the fourth quartile according to model 1 moved to the first, and only one school to the second quartile when model 3, with more variables, was used.

Assumption of VRS seemed to lead to somewhat more instability than that found with the CRS results. According to Table 6 (VRS), out of 291 schools some 52 per cent (151 schools) stayed in the same quartile (diagonal elements) and the remaining 140 schools (off-diagonal elements) changed quartile when results of model 3 were used instead of model 1. In general, the pattern of changes in rankings is very similar to those obtained assuming CRS.

Similar comparisons can be made between any two pairs of models. Instead of presenting and commenting on them all, we feel that Table 5 and Table 6 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 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 variables of interest. In

Table 6. Frequencies (and range of efficiency scores) of schools in quartiles according to their efficiency scores in model 1 and model 3 assuming VRS

	Model 3				
	I Quartile (58.4–75.8)	II Quartile (75.8–82.6)	III Quartile (82.6–93.5)	IV Quartile (94.0–100)	Total
Model 1					
I Quartile (40.6–64.1)	45 (15.5)	16 (5.5)	4 (1.4)	8 (2.8)	73 (25.1)
II Quartile (64.2–69.8)	25 (8.6)	27 (9.3)	12 (4.1)	9 (3.1)	73 (25.1)
III Quartile (69.9–80.3)	3 (1.0)	30 (10.3)	32 (11.0)	8 (2.8)	73 (25.1)
IV Quartile (80.6–100)	0 (0.0)	0 (0.0)	25 (8.6)	47 (16.2)	72 (24.7)
Total	73 (25.1)	73 (25.1)	73 (25.1)	72 (24.7)	291 (100.0)

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 in which 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 differences 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 in 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 viewpoint differs, the evaluation and rankings may also 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 quality of senior secondary schools.

We see from Figure 4 that efficiency scores and the average grades in matriculation examination are almost unrelated (Pearson correlation coefficient 0.20). A school with poor matriculation examination results can be efficient because it may also have a small amount of inputs. On the other hand, high-performing schools (using matriculation examination scores) can be inefficient since their use of inputs is large.

In public debates concerning schools it is often claimed that school size and class size are related to efficiency. Here, we shall consider this question with scatter diagrams recognizing that these (or related) school variables are implicitly incorporated already in the DEA models. In Figure 5 we have plotted efficiency scores and school size measured by the number of pupils. The correlation (Pearson) 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 efficiency varies in nearly all sizes of school. 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.

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 per week in the school by 20 (our estimate of the average teaching hours of each teacher per week) and then dividing the number of pupils by the number obtained. The correlation between the class size and efficiency scores is 0.55

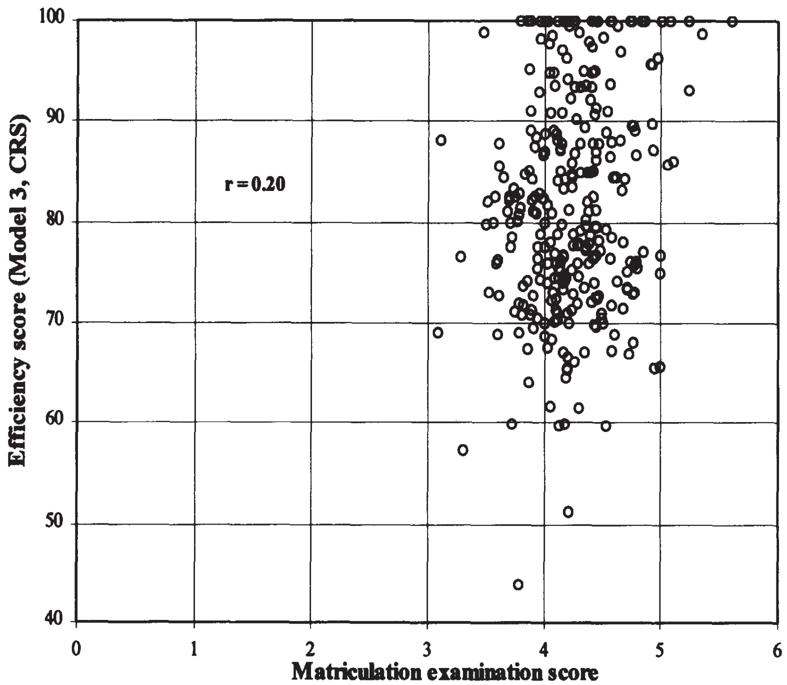


Figure 4. Scatter diagram for efficiency scores (model 3, CRS) and matriculation examination scores.

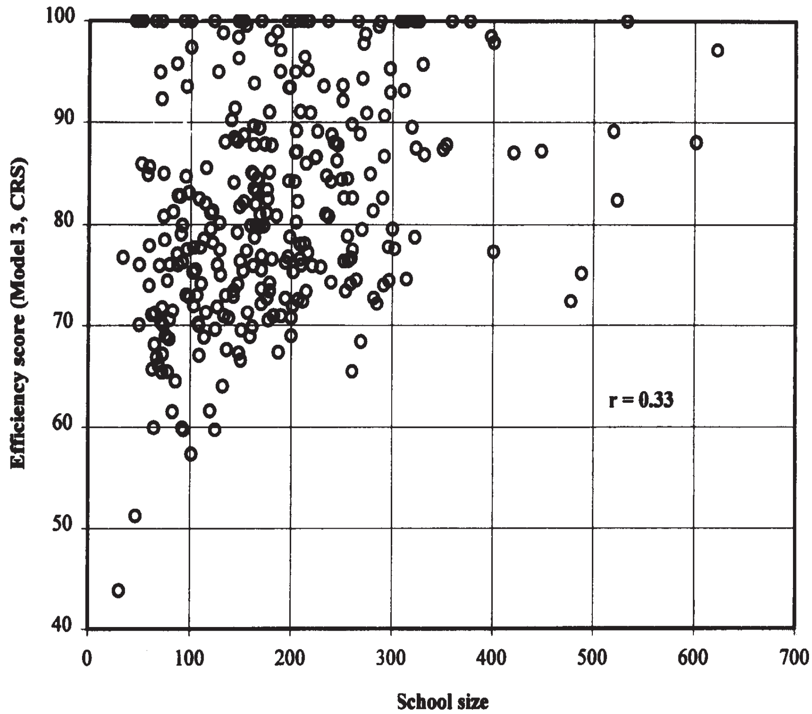


Figure 5. Scatter diagram for efficiency scores (model 3, CRS) and school size (measured by number of pupils).

suggesting that schools with larger classes are more efficient. Despite this general relation, one must note again that there are some fully efficient schools with small classes.

5.3.2. A Tobit model. Here, we shall present results of a statistical exercise in which efficiency differences are explained by some variables not directly included in the DEA analysis. As our basic rule, factors that can more easily be influenced by the school are included in the original DEA models and those outside the decision making power of schools, i.e. the influence of environment, are taken as factors explaining efficiency differences (see e.g. Ray, 1991 and McCarty and Yaisawarng, 1993).

Inefficiency may also be caused by an unoptimal scale of operation. This can be taken into account either by considering efficiency differences obtained under the assumption of CRS and by explaining them by such variables as measuring the scale of operation (e.g. class size and school size), or by considering efficiency differences obtained under the assumption of VRS (the difference between efficiency scores from CRS and VRS models is an indication of scale inefficiency). We ran our models in both ways. In the following we shall concentrate on discussing the results of the first alternative and only shortly comment on the results of the second.

As efficiency scores are limited to range to the [0,1] interval, and there are always one or more observations at the upper limit, we wish to take this fact into account. In a standard Tobit model the dependent variable is either zero or some positive number (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

$$\begin{aligned} &Y_i^* = X_i\beta + \mu_i, \\ &Y_i = Y_i^*, \text{ if } Y_i^* > 0, \\ &Y_i = 0, \text{ otherwise.} \end{aligned} \tag{11}$$

In Equation (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 the Tobit model as a second

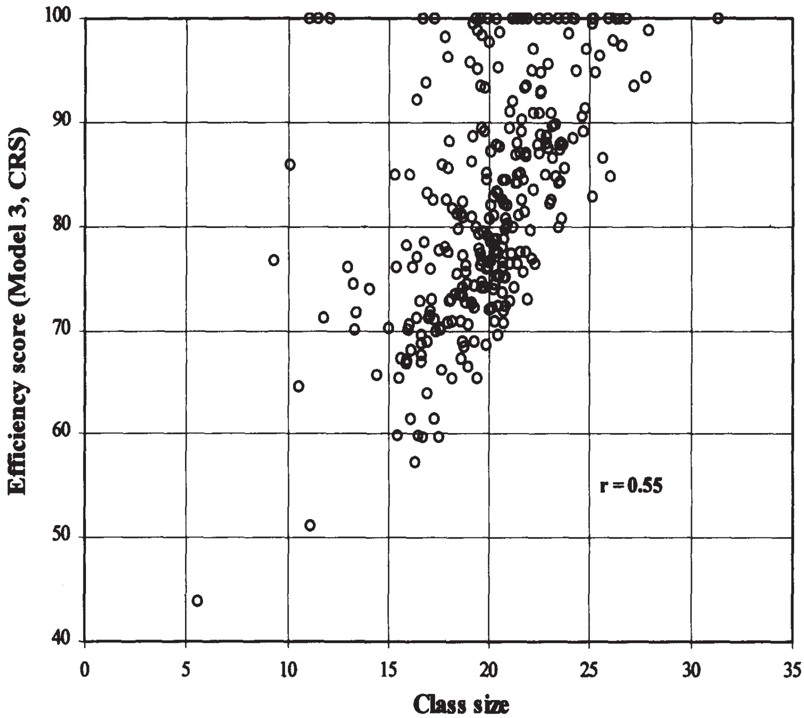


Figure 6. Scatter diagram for efficiency scores (model 3, CRS) and class size.

step after deriving efficiency scores by DEA has been used, e.g. by McCarty and Yaisawarng (1993) in a school context as pointed out earlier, and also by Luoma *et al.* (1996), in studying the efficiency of health centres 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 results here to calculate efficiency scores corrected by variables related to school environment as did McCarty and Yaisawarng.

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

As mentioned above we use school size (School size) and class size (Class size) in our Tobit models to explain inefficiency stemming from unoptimal scale of operation in CRS models. If there is a school size or a class size (measured by number of pupils) that is efficient, inefficient schools have either smaller or larger respective sizes than these optimal sizes. 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 1989–92 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. expenditure). 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 40.8 to 100.6 per cent 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 (having grants as public schools) in our data and we have a dummy variable (Private) to test its relation to (in)efficiency. As other 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

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

	Tobit 3A		Tobit 3B		Tobit 3C	
	Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio
Constant	2.34	1.59	2.49	1.53	0.92	0.47
School size	0.0006	0.23	0.0016	0.68	0.0016	0.70
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' education	−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. expenditures					0.99	1.05
Grant ratio2			−0.0027	−0.37		
Proportion of efficient schools	29/291		29/291		29/291	
R^2	0.361		0.358		0.362	
Log-likelihood	211.82		210.98		211.51	

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 generosity 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 the Appendix B (Table 10).

The results of three Tobit models explaining the inefficiency scores of CRS model 3 are given in Table 7. 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 are 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 for the range of class size in our data. According to these results inefficiency first increases as class size increases. After a class size of on average 11 students is reached, inefficiency starts to decrease as class size further increases and when class size is on average 27 students inefficiency is minimized.

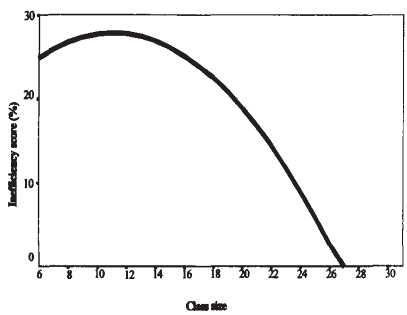


Figure 7. The relation between inefficiency and class size according to Tobit model 3A.

Somewhat surprisingly, and other things equal, private schools are less efficient than public schools. And parents' educational level has a clear positive effect on efficiency.

Next, we shall take the inefficiency scores from model 4 (assuming CRS), i.e. a model where parents' educational level is treated as an input, and explain the differences using Tobit models. The results are presented in Table 8. Note that the number of schools that are fully efficient (inefficiency score 0) is greater and the average inefficiency is lower here compared with the findings of model 3 (assuming CRS). Thus it is not surprising that the R^2 values are lower here than in the models in Table 7.

When parents' educational level has been already included in the DEA model and is not an explanatory variable in the 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 achieve a significant coefficient, heterogeneity in grades is positively related to inefficiency. When state grant variables are used instead of municipality type variables the grant ratios, surprisingly, get negative coefficients which are close to being significant.

As a final alternative we ran the Tobit models using the inefficiency scores obtained from VRS models as dependent variable.[11] In type 3A and 3B Tobit mod-

Table 8. Parameter estimates (normalized coefficients) of Tobit models explaining school inefficiency (model 4, CRS)

	Tobit 4A		Tobit 4B		Tobit 4C	
	Coefficient	*t*-ratio	Coefficient	*t*-ratio	Coefficient	*t*-ratio
Constant	–4.91	–2.94	–3.51	–2.08	–3.40	–1.68
School size	0.00009	0.04	0.0023	0.96	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	3.87	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	0.17	0.86				
Grant ratio1					–0.014	–1.80
Act./Acc. expenditure					0.063	0.07
Grant ratio2			–0.012	–1.81		
Proportion of efficient schools	61/291		61/291		61/291	
R^2	0.253		0.229		0.229	
Log-likelihood	214.34		209.69		209.66	

els, excluding the school size and class size variables, the only significant explanatory factor was the educational level of students' parents. In the type 3C Tobit model also other factors started to be of influence. Heterogeneity of the student body, the Grant ratio1, ratio of actual to acceptable expenditures, and being a private school all increased inefficiency and were statistically significant. The respective results of type 4A, 4B, and 4C Tobit models, which did not include parents' educational level (included in DEA), indicated that heterogeneity of student body was the only significant factor for the two first models, and, in the type 4C model, private schools again turned out to be inefficient. The two state grant variables gave positive coefficients, but their t-values were only 1.5 (Grant ratio1) and 1.45 (Act./acc. expenditure).

6. CONCLUSIONS

In this paper we studied the efficiency differences among 291 Finnish senior secondary schools by DEA. Four different models with school level data were used. Average efficiencies in modelling with the most extensive list of school related inputs and outputs (model 3) were 82–84 per cent, ranging from 44 to 100 per cent. When parents' educational level was treated as an additional input (model 4), 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 efficiency frontier.

One purpose of the study was to analyze how efficiency rankings of schools depend on inputs and outputs included in DEA. By classifying schools into quartiles according to their efficiency scores, and comparing the simplest model 1 and model 3 with some additional variables, 60–70 per cent of schools in the topmost and the lowest quartiles tended to remain in the same quartile. 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 the efficiency frontier in regard to outlier schools we applied a jackknifing technique. Changes in efficiency scores and rankings (according to Spearman rank correlation coefficients) turned out to be minor.

The results also showed that, for example, 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 a schools' student body, scale of operation, and state grants but were not included in the DEA models. 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. However, 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 became significant so that schools in urban areas were inefficient relative to those in rural and other less populated areas. Also student body composition began to matter. Although the share of female students did not achieve a significant coefficient, heterogeneity in grades was significant and positively related to inefficiency. When grant variables were used instead of municipality type variables they, surprisingly, gave negative coefficients which were close to being significant. Also the dummy variable related to private schools got a positive and significant coefficient.

When these analyses were carried out taking the efficiency scores from VRS models as dependent variable and dropping out the variables controlling for the scale of operation (i.e. class and school size) from Tobit models, the most noticeable change in results was that the ratio of grants to educational expenditures, as well as the ratio of actual to acceptable expenditures, now became positive and in some cases had statistically significant coefficients. Here, heterogeneity of student body and being a private (instead of public) school were related to inefficiency.

In conclusion, the results of our DEA analyses indicate the importance of variable choices both on the input and the output side. Average efficiency and rankings vary to some extent from one model to another preventing too straightforward conclusions. Furthermore, we view these results only as a step in a research process which involves finding out what characterizes observations in opposite ends of efficiency distributions. In addition, the results of Tobit models suggest the need for future research, since the results change depending among other things on the DEA-model used. That the educational level of pupils' parents is important for efficiency seems to be a robust result, but state grant variables, for example, behave very unpredictably in alternative models. Also the unexpected outcome of private schools being less efficient than public ones can hopefully be tested with alternative data.

Acknowledgements—The authors would like to thank Elchanan Cohn and two anonymous referees for their valuable comments. Responsibility for any errors remaining rests with the authors.

NOTES

1. This study is part of the "Effectiveness of schooling" research program of the Academy of Finland.
2. 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 e.g. Hanushek (1986) and Cohn and Geske (1990).
3. This is also the problem of many production function studies using statistical methods (see e.g. Hanushek, 1979).
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 the latter provides 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 i.e. keeping their output constant. Therefore we do not discuss the other alternative i.e. the determination of efficiency scores in the case where schools are assumed to maximize their outputs.
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); the results are obtainable from the authors upon request.
6. There are also other 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). Bonesrønning 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 consists of mother tongue, the second national language of the country, foreign language, and mathematics (comprehensive or a short course) or science and humanities. Additional subjects may include grades in foreign language, mathematics (short course), and science and humanities. The 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 the matriculation examination can be achieved in various ways.
8. The following points were given to each degree: 1.5 = no degrees other 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, but not a university degree (13–14 years of schooling); 6 = bachelor's degree; 7 = master's degree; 8 = postgraduate degree. The surprisingly detailed information on parents' educational level was obtained from Statistics Finland.
9. Also Bonesrønning and Rattsø (1994) and Färe *et al.* (1989) used jackknifing in their analyses. The

difference in our analysis is that they dropped each school one at the time whereas 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 value of 0 indicates no relation between the rankings and the value of − 1 indicates reverse rankings.
11. These results are not reported here. They are available from the authors upon request.

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.

Cohn, E. and Geske, T. G. (1990) *Economics of Education*. 2nd edn. Pergamon Press, Oxford.

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.

Cooper, S. T. and Cohn, E. (1997) Estimation of a frontier production function for the South Carolina educational process. *Economics of Education Review* (in press).

Färe, R., Grosskopf, S. and Weber, W. (1989) Measuring school district performance. *Public Finance Quarterly* **17**, 409–428.

Fried, H., Lovell, K. C. A. and Schmidt, S. S., (Eds) (1993) *The Measurement of Productive Efficiency. Techniques and Applications*. Oxford University Press, New York.

Hanushek, E. A. (1986) The economics of schooling: production and efficiency in public schools. *Journal of Economics Literature* **26**, 1141–1177.

Hanushek, E. A. (1979) Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources* **14**, 351–388.

Hanushek, E. A. and Taylor, L. L. (1990) Alternative assessment of the performance of schools. *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.

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. Government Institute for Economic Research, Helsinki.

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. (1996) Financial incentives and productive efficiency in Finnish health centres. *Health Economics* **5**, 435–445.

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*, eds H. Fried, K. C. A. Lovell and S. S. Schmidt. Oxford University Press, New York.

Ray, S. C. (1991) 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. (1987) Measuring efficiency in the public sector. *OMEGA International Journal of Management Science* **15**, 181–189.

APPENDIX A

Table 9. Summary statistics of the variables used in DEA analyses

	Mean	Standard deviation	Minimum	Maximum
		Inputs		
Teaching hours per week	182.2	89.2	43.8	567.0
Other than teaching hours per week	21.0	8.7	3.5	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	61.0	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)	178.2	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 additional subjects in matriculation examination	276.7	173.2	42.0	1258.0

APPENDIX B

Table 10. Summary statistics of the variables used in Tobit analyses

	Mean	Standard deviation	Minimum	Maximum
Parents' educaytion	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	72.5	10.9	50.0	86.0
Act./acc. expenditure	1.03	0.06	0.75	1.26
Grant ratio2	71.0	11.1	40.8	100.6

Article II

EFFICIENCY OF SCHOOLS: EXPLAINING HIGH AND LOW PERFORMANCE WITH ORGANIZATIONAL CHARACTERISTICS

Tanja Kirjavainen

Abstract

In this paper, the efficiency of Finnish general upper secondary schools is first studied with Data Envelopment Analysis (DEA) using four different models. In the simplest model input and output variables were quantitative. In the most extensive model, parental socioeconomic background and earlier achievement were controlled. Average efficiency varied from 71% to 94%. At the second stage, schools at the top and bottom 15% in efficiency distribution were studied closer with survey data. The probability of being in the group of the most efficient schools was explained with statistical probit and tobit models using variables depicting environmental and organizational practices of the schools. Concerning environmental factors, schools with higher admittance level, less heterogeneous student body and higher proportion of female students were more likely to be efficient. Probability of being efficient increased when the school had an organizational structure with a management group or other similar body, the less it used such management tools as formal mission, goals or strategy, the lower the participation of teachers in training, the lower the number of joint teaching projects across subjects and the more there were active clubs in the school. Other organizational factors were not related to efficiency. The results of statistical models explaining efficiency differences were fairly unstable.

Key words: Efficiency, Data Envelopment Analysis, Secondary education (JEL I21)

Contents

1. Introduction

Data envelopment analysis (DEA) has been widely used in efficiency measurement of schools. In such a context, schools are viewed as productive units using multiple inputs and outputs. The method calculates a relative efficiency measure for each school by determining an efficiency frontier through the most productive schools of the data and by measuring the distance of each school to the frontier. Schools at the frontier get an efficiency score of 1 (100% in this study) and schools below the frontier get efficiency score between 0 and 1 (0 and 100%) depending on their distance to the frontier.

The advantage of DEA compared to statistical efficiency measurement methods is that it is easy to incorporate several inputs and outputs into the analysis. Furthermore, because of the nonparametric nature of the method, no strong assumptions about the production technology are needed. Instead, it is determined from the data. The assumption about the distribution of efficiency is also unnecessary.

Most often DEA has been used for determining relative efficiency differences of schools.[1] As quite considerable efficiency differences seem to characterize the school sector, question arises why some schools get more out of their resources than others. For investigating the causes of inefficiency, a two-stage approach has been most often applied. Ray (1991) and McCarty and Yaisawarng (1993) corrected the efficiency scores obtained with DEA with student background factors. Kirjavainen and Loikkanen (1998) studied the effect of heterogeneity of student body, class size, school size and the location of the school on efficiency measured with DEA. Other similar studies adopting a two-stage approach have examined the effect of competition (Bradley

[1] See e.g. Bessent and Bessent (1980), Charnes *et al*. (1981), 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. Jesson *et al.* (1987) and Smith and Mayston (1987) studied the efficiency of school districts (LEAs) in United Kingdom. Other European studies are an efficiency analysis of Norwegian high schools by Bonesrønning and Rattsø (1994) and Finnish general upper secondary schools by Kirjavainen and Loikkanen (1993). In some of the studies efficiency differences calculated by DEA have also been compared to results of statistical methods (see e.g. Mayston and Smith, 1988 and Sengupta and Sfeir, 1986).

et al., 2001; Duncombe *et al*., 1997; Waldo, 2003), voter monitoring (Duncombe *et al*., 1997) and political context of the municipality (Waldo, 2003) on efficiency.

In school effectiveness studies, a wide range of other factors related to schools' organizational and instructional practices have been addressed. Instead of DEA, these studies often use statistical multilevel modeling for determining effective and ineffective schools. The results of school effectiveness studies have been reviewed several times (see e.g. Teddlie and Reynolds, 2000; Sammons, 1999; Scheerens and Bosker, 1997; Reynolds *et al*., 1994; Lee *et al*., 1993; Reynolds, 1992; Levine, 1992; Purkey and Smith, 1983) and reviews have come up with different lists of characterizations of effective schools. Factors that are in many reviews associated with effective schools include strong educational leadership, high expectations of student achievement, emphasis on basic skills, maximized time on learning, a safe and orderly climate, practice-oriented staff development at the school site, parental involvement and support, monitoring of student progress, and clear school goals. Other correlates include e.g. school-site management that allows schools to decide its ways to improvement and academic performance, instructional leadership that initiates and maintains development process, staff stability, clear curriculum articulation and organization, school wide recognition of academic success, district support, collaborative planning and collegial relationship among staff, and sense of community.

Some studies using DEA have taken similar case study approach as school effectiveness studies and investigated the most efficient schools measured with DEA more intensively (Dodd, 2006; Portela and Camanho, 2007). The findings of these studies are similar to many school effectiveness studies. As these studies concentrate on efficient schools, it is not certain if the same characteristics are also present in inefficient schools.

Contrary to earlier studies that concentrate on taking closer look on schools at the efficiency frontier, in this paper schools at both ends of the efficiency distribution are studied. Efficiency of Finnish general upper secondary schools is first measured with DEA using a set of usual input and output variables. In the second stage, survey information concerning organizational practices was collected from schools that were mostly at the upper and lower tails of the efficiency distribution in these models. The survey data cover information on students, communication between parents and school, management of school, staff development, teacher and student monitoring and school facilities from 73 schools. The differences in these factors are studied statistically using descriptive statistics and statistical models. Sensitivity of the results is also tested with a more extensive DEA model and by grouping schools as efficient and inefficient in alternative ways.

The results of this study show, that there were clear efficiency differences between schools. However, these differences were not related to most of the organizational practices examined. Mission, goals and strategy were less seldom determined in efficient schools. If they were determined, principals of efficient schools expressed higher consensus about them. In some of the models, schools with management or other assisting group were more efficient. Furthermore, efficient schools had less teaching projects across subjects and teachers participated less in training but had larger number of active clubs according to some of the models. Most of the results were not robust to model specification and grouping. Concerning the factors related to student body composition, such as heterogeneity and proportion of female students, they were more robustly related to efficiency. Hence, the results of this study demonstrate that concentrating only on top performing schools when examining the causes of efficiency may give distorted picture and provide little guidance for improving efficiency.

The paper continues as follows. In section 2 Finnish general upper secondary schools are described. Section 3 presents the research design and data. In section 4 the models and variables used in DEA analysis are introduced. This section concludes with the presentation of the results of efficiency differences calculated with DEA. Section 5 describes the variables explaining efficiency differences and provides the results of the estimations and section 6 concludes.

2. Finnish general upper secondary schools

Since this study uses data on the 1990's, a short description of the system at that time is provided in the following. Certain core characteristics of the Finnish general upper secondary schooling were the same at the beginning of the 1990's as they are today but there were also many differences.

Schools provided general education after the nine - year comprehensive school. Some 55% of the comprehensive school-leaving pupils continued their studies in general upper secondary schools. There were about 450 general upper secondary schools in Finland and they provided education for about 100 000 students. Students followed year classes and the education was usually completed in three years. Studies were divided into courses so that each course consisted of 38 lessons. The school year was usually divided into five or six periods. A schedule was devised for each period focusing on certain subjects.

The school concluded with a national school-leaving examination, the matriculation examination. Passing of the examination gave general eligibility for university studies and vocational education intended for matriculated students. The examination comprised of four compulsory tests: the student's mother tongue (Finnish/Swedish/Sami), the second national language of the country (Finnish/Swedish), a foreign language (usually English, French, German or Russian) and either the mathematics or science and humanities. The latter comprised questions in several subjects and the pupil was free to choose any combination of questions and subjects: religion, psychology, philosophy or ethics, history, civics, physics, chemistry, biology and geography. Besides compulsory tests, the candidate was able to take optional tests. Teachers carried out the initial grading of the exams; the final grading was done by the national matriculation examination board.

During 1985-1994 general upper secondary schools were quite tightly regulated at the national level. National core curriculum established by the Finnish national board of education was very detailed. It provided little freedom for schools and teachers to direct the content of instruction. Schools were mainly maintained by the municipalities and financed by local taxes and state grants to municipalities. The state grant system was also very centralized in nature and the grants to municipalities and schools were earmarked.

In mid 1990's the tight governmental control was released and decision making power was decentralized into local level. The new national core curriculum issued in 1994 provided a framework for instruction. Each school had a two year transition period during which they had to start devising their own school level curriculum based on the national core curriculum. In addition, the system with fixed classes was abolished and schools started following the non-graded system. The renewed state grant system supported the local level decision making by making the grants general.

3. Research design and data

This study was carried out in two phases. Efficiency of Finnish general upper secondary schools was first measured using DEA. Based on the results of three different DEA models, schools at the tails of the efficiency distributions were studied closer concentrating on their organizational practices. The analysis was statistical and it was based on survey data.

The data used in efficiency measurement was a cross section aggregated to the school level. It was compiled from several official registers and complemented with a survey to schools. The data consisted of 309 general upper secondary schools located all over Finland. There were some 450 general upper secondary schools providing education for students younger than 18 years of age in 1991, but unfortunately not all the necessary information was available from all the schools.

The key school output measures concerned students completing their studies in 1991. Since the general upper secondary school lasted for three years, students matriculating in 1991 started their studies in the fall of 1988. Therefore, the input data covered the years 1988-1991. Whenever possible, all input variables were measured as averages over the whole period. One of the output variables, the number of students passing their grade, was also an average over the three years. For variables depicting the environment and student body composition of the schools information for the year 1991 was used.

The second stage survey data, which was used for studying the organizational practices, was designed by taking into account the findings of production function and school effectiveness studies. In particular, the characteristics of effective schools listed by Purkey and Smith (1983) and Levine (1992) were considered. The survey consisted of both structured and open questions. Open format was used in connection of those school practices which were difficult to capture in closed format either because of the nature of the phenomenon or because of limited knowledge of the actual school practices. When coding the responses of open questions, information on the efficiency of the school was hidden. The classifications of open questions were created from the responses.

The survey data was collected in the summer of 1995. The survey was addressed to principals of 130 schools which were classified either as efficient (58 schools) or inefficient (72 schools) based on their efficiency scores in three different DEA models.[2] Some 62% of the schools returned the survey after two rounds. Efficient schools were little more likely to respond (71%) than the inefficient schools (54%). Some of the responses were so incomplete that they were dropped out.

[2] School was classified as efficient (or high performer) if it was among the 15% most efficient schools in model 2, 3, or 4 assuming variable returns to scale. Respectively, school was classified as inefficient (or low performer) if it was among 15% least efficient schools in model 2, 3, or 4 assuming variable returns to scale. The efficiency scores assuming variable returns to scale were used in order to control for the scale effects.

The final data used in the second stage analyses comprised of 73 schools, of which 40 were efficient and 33 inefficient. In other words, some 69% of the efficient and 46% of inefficient schools were used for analyzing the causes of efficiency differences. This result can be considered as satisfactory taking into account the delicate nature of the survey.[3]

The respondents were compared to all schools selected to the survey in terms of efficiency scores, matriculation examination scores, school size and type of municipality.[4] These results are reported in Appendix 1. They show that the average efficiency scores, average scores in the matriculation examination, and school size did not differ in the two groups. Only the schools in urban areas were somewhat less likely to respond. Correspondingly, the proportion of schools in densely populated and rural areas was somewhat higher among the respondents. However, these differences were not statistically significant.

In an ideal situation, one would collect data linking variables used in efficiency measurement to variables describing the school processes so that there is no time lag between these two sets of variables. In practice, this is rarely possible. In most cases, because register data lags behind, the collection of information on school processes and practices takes place few years after the efficiency measurement. This means that there are limitations in the data describing school practices and processes that are related to the continuing availability of relevant staff in school, and potential problems of recall and post-

[3] The research design was not mentioned in the covering letter. Thus, principals did not know the efficiency score of their school when responding to the survey. Only information given was related to the purpose of the study. A copy of the summary of an earlier study by Kirjavainen and Loikkanen (1993) was also attached for providing information on efficiency measurement and differences in efficiency scores. Since names of the schools were not mentioned in the study, principals were not able to identify the placing of their school in the efficiency distributions. It was also emphasized that the responses would be analyzed statistically so that it would not be possible to identify individual schools.

[4] The distribution of the matriculation examination scores, school size, parents' education and admission level were also very similar among 309 schools and those returning the survey. These results are not reported here but they are available upon request.

hoc rationalization (see Sammons *et al.*, 1998). That is also the case in this study.

In order to minimize problems caused by post-hoc rationalization, principals were reminded at the beginning of the survey to recall the situation during the period of efficiency measurement in 1998-1991. In addition, to control for the staff turnover, there were questions related to it. They revealed that it was not a severe problem. Concerning the turnover of principals, over 80% of those returning the survey stated that they had been in the school for the past five years acting as a principal between 1988-1991.[5] Among those not returning the survey, 11 schools (9%) informed that they were not able to return the survey because the change of the principal and the former principal being out of reach or not willing to respond. Overall, teachers' and principals' turnover is quite low in Finland. There were also three schools that were closed between the research period and the point of the survey.

Changes taking place after the period of efficiency measurement may also influence the ability to recall past practices. There was one question about the changes after 1991 in the survey. Some 57% of the principals mentioned that their school had started to follow the non-graded system and some 43% mentioned that there was a process of school level curriculum development taking place in their school.

Concerning the post-hoc rationalization, principals may have had the temptation to prettify the actual situation in the school. Teacher survey is one way to control for this kind of bias.[6] This was not possible in this study because of limited resources. It must be, however, noted that in an earlier study (Jakku-Sihvonen, 1994) instructional leadership in general upper secondary schools was studied using both principal and teacher surveys. According to the results, there were no major differences between the responses given by principals and teachers.

[5] If the principal had been in duty for a short period, they responded only to those questions they had enough information.

[6] Also the problems caused by time gap between efficiency measurement and survey would have been reduced by comparing the responses of these two groups.

4. Efficiency analysis with DEA

4.1. Method

Economic theory recognizes several efficiency concepts. In this study, the concept of *technical efficiency* is used. It is determined either as the ratio of observed to maximum potential outputs obtainable from the given inputs or as the ratio of minimum potential to observed inputs required to produce the given outputs. In the former case, school is viewed as maximizing its outputs with the given inputs and in the latter case as minimizing the use of inputs when the outputs are fixed. When efficiency is measured with DEA either of the assumptions can be applied. In this study the efficiency scores are calculated assuming that schools are minimizing their use of inputs i.e. their outputs are fixed.[7]

In DEA, a relative efficiency measure is calculated for each school by determining an efficiency frontier through the most productive schools in the data and by measuring each school's distance to the frontier. Schools at the efficiency frontier get an efficiency score of 1 (100% in this study) and schools below the frontier get an efficiency score between 0 and 1 (between 0 and 100%) depending on their distance to the frontier. Efficiency scores can be calculated either by assuming constant returns to scale (c.f. Charnes *et al*., 1978) or variable returns to scale (c.f. Banker *et al*., 1984). Efficiency scores are higher when variable returns to scale are assumed because there are more schools at the efficiency frontier. In this study, efficiency scores are calculated using both assumptions but the selection of the most efficient and inefficient schools is based on models assuming variable returns to scale.[8]

The formal optimization problem is the following when schools are minimizing their use of inputs. If we consider n schools so that school j uses the amount of x_{ij} of input i and produces the amount of y_{rj} of output r. Then, by denoting the input weights by v_i (i=1,…,m) and

[7] The assumption of output maximization could also have been used but here schools were seen as units minimizing their use of resources. In practice, the results were almost identical.

[8] For more of DEA, see e.g. Fried *et al*. (2008).

the output weights by μ_r (r =1,…,s) and by constraining them to be greater than an infinitesimal ε (in order to ensure unique solution), our basic DEA linear programming problem for the school j_0 can be written as follows, assuming constant returns to scale:

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

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

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

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

Model (1) maximizes the efficiency score for the school j_0 so that the score varies between [0,1]. Each school gets the optimal weights for inputs and outputs so that they are non-negative.

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

4.2. Models and variables

For the purposes of this study, there are two factors that complicate the choice of input and output variables in DEA. First, because of non-parametric nature of the method, no statistical tests are available to test the statistical significance of the inputs and outputs. Second, as the number of inputs and outputs increase the efficiency of the units in DEA increase because of larger number of efficient input - output combinations. Hence, the discriminating power of the method diminishes as new inputs and outputs are included. In this study, the choice of schools for the second stage is based on models with basic inputs and outputs. To test the stability of the results, a modeling strategy from simple to more complicated one is followed. An additional DEA model with inputs controlling the quality of students and teachers is also calculated. This model is used for testing the robustness of the results of other models. The models and variables are presented in Table 1. Summary statistics are reported in Appendix 2 and the precise descriptions of all the input and output variables are provided in Appendix 3.

In the simplest model 1, including only quantitative inputs and outputs, schools use teaching and other hours as their input and produce students moving to upper grade levels and students participating in the matriculation examination. Since the parental socioeconomic status strongly affects student achievement, it is added in model 2. Usually it is measured with education and income level of parents. In this study, only information on education level of parents was available and it was used in models 2-5.

Table 1. Models and variables in DEA efficiency measurement

	Model 1	Model 2	Model 3	Model 4	Model 5
Inputs					
Number of teaching and other hours per week	X	X	X	X	X
Parental education		X	X	X	X
Admission level					X
Education of teachers					X
Experience of teachers					X
Outputs					
Number of students who were moved up	X	X	X	X	X
Number of students taking the matriculation examination	X	X	X		X
Total number of scores in the matriculation examination			X	X	X

In most of the production function studies output is measured with some standardized achievement test or examination score. In this study, the total number of scores in the matriculation examination was used. It was added as an output in model 3. The grades in each test in the matriculation examination range from improbatur (failed) to laudatur (excellent) and they were converted to scores using a scale from 0 to 6. Since the quantity aspect is also present in the matriculation examination scores, the number of students taking the matriculation examination was excluded in model 4.

Model 5 is the most extensive model with three additional inputs. It takes into account the quality of teaching staff (education and experience of teachers) and earlier student achievement. Education of the teachers was measured with the number of teachers having at least a master's degree. Experience of the teachers was calculated by giving one point for each professional five year period and adding these points in each school. Earlier student achievement was measured with the average grade in comprehensive school report of the weakest accepted student (admission level). Naturally, information on all accepted students would have been a better choice but unfortunately it was not available.

4.3. Results of DEA analysis

In discussing the results of DEA efficiency measurement, the emphasis will be in demonstrating the magnitude of efficiency differences using basic summary statistics and in testing the stability of the results with Spearman rank correlation coefficients and cross-tabulations.[9] Efficiency scores for each school are calculated assuming that schools

[9] Possible outlier cases may also affect the stability of the results of DEA, i.e. there may be some schools at the efficiency frontier whose production technology is clearly different from other schools. It is not desirable that such schools serve as a reference case for other schools. Using a method called jackknifing it is possible to detect such outlier cases in the data. In jackknifing, DEA is run by dropping out each efficient unit one at the time in order to see whether the results change. This procedure is not done in this study because mostly the same data is used as in Kirjavainen and Loikkanen (1998), in which this analysis was conducted. Their analysis showed e.g. that the results were robust and the robustness increased as the number of variables in analysis increased. See also Bonesrønning and Rattsø (1994) and Färe *et al*. (1989) for their results.

are minimizing their use of inputs keeping the output constant. In such a case, the deviation of efficiency scores from one (100% in this study) indicates savings possibilities in the use of inputs. In addition, efficiency scores assuming both constant returns to scale (CRS) and variable returns to scale (VRS) are calculated. The summary statistics of efficiency scores are presented in Table 2.

Table 2. Summary statistics of DEA efficiency measurement

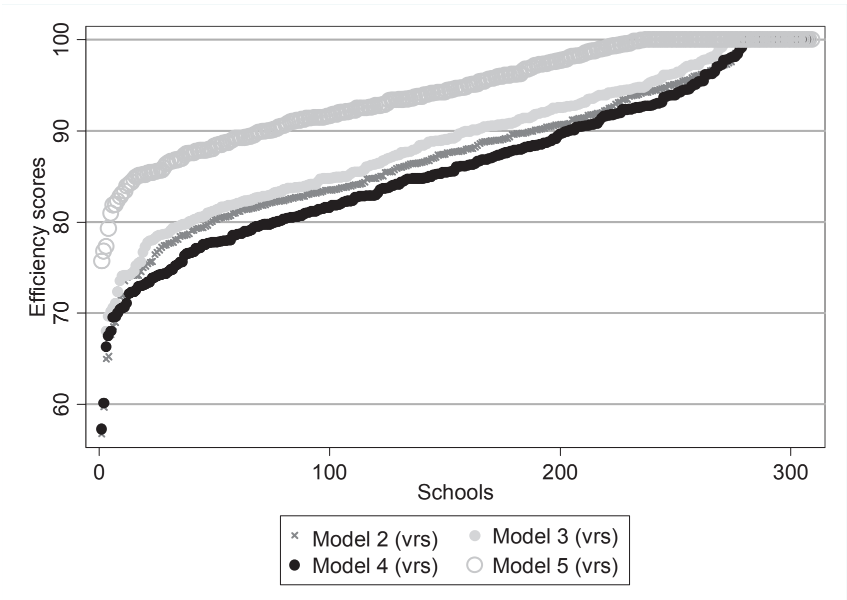
	Model 1		Model 2		Model 3		Model 4		Model 5	
	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
Average	70.85	73.94	84.61	87.54	85.73	88.79	83.19	86.08	92.09	94.13
Minimum	19.00	41.34	55.59	56.74	55.89	57.31	54.75	57.28	75.92	75.68
Maximum	100	100	100	100	100	100	100	100	100	100
Number of efficient schools (Percentage)	4 (1.3)	10 (3.2)	13 (4.2)	28 (9.1)	17 (5.5)	36 (11.7)	11 (3.6)	28 (9.1)	40 (12.9)	73 (23.6)
Average efficiency of highest 15% (Range)	90.58 (83.76 -100)	93.56 (86.69 -100)	96.43 (92.35 -100)	99.16 (96.31 -100)	97.55 (93.73 -100)	99.73 (97.87 -100)	95.96 (91.54 -100)	99.14 (96.20 -100)	99.97 (99.46 -100)	100 (100)
Average efficiency of lowest 15% (Range)	49.85 (19.00-57.32)	56.95 (41.34-62.22)	73.03 (55.59-77.24)	74.68 (56.74-79.58)	74.19 (55.89-78.34)	76.10 (57.31-80.72)	71.73 (54.75-76.04)	72.88 (57.28-77.52)	83.51 (75.92 86.44)	84.85 (75.68-87.89)

The average efficiency varied from 71% in model 1 to 92% in model 5 assuming CRS. The largest increase in average efficiency was between models 1 and 2, from 71% to 85%. In model 3, the addition of matriculation examination scores measuring the quality of output had only modest impact on average efficiency. In model 4, the dropping of the number of students taking the matriculation examination decreased the average efficiency to 83%. Adding admission level controlling for the quality of students and education and experience of teachers controlling for the quality of teachers increased the average efficiency to 92%. The variability in average efficiencies was only slightly smaller when VRS assumed. The pattern of changes was identical to the assumption of CRS.

The differences in efficiency scores between the most and least efficient schools are quite large especially in models 1, 2, 3, 4 (see also Figure 1). In model 1, the least efficient school had an average efficiency of 20% meaning that it could have used 80% less of inputs to produce its output. In model 5, the differences were smaller so that the least efficient school had an efficiency score of 76% indicating a saving potential of 24%.

As the number of inputs and outputs increase the number of schools at the efficiency frontier also increases. In model 1 there were four schools (1.3%) at the frontier and in model 5 the corresponding figure is already 40 (13%) assuming CRS. When VRS was assumed, the number of efficient schools further increased as was indicated earlier. In model 1 assuming VRS there were 10 (3%) and in model 5 already 73 (24%) schools at the efficiency frontier.

Figure 1. *Efficiency distributions of DEA models 2, 3, 4, and 5 assuming VRS*



The changes in rankings based on efficiency scores in different DEA models were examined using the Spearman rank correlation coefficient measuring the similarity of rankings between two different models.[10] The results show (see Table 3) that the rankings in model 1 are very different from those in other models. The largest differences were between models 1 and 4 with rank correlation coefficient of 0.29 when CRS was assumed. The differences were smaller when VRS was assumed. Rankings in models 2, 3 and 4 were more similar regardless of scale assumption. For instance, the rank correlation coefficient was 0.95 between the models 2 and 3 indicating that very few schools changed their rankings in these models.

Table 3. Spearman rank correlation coefficients between different models. Scale assumption in parentheses

	Model 1	Model 2	Model 3	Model 4	Model 5
Model 1	1.00 (CRS)				
	1.00 (VRS)				
Model 2	0.55 (CRS)	1.00 (CRS)			
	0.56 (VRS)	1.00 (VRS)			
Model 3	0.44 (CRS)	0.95 (CRS)	1.00 (CRS)		
	0.51 (VRS)	0.95 (VRS)	1.00 (VRS)		
Model 4	0.31 (CRS)	0.77 (CRS)	0.87 (CRS)	1.00 (CRS)	
	0.42 (VRS)	0.80 (VRS)	0.89 (VRS)	1.00 (VRS)	
Model 5	0.18 (CRS)	0.66 (CRS)	0.73 (CRS)	0.71 (CRS)	
	0.29 (VRS)	0.64 (VRS)	0.72 (VRS)	0.70 (VRS)	1.00 (VRS)

Schools among the 15% most or 15% least efficient schools in models 2, 3, or 4 assuming VRS were selected for the second stage analysis. By using this rule, around one third of the schools were

[10] The value of 1 of Spearman rank correlation coefficient indicates that the rankings are exactly the same. The value of 0 indicates that there is no connection between the rankings, and the value of –1 indicates that the rankings are reverse.

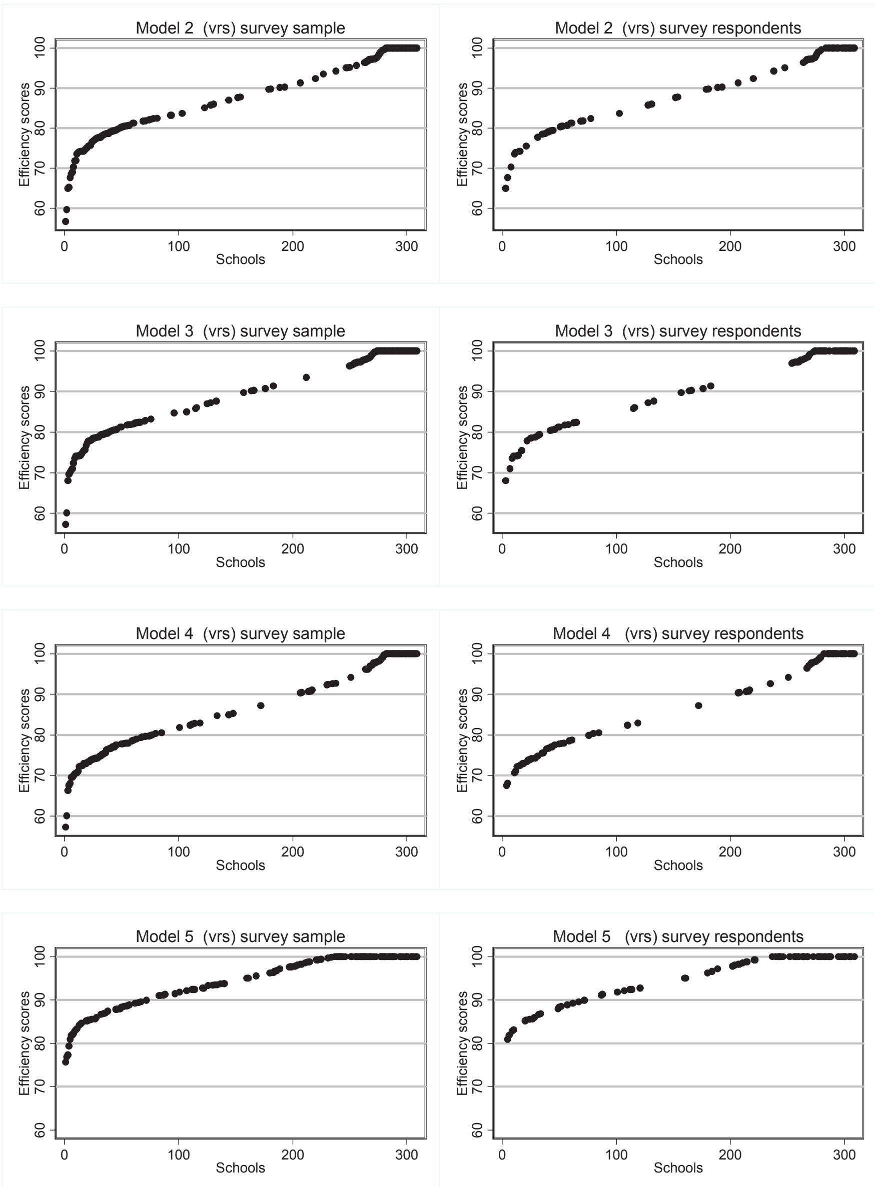
included. This sample was considered large enough. It also separated individual schools.

The average efficiency of the high performers exceeded 90% in all the models (see Table 2). In model 1 assuming CRS it was 91% and in model 3 assuming VRS it was close to 100%. The average efficiency of the low performers ranged from 50% in model 1 (assuming CRS) to 76% in model 3 (assuming VRS). Hence, there were clear differences in average efficiency across these two groups.

The stability of the groups of high and low performers across different DEA models was studied by cross-tabulation (see Appendix 4 and Appendix 5). The results showed that the group of high performers was more stable than the group of low performers. Concerning the effect of scale assumption, the group of top performers was somewhat more stable assuming VRS, whereas the group of low performers was more stable when CRS was assumed. The rankings of most of the schools especially at the top of the distribution remained relatively unchanged. The lower tail of the distribution is more vulnerable to the addition of new inputs and outputs. This is a fact that complicates the analysis at the second stage.

Distributions of efficiency scores for the high and low performers and for the schools responding to the survey are depicted in Figure 2. The figure shows that schools returning the survey are mostly located at the tails of the distributions. Because schools were selected based on their performance in three different models, there are always some schools that are also in the middle of the distribution. The figure also shows that non-responding was quite randomly distributed.

Figure 2. Distributions of efficiency scores in models 2-5 assuming variable returns to scale for schools in the survey sample and schools responding to the survey



5. Efficiency differences and organizational characteristics

Since most of the DEA models included only few inputs and outputs, the effect of some additional factors related to school resources, student body and school environment on efficiency are first tested. Thereafter, descriptive statistics and statistical models examining the influence of organizational factors or practices related to structure, goals, decision making, staff development, monitoring of teachers and students, instructional development, policy concerning parents and satisfaction of principals over the operation of school on efficiency are estimated.[11]

Earlier, the division of schools into two groups (efficient vs. inefficient) based on their efficiency scores in DEA models 2, 3 and 4 were described. Additional groups were also formed taking into account the results of model 5 in order to test the stability of the results. These groupings are described in Appendix 6. Descriptive statistics of the organizational characteristics based on these groupings are reported in the Appendices 7-11. They are also shortly commented in the text but the main emphasis will be on describing the results of statistical probit and tobit models.

When the schools at the tails of the efficiency distribution are compared, the dependent variable can be described as a dichotomy i.e. school is classified either as efficient (it was among the 15% most efficient schools in at least one of the three models) or as inefficient (it was among the 15% least efficient schools in at least one of the three models). In such a case, the dependent variable is not a continuous one and therefore the standard OLS estimation would produce biased results. Discriminant analysis is usually used in this kind of setting to determine factors discriminating the two groups. According to Maddala (1983) it would not, however, produce consistent estimates when the assumption of normality is not true for the independent variables (in this study e.g. dummy variables are used). Instead, probit-

[11] The results of the survey have been reported in more detail in Kirjavainen (1997).

models are considered more appropriate.[12] They yield probabilities as a function of explanatory variables for the occurrence i.e. for the school being efficient. The model is illustrated in the following.

Assume that there exists a continuous index Y_i^* that is determined by the vector of explanatory variables X_i of the form

$$Y_i^* = \beta X_i + \mu_i \tag{5}$$

Observations on Y_i^* are not, however, available. Instead, the available data only distinguish whether school is efficient (high values of Y_i^*) or inefficient (low values of Y_i^*). Probit models produce estimates for the parameters β while at the same time obtaining information about the underlying index Y_i^*. Specifically, if the dependent variable for school i is a dummy variable Y_i having values of 1 (school is efficient) or 0 (school is inefficient). Then it is assumed that Y_i^* represents the critical cutoff value translating the underlying index into occurrence observed i.e.

$$Y_i = 1 \text{ if } Y_i^* > 0 \tag{6}$$

$$Y_i = 0, \text{ otherwise.}$$

This specification enables the model to be estimated by the maximum likelihood method by assuming normally distributed errors μ_i.

In addition to probit models, the robustness of the results were tested with a tobit models using the efficiency scores of model 5 assuming VRS as the dependent variable. As a consequence of including admission level and education and experience of teachers in DEA model (model 5) the efficiency of some schools increased so that the efficiency distribution was continuous (see Figure 2). Because

[12] It must be noted that there are also problems associated with the use of probit models that are related to the fact that observations are not normally distributed as is assumed in probit models. Another problem is related to the small number of observations which may cause the estimates to be unstable.

the efficiency scores are continuous and limited to 1 in the upper tail of the distribution and to 0 in the lower tail, tobit models were considered appropriate for the estimation. The estimated model is thus following (Maddala, 1983)

$$Y_i^* = \beta X_i + \mu_i \qquad (7)$$

where Y_i^* is a latent variable referring to the technical efficiency of schools. The observed data is, however, given in the following

$$\begin{aligned} Y_i &= L_{1i} \quad \text{if } Y_i^* \leq L_{1i} \\ Y_i &= Y_i^* \quad \text{if } L_{1i} < Y_i^* < L_{2i} \\ Y_i &= L_{2i} \quad \text{if } Y_i^* \geq L_{2i} \end{aligned} \qquad (8)$$

where L_{1i} and L_{2i} are the lower and upper limits of the data. Since there are no observations at the lower limit the data is estimated with right censored tobit model in Stata.

5.1. Resources and environmental characteristics

Efficiency measurement and subsequent choice of schools included in the second stage analysis was based on three different models with limited number of inputs. As for student background, these models only controlled parental education. Therefore, it is tested if the group of efficient and inefficient schools differed in terms of *admission level* in the second stage.[13]

Additional measures that were not possible to include in DEA analysis include *heterogeneity of the student body* which takes into account the different demands set for the instruction and the use of

[13] In Kirjavainen (1997) also estimate of principals concerning the percentage applicants accepted as well as estimate concerning the percentage of accepted students placing the school as their first alternative were considered. However, these variables did not discriminate efficient and inefficient schools.

resources through e.g. different grouping practices. It is measured with standard deviation in grades in the matriculation examination within the school. *Proportion of female students* is another factor related to student body composition. It tests the gender effect on efficiency. As for *education* and *experience of teachers*, since they were included only in model 5, their effect is also tested in the second stage.

In production function studies, the effect of class size and school size is of primary interest. The class size effect is not tested in this study and scale effects are already taken into account in DEA models since schools were selected on the basis of their efficiency scores assuming variable returns to scale (VRS).

Physical facilities may affect student learning if there is lack of equipment or materials. In addition, if the physical condition of the school premises is poor or the space in the classrooms is too limited it may have consequences on learning through students' and teachers' satisfaction. The data used in efficiency measurement did not contain any information on school facilities or equipment. In the survey, some indirect information about the physical facilities of schools was collected. It concerned principal's assessment of the space in classrooms and instructional equipment. Its effect is tested in statistical models.

Parental involvement may be considered from different point of views. On the one hand, there are the actions of parents in the form of encouragement and support for their children. On the other hand, there are the actions taken by the school in enhancing involvement of parents in the schoolwork. The former one is more related to environmental characteristics of schools, whereas the latter one can be considered as a policy chosen by the school. In this study, the first view is considered by using the estimate of principals concerning *the attendance of parents in parent meetings* and testing its effect on efficiency.

Finally, there is also information related to the type of municipality available and it is used as a dummy structure indicating whether the school is located in the *urban*, *densely* populated or in *rural* area.

5.2. Organizational characteristics

The main concern of this study is to examine the effect of certain organizational factors and practices on efficiency. Formal structure of the organization is important and it influences on the way routines are carried out in the school. It may also facilitate for instance participation of teachers in school matters and decision making. Information concerning the *structure of organization* was obtained from the survey by requesting principals to draw a chart on their school and to specify the tasks of each group mentioned in it. When the responses were coded, two groups of schools emerged. In most of the schools, the formal organization was described as traditional. In such a case, principals described that there were teachers, principal, students and other staff in the school. Another group was formed of schools that had a management group and/or working groups consisting of teachers responsible for the preparation and planning of various activities related to school matters and instruction.

The results concerning effective schools emphasize the importance of clearly defined goals as one of the features of good leadership. These goals should also have a wide acceptance among the teachers and other staff in order to be effective. Other related "instruments" - perhaps more common in corporate settings - are a mission and a strategy for carrying out the goals. In this study, there is information on whether schools had officially defined *mission, goals, and/or strategy*. In addition, the assessment of principals concerning the *consensus about the goals* among teachers is also used.

Participative decision making involving teachers is one of the characteristics of effective schools. There are two dimensions related to it. One is the possibility of teachers to influence decision making at the school site and another is the autonomy of schools, especially in relation to municipalities, in making decisions concerning themselves. Variables depicting both of these dimensions are used. The *influence of teachers in decision making* is tested with two alternative variables. The first one is constructed from the statements of principals concerning the proposal and decision making in the following matters: Goals and strategy, resource allocation, teacher resources, content of instruction, teaching methods and instructional materials.

The second variable describes the process of short range planning. It is constructed from the question asking principals to describe how the process was carried out, i.e. who were involved and what were the responsibilities of each of the participant. In the coding, schools were classified as either following the *bottom-up* or *top-down approach*. The first one is characterized as a process in which teachers take part and make proposals right at the beginning. The final working plan of the coming year was a product of the proposals of the teachers and other staff. The latter approach was characterized as one in which principal took care of all the major phases of the process. The role of the teachers was more or less to formally approve the proposals of the principal.

Autonomy of schools in decision making has been one of the arguments in favor of better performance of private schools in the U.S. (see e.g. Chubb and Moe, 1990). It allows schools to be more responsive for instance to problem situations. In the survey, it turned out that in Finland only the hiring of new teachers was an issue differentiating schools so that proposals or decisions were also made by the school board or education committee[14]. All the other decisions were mainly made at the school site. Therefore, only variable depicting the *hiring decisions of new teachers* is considered in this study.

On the basis of earlier research, staff development and monitoring are important features of effective schools. In this study, some information related to these factors is available. There is data on whether or not schools made *staff development plans* indicating the consistency of staff development. In addition, information on the *teachers' participation in training* is used. As for *monitoring of performance of teachers*, information on how often principals conducted personal monitoring sessions with the teachers is used.

The performance *monitoring of students* was also one of the features of effective schools. Levin (1997) also mentions it as one of the ways to improve x-efficiency. In this study, there is information on how

[14] The education committee is a local-government body that directs and supervises the municipal school system. The education committee is elected by the local council.

often the performance of students was monitored by the principal and staff during the school year.

Instructional matters were for the most part left out of this study. However, the effect of some factors related to it is tested. They concern the developing of instruction. In Finland, schools often proceed in this matter by *participating* in locally or nationally organized *development projects* or experiments. In the survey, principals were asked if their school had taken part in any such projects during the research period. Another factor is the introduction of new instructional methods. The projects crossing the subject boundaries were especially popular. In this study, there is information on the number of *joint projects across different subjects* in a school year.

Parental involvement was earlier considered from the point of view of parental support for their children and schoolwork. Another dimension concerns the role of schools in promoting parents' participation. In other words, schools may have policies that encourage parents to take actively part in schoolwork and inform parents about the school matters. The variable used in this study depicts whether schools had policies going beyond the usual parent meetings i.e. if schools had parent associations, bulletins or other practices that increased the involvement of parents in school matters.

Other factor, which was tested concerned after school activities and it was measured with the *number of clubs* in the school. In addition, evaluation of the principals on the following matters was considered: *Satisfaction with the teachers' preparation of* classroom work, *how easy it was to get things done* with the teachers, and an *overall satisfaction* on a scale from 4 to 10.

5.3. Estimation results

Results concerning the student body and environmental characteristics are very much in line with earlier production function studies (see Table 4). Admission level and share of female students was somewhat higher, and heterogeneity of student body somewhat lower in efficient schools. These results were also quite robust when schools were grouped in different ways based on their efficiency (see Appendix 8). In probit model 1A, the probability of being efficient (PE) increased if admission level (at 1% significance level) and the percentage of female students (at 5% significance level) were high whereas heterogeneous student body decreased the PE at 5% significance level.

Teachers' education and experience was not related to PE in probit model 1A, whereas principals' estimate of lack of equipment and school facilities increased PE at 10% significance level. Measure for parental involvement had statistically insignificant coefficient. Schools in densely populated areas were more likely to be efficient than schools in urban and rural areas at 5% significance level in probit model.

Table 4. Parameter estimates of probit and tobit models. Explanatory variables related to student body, school resources and school environment

	Probit 1A		Tobit 1A (Model 5)	
	Coefficient	z-ratio	Coefficient	t-ratio
Constant	-10.953	-2.41**	0.9354	7.98
Student body composition				
Admission level	1.505	3.04***		
Heterogeneity	-3.215	-2.04**	-0.182	-1.87*
Females (%)	0.0477	2.42**	0.002	1.82*
Education of teachers	-1.365	-0.73		
Experience of teachers	0.0742	0.24		
Parental involvement				
Attendance in parent meetings	0.0844	0.27	0.008	0.38
Physical facilities	0.545	1.66*	0.040	2.13**
Municipality type				
Rural	0.470	0.97	0.021	0.81
Dense	1.565	2.23**	0.039	1.08
Proportion of efficient schools	38/71		32/71	
Proportion of right predictions	76.1			
Pseudo R^2	0.318			
Log-likelihood	-33.456		19.623	

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

The results of tobit model using the efficiency scores of model 5 (VRS) as the dependent variable supported the results of probit models. Heterogeneity of the student body decreased efficiency and share of female students increased it. Both coefficients were statistically significant at 10% significance level. In efficient schools, principals were more likely to report that there was lack of space and equipment in the school (at 5% significance level). Municipality type was not related to efficiency in this model.

Because in both probit and tobit models 1A variables measuring the student body composition were statistically significant, they were also included in models testing the effect of organizational factors in order to control for student body. Since a number of factors were used to characterize the organization of schools, separate models for different groups of variables were estimated. The first group of variables was related to the structure of organization, formal management practices, and decision making. The second group of variables tested the effect of development and monitoring practices. Finally, the third group consisted of variables which describe the policy concerning parent-school relations, after-school activities and the satisfaction of principals for various aspects of the school. The estimation results are presented in Table 5 and Table 6.

Overall, the results show that most of the variables measuring organizational characteristics and practices were not related to efficiency. Furthermore, in some of the cases the sign of the effect was unexpected. The variables measuring student body composition were more robust but in some models also their effect was not statistically significant.

Probit and tobit models 2A and 2B tested the effect of student body, structure of organization, mission and goals and decision making practices on efficiency. Means of the variables based on different groupings are presented in Appendix 9. They show that in over half of the schools organization structure was characterized as one without management or other groups assisting and preparing matters related to whole school. Use of management or other assisting groups was more common in efficient schools according to grouping 1 and 4. In groupings 2 and 3 these differences diminish. Planning was characterized as bottom-up in over half of the schools. This approach was more common in inefficient schools. The results were robust to different groupings of schools.

Concerning the autonomy of schools in hiring teachers, most often schools were able to make proposals that were confirmed by the local school board. This procedure was more common in efficient schools based on grouping 1. However, the results changed in other groupings and the differences between efficient and inefficient schools became smaller. In almost all schools principals responded that they had shared, clearly defined goals. Most of the principals also felt that there was consensus about them. Principals of inefficient schools had more often the view that teachers were very unanimous about these goals.

The results were similar when these factors were tested together with the student background factors in probit model. Higher admission level increased the PE at 1% significance level. PE also increased if school had an organization structure with management group or other working groups. The more schools used such management tools as mission, goals and strategy the less likely they were to be efficient. The coefficients of these two variables were statistically significant at 1% level. The autonomy in hiring new teachers increased the PE at 10% significance level. Other factors depicting decision making were not related to efficiency in probit models.

In tobit models 2A and 2B factors related to student body were not statistically significant. Schools with mission, goals and strategy were less likely to be efficient also in these models at 5% significance level. The higher unanimity about the school goals decreased the PE at 5% significance level. Other factors were not related to efficiency i.e. autonomy of school in hiring teachers and influence of teachers in decision making were not statistically significant in explaining efficiency.

Table 5. *Parameter estimates of probit and tobit models. Explanatory variables related to student body, structure of organization, school goals, and decision making*

	Probit 2A		Tobit 2A (Model 5)		Probit 2B		Tobit 2B (Model 5)	
	Coefficient	z-ratio	Coefficient	t-ratio	Coefficient	z-ratio	Coefficient	t-ratio
Constant	-7.613	-1.72*	1.049***	7.70	-6.330	-1.38	1.038	7.29***
Student body composition								
Admission level	1.327	2.48***			1.051	1.85*		
Heterogeneity	-3.631	-1.85*	-0.161	-1.33	-2.801	-1.47	-0.151	-1.20
Female (%)	0.018	0.79	0.001	0.63	0.027	1.21	0.001	0.97
Organizational structure	1.325	2.55***	0.036	1.31	1.340	2.50***	0.035	1.21
Mission, goals and strategy	-1.973	-2.46***	-0.109	-2.45**	-2.011	-2.26**	-0.108	-2.26**
Consensus about the goals	0.278	0.84	0.041	1.96*	.256	0.77	0.044	2.01**
Decision making								
Influence of teachers	0.613	0.83	0.035	0.78				
Bottom-up planning					-.338	-0.58	-0.013	-0.38
Principal proposal, educ. committee decisions in hiring new teachers	1.253	1.88*	0.016	0.49	1.221	1.69*	0.015	0.40
Educ. committee proposals and decisions in hiring new teachers	0.568	0.85	-0.005	-0.12	0.649	0.93	0.009	0.21
Proportion of efficient schools	32/55		27/55		33/55		28/55	
Proportion of right predictions	72.7				76.4			
Pseudo R^2	31.7				33.0			
Log-likelihood	-27.406		11.152		-25.061		9.904	

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

Staff development plans were completed in some 40% of the schools (see Appendix 10). They were made more often in efficient schools. Estimates of principals concerning the participation of teachers in training were higher in inefficient schools. In about half of the schools principals reported that they did not have any individual discussions or monitoring sessions with teachers. In about one third of the schools these sessions were conducted once a school year. There were no differences between efficient and inefficient schools with these practices. Some 20% of the schools in both groups had been involved in some development projects.

Staff development and monitoring practices were tested in probit model 2C. In this model admission level was not statistically significant whereas the heterogeneity of the student body affected negatively on PE at 5% significance level. Concerning the development and monitoring practices, PE increased the less teachers participated in training[15] (at 5% significance level) and the less there were joint teaching projects across different subjects (at 5% significance level). Staff development plans, monitoring of the performance of teachers and students were not related to efficiency. In tobit models, only participation of teachers in training affected negatively on PE at 10% significance level. Other organizational factors were statistically insignificant in explaining PE.

[15] In Kirjavainen (1997) the content of training and organizers of training were also discussed. In efficient schools principals estimated the amount of training taking place in universities and training concerning pedagogic skills somewhat higher than principals of inefficient schools. Apart from these results there were no differences between efficient and inefficient schools.

Table 6. *Parameter estimates of probit and tobit models. Variables related to student body, staff development, monitoring, development of instruction, parent-school relations, and satisfaction*

	Probit 2C		Tobit 2C (Model 5)		Probit 2D		Tobit 2D (Model 5)	
	Coefficient	z-ratio	Coefficient	t-ratio	Coefficient	z-ratio	Coefficient	t-ratio
Constant	0.031	0.01	1.093	7.69***	6.517	1.23	1.338	4.84***
Student body composition								
Admission level	0.427	0.96			0.218	0.45		
Heterogeneity	-3.918	-2.08**	-0.197	-1.74*	-4.365	-2.42**	-0.227	-2.05**
Female (%)	0.024	1.50	0.002	1.46	0.030	1.50	0.002	1.58
Staff development and monitoring								
Plans	0.623	1.43	0.035	1.25				
Participation in training	-0.427	-2.10**	-0.024	-1.90*				
Monitoring of teachers	0.090	0.46	0.014	1.08				
Monitoring of students	0.131	1.10	0.010	1.13				
Development of instruction								
Participation in projects	-0.044	-0.08	-0.052	-1.49				
Teaching across subjects	-0.350	-2.06**	-0.013	-1.32				
Policy of parent-school relations					0.072	0.17	0.009	0.29
Active clubs					0.146	1.91*	0.004	0.85
Overall satisfaction					-0.231	-0.59	-0.025	-0.92
Satisfaction with teachers					-0.536	-1.43	-0.010	-0.44
Getting things done with teachers					-0.111	0.26	-0.002	-0.07
Proportion of efficient schools	32/61		29/61		32/60		26/60	
Proportion of right predictions	70.5				76.7			
Pseudo R^2	28.6				23.8			
Log-likelihood	-30.150		14.687		-31.581		16.741	

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

In probit and tobit models 2D a mixture of additional factors were tested (see also Appendix 11). Concerning policy of parent-school relations, one third of the schools had school bulletins, parent association and other activities in addition to parent meetings. They were more common in efficient schools. There were on average some 3 active clubs in the schools. The number of active clubs was higher in efficient schools than in inefficient schools. Overall, principals were quite satisfied with their school and teachers and there were no clear differences in this matter between efficient and inefficient schools.

In probit and tobit models including these factors as explanatory variables, heterogeneity of the student body decreased efficiency at 5% significance level. Other factors related to student body composition were statistically insignificant in explaining efficiency. Variables related to policy of parent-school relations and principals' satisfaction were also statistically insignificant in explaining efficiency. The number of active clubs and efficiency were positively related in probit model 2D at 10% significance level.

6. Conclusions

This study first estimated the efficiency differences of Finnish general upper secondary schools with five different DEA models. According to results there were efficiency differences between the schools. These differences varied depending on the model. The average efficiency was between 71% and 94% indicating that for instance in the latter case schools could have used on average 6% less of their inputs to produce their outputs.

In the second stage, schools among the 15% most (high performers) and least efficient (low performers) ones in at least one of three different DEA models were further analyzed with statistical probit and tobit models. The second stage analysis contained information on 73 schools and it was based on survey data. The survey included information on student body, parent school relations, school management, staff development and monitoring, development of instruction, and materials and facilities of the schools.

Admission level and some other student characteristics were not included in DEA models affecting the choice of the schools for the second stage analysis. Therefore, they were included as controls in the second stage analysis. According to results, schools with higher admittance level, less heterogeneous student body and higher proportion of female students were more likely to be efficient in some models.

Concerning teachers' education and experience and parental involvement, they were statistically insignificant in predicting the probability of being efficient. Schools reporting lack in facilities were more likely to be efficient. Schools in densely populated areas were somewhat more likely to be efficient than schools in urban areas, whereas schools in rural areas were not different from schools in urban areas.

The effect of organizational characteristics was studied in four different models. In general, coefficients depicting these factors were fairly unstable. Schools having organization structure with management group or other working groups responsible for the preparation of tasks related to school and instruction were more likely to be efficient. The use of such management tools as mission statement, goals, and strategy decreased the probability of being efficient. Consensus about the school goals lowered efficiency in some models. Other factors, such as influence of teachers in decision making and autonomy of schools in making hiring decisions of new teachers, were only weakly related to efficiency in some models.

The effect of development and monitoring practices were also fairly unstable and somewhat surprising. In some models, the probability of being efficient increased the less teachers participated in training and the less there were joint teaching projects crossing subject boundaries. The preparation of staff development plans or monitoring performance of teachers and students were not related to efficiency.

Policies concerning parent-school relations and the satisfaction of principals on the overall operation of the school, on the work of the teachers and on the cooperation with the teachers were not related to efficiency. Schools with higher number of active clubs were more likely to be efficient.

To conclude, there are efficiency differences between schools and they are dependent on the DEA model used. However, only few organizational practices were related to these differences and these results were not robust to model specifications. Therefore, studies concentrating only on the top performing schools when investigating factors related to efficiency may give somewhat distorted picture on the possibilities to improve efficiency. Concerning student background factors, their influence is more robust and in accordance with the expectations.

References

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

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

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

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

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

Bradley, S. – Johnes, G. – Millington, J. (2001): The Effect of Competition on the Efficiency of Secondary Schools in England. *European Journal of Operational Research* 135, 545–568.

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

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

Chubb, J. E. – Moe, T. M. (1990): *Politics, Markets and America's Schools*. Washington D.C.: The Brookings Institution.

Dodd, A. (2006): Investigating the Effective Use of Resources in Secondary Schools. *Research Report to DfES*. Reserch Report No 779. London: Hedra plc.

Duncombe, W. – Miner, J. – Ruggiero, J. (1997): Empirical Evaluation of Bureaucratic Models of Inefficiency. *Public Choice* 93, 1–18.

Fried, H. – Lovell, K. C. A. – Schmidt, S. S., (eds.) (2008): *The Measurement of Productive Efficiency and Productivity Growth.* New York: Oxford University Press.

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

Jakku-Sihvonen, R. (1994): General Upper Secondary School as a Working Environment of Principal and Teacher (in Finland). In Jakku-Sihvonen, R. – Blom, H. (eds.) *Lukion tila 1994*. Helsinki: Opetushallitus.

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

Kirjavainen, T. (1999): Efficiency of Finnish Senior Secondary Schools: Explaining High and Low Performance with Resources, Environmental, and Organizational Factors. *Helsinki School of Economics and Business Administration Working Papers W-221*. Helsinki: Helsinki School of Economics and Business Administration.

Kirjavainen, T. (1997): *Efficiency of Schools: The Impact of Environmental and Organizational Factors on Efficiency Measured by DEA*. Licentiate Thesis. Helsinki: Helsinki School of Economics and Business Administration.

Kirjavainen, T. – Loikkanen, H. A. (1998): Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit-Analysis. *Economics of Education Review* 17, 377–394.

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

Lee, V. E. – Bryk, A. S. – Smith, J. B. (1993): The Organization of Effective Secondary Schools in L. Dardling Hammond (ed.) *Review of Research in Education* 19. Washington D.C.: American Educational Research Association.

Levin, H. M. (1997): Raising School Productivity: An X-Efficiency Approach. *Economics of Education Review* 16, 303–311.

Levine, D. U. (1992): An Interpretive Review of US Research and Practice Dealing with Unusually Effective Schools. In D. Reynolds and P. Cuttance (eds.) *School Effectiveness. Research, Policy and Practice*. London: Cassell.

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

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

McCarty, T. A. – Yaisawarng, S. (1993): Technical Efficiency in New Jersey School Districts, in Fried *et al.* (eds.) *The Measurement of Productive Efficiency. Techniques and Applications.* New York: Oxford University Press.

Mayston, D. – Jesson, D. (1988): Developing Models of Educational Accountability. *Oxford Review of Education* 14, 321–339.

Portela M. C. – Camanho, A. S. (2007): Performance Assessment of Portuguese Secondary Schools. Working paper, economics 07/2007. Universidade catolica Portuguesa (Porto). http://dspace.feg.porto.ucp.pt:8080/dspace/bitstream/2386/100/1/072007+Portela+e+Camanho+-+Performance+assessment+of+portuguese+secondary+schools.pdf.

Purkey, S. C. – Smith, M. S. (1983): Effective Schools: A Review. *The Elementary School Journal* 83, 427–452.

Ray, S. C. (1991): Resource-use Efficiency in Public Schools: A Study of Connecticut Data. *Management Science* 37, 1620–1628.

Reynolds, D. – Creemers, B. – Nesselrodt, P. – Schaffer, E. v Stringfield, S. – Teddlie, C. (1994): *Advances in School Effectiveness Research and Practice*. Oxford: Pergamon.

Reynolds, D. (1992): School Effectiveness and School Improvement: An Updated Review of the British Literature. In D. Reynolds and P. Cuttance (eds.) *School Effectiveness. Research, policy and practice*. London: Cassell.

Sammons, P. (1999): *School Effectiveness. Coming of Age in the Twenty-First Century*. Lisse: Swets & Zeitlinger.

Sammons, P. – Thomas, S. – Mortimore, P. – Walker, A. – Cairns, R. – Bausor, J. (1998): Understanding Differences in Academic Effectiveness: Practitioners' Views. *School Effectiveness and School Improvement* 9, 286-309.

Scheerens, J. – Bosker, R. (1997): *The Foundations of Educational Effectiveness*. Oxford/New York/Tokyo: Pergamon.

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

Smith, P. – Mayston, D. (1987): Measuring Efficiency in the Public Sector. *OMEGA International Journal of Management Science* 15, 181-189.

Teddlie, C. – Reynolds, D. (2000): *The International Handbook of School Effectiveness Research*. London/New York: Routledge.

Waldo, S. (2003): Efficiency in Education. Ph.D. Thesis. Department of Economics, Lund University.

Appendices

Appendix 1. Average efficiencies, matriculation examination grades, school size, and type of municipality of efficient and inefficient schools (Grouping 1) in the survey sample and respondents

	Efficient schools		Inefficient schools	
Efficiency scores	Original group (N=58)	Final data (N=40)	Original group (N=72)	Final data (N=33)
Model 1 (CRS)	76.78	74.62	62.19	62.45
Model 1 (VRS)	82.00	80.78	65.53	66.24
Model 2 (CRS)	92.39	93.05	76.52	78.25
Model 2 (VRS)	97.49	97.84	78.02	79.84
Model 3 (CRS)	94.04	94.85	77.82	79.49
Model 3 (VRS)	99.19	99.28	79.41	81.01
Model 4 (CRS)	91.94	92.93	74.56	75.26
Model 4 (VRS)	97.35	97.56	75.71	76.19
Model 5 (CRS)	96.93	97.56	88.07	88.80
Model5 (VRS)	99.59	99.70	89.67	90.20
Average scores in the matriculation examination	4.28	4.31	4.17	4.04
School size	231	207	139	131
Parental education	2.92	2.85	3.16	3.04
Admission level	6.99	7.02	6.78	6.79
Type of municipality				
Urban (%)	44.8	32.5	44.4	39.4
Densely populated area (%)	17.2	22.5	13.9	9.1
Rural (%)	37.9	45.0	41.7	51.5

Appendix 2. Summary statistics of the variables in DEA analyses

	Mean	Std. dev.	Min.	Max.
	Inputs			
Number of teaching and other hours per week	202.4	97.2	50.8	616.7
Parental education	168.0	113.9	18.0	894.6
Admission level	369.22	221.22	59.76	1487.5
Education of teachers	10.7	4.3	3	32.5
Experience of teachers	54.7	27.6	2	174.0
	Outputs			
Number of students who were moved up	177.1	97.2	26.5	603.5
Number of students taking the matriculation examination	53.5	30.1	9.0	186.0
Total number of scores in the matriculation examination	1175.5	701.2	182.0	4788.0

Appendix 3. Description and source of variables

Variable name	Definition	Source
	INPUTS AND OUTPUTS IN DEA-MODELS	
Number of teaching and other hours per week	Measured in weeks. Average over the school years of 1989-91. Unfortunately, it was not possible to obtain information on teaching and other hours concerning the first school year 1988-89 when the students graduating in the spring of 1991 started their school.	Official register
Education of teachers	Number of teachers with at least master's degree. Average over 1989-91.	Official register
Experience of teachers	Cumulative teachers' experience. Average over the years 1989-91. Lacking direct information salary information is utilized since teachers get upperity bonuses according to a certain time schedule. The information on teacher specific bonuses indirectly based on the number years they have taught is used to calculate a proxy variable of teacher experience.	Official register
Parental education	The variable is constructed by multiplying the average educational level of both biological parents whose children matriculated in the spring of 1991 by the number of students taking matriculation examination. Parents' educational level is measured by giving points for degrees as follows: 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. The detailed information on parents' educational level was obtained from Statistics Finland.	Official register
Admission level	Students to general upper secondary schools are chosen on the basis of comprehensive school grade point average. The grades are given by teachers. Admission level is the grade of the last accepted student. Grades range from 4 to 10. The grade is multiplied by the number of students in the first grade.	Official register
Number of students who were moved up	An average over the school years from 1989 to 1991.	Official register
Number of students taking the matriculation examination	Spring of 1991.	Official register
Total number of scores in the matriculation examination	A cumulative sum over the scores of all students participating in matriculation examination both in compulsory and optional subjects in the spring of 1991. The matriculation examination score in each subject has a range from one (improbatur = fail) to six (laudatur). The grade in compulsory subjects consists of mother tongue, the second national language of the country, foreign language, and mathematics (comprehensive or a short course) or science and humanities. Optional subjects may include grades in foreign language, mathematics (short course), and science and humanities. The science and humanities examination covers a wide range of subjects. Because students may choose the questions they answer the same score in the matriculation examination can be achieved in various ways.	Official register

Appendix 3. continues

VARIABLES EXPLAINING EFFICIENCY		
RESOURCES AND ENVIRONMENTAL CHARACTERISTICS		
Admission level	Comprehensive school grade point average of the last accepted student.	Offical register
Education of teachers	Proportion of teachers with at least a master's degree. Average over the years 1989-91.	Official register
Experience of teachers	Average experience of teachers. Average over the years 1989-1991.	Official register
Heterogeneity	Within school standard deviation of average grades given by teachers at the end of general upper secondary school. Grades range from 4 to 10.	Official register
Female	Proportion of female students.	Official register
Attendance of parents in parent meetings	Coded from closed question of how large proportion of parents attended parent meetings in general. 1=less than 25 %; 2=25-50 %; 3=over 50 %. Principals' estimate.	Survey
Facilities	An average over following three questions concerning facilities of schools. Did you have large enough classrooms? 1=Yes; 2=Yes, with some exceptions; 3=There was not enough room in the classrooms. Did the quality of classrooms (e.g. the shortage of special classes) restrict the schoolwork? 0=Never; 1=Sometimes; 2=Often; 3=Very often. Did the quality of equipment restrict the schoolwork? 0=Never; 1=Sometimes; 2=Often; 3=Very often. Principal's estimate.	Survey
Municipality type	A dummy structure. Urban area=Urban; Densely populated area=Dense; Rural area=reference case.	Official register

ORGANIZATIONAL CHARACTERISTICS		
Organizational structure	A variable indicating whether the school had something else than a traditional organizational structure. Coded from an open question in which principals were asked to characterize their organization of school by drawing a chart. 0=No; 1=Yes.	Survey
Mission, goals and strategy	Average over three questions. Did the school have a mission statement? 0=No: 1=Yes; Did the school have determined goals? 0=No; 1=Yes; Did the school have a strategy? 0=No; 1=Yes.	Survey
Consensus about goals	Estimate of the principal about the consensus of goals among teachers. How much there was consensus among teachers about the school goals? 1=Very much; 2=To some extent; 3=Disagreement to some extent; 4=Disagreement to large extent.	Survey
Influence of teachers in decision making	Indicates if teachers were involved in making proposals or decision in the following matters. An average over the following four factors: a) Goals and strategy of the school, b) allocation of resources, c) allocation of teacher resources. 0=No; 1=Yes.	Survey

Appendix 3. continues

Bottom-up planning	Indicates if the school was classified as having a bottom-up short-range planning process. 0=No; 1=Yes. Coded from an open question.	Survey
Hiring of teachers	A dummy structure. Reference case=Principal made proposals and principal/local school board made decisions; Principal made proposals and education committee decisions; Education committee made proposals and decisions.	Survey
Plans	Indicates if development plans were made in the school. 0=No; 1=Yes.	Survey
Participation in training	Coded from a closed question. Indicates how large proportion of teachers participated more than three days a school year in training. 0=None of the teachers participated at least three days a year in training; 1=Less than 25 %; 2=25-50 %; 3=51-75 %; 4=over 75 %.	Survey
Monitoring of teachers	Coded from a closed question. Indicates how often principals had personal monitoring sessions with teachers. 0=Did not have sessions; 1=Had a session once a school year; 2=Had a session once a semester; 3=Had a session more than once a semester.	Survey
Monitoring of students	Indicates the number of times the performance of students was monitored together with the teachers in a semester.	Survey
Participation in projects	A dummy indicating if school was taking part to any larger development project organized by the school, municipality or e.g. Finnish national board of education. 0=No; 1=Yes.	Survey
Teaching across subjects	Number of joint projects across different subjects. Principal's estimate.	Survey
Policy of parent-school relations	A dummy variable. Coded from an open question stating the policy of parent-school relations. Schools indicating as having no particular policy, or just having regular parent meetings were the reference case. In the comparison group schools had in addition to parent meetings school bulletins, parent associations, or other activities.	Survey
No. of active clubs	Average number of active clubs during the research period. Estimate of the principal.	Survey
Overall satisfaction	Grading of the principal (from 4 to 10) for the following factors: overall operation of the school, work of teachers, work of counselor, work of students, work of other staff, support of parents for the school, relations between school board and the school, and relations between the municipality and the school. Average over all factors.	Survey
Satisfaction concerning preparation of teachers	Indicates satisfaction of the principal (grade from 4 to 10) with the preparation of classroom work of teachers.	Survey
Getting things done	Indicates opinion of the principal of how easy it was to get things done with the teachers. 1=Easy; 2=Fairly easy; 3=Fairly difficult; 4=Difficult.	Survey

Appendix 4. *Changes in the group of 15% most and least efficient schools in models 2, 3, or 4 assuming variable returns to scale (Grouping 1). All schools. Percentage share in parentheses*

Group of inefficient schools					
	Model 1	Model 2	Model 3	Model 4	Model 5
Model 1	46 (100)				
Model 2	20 (44)	46 (100)			
Model 3	17 (37)	38 (83)	46 (100)		
Model 4	17 (37)	24 (52)	28 (61)	46 (100)	
Model 5	9 (20)	22 (48)	25 (54)	24 (52)	46 (100)
Group of efficient school					
	Model 1	Model 2	Model 3	Model	Model 5
Model 1	46 (100)				
Model 2	22 (48)	46 (100)			
Model 3	22 (48)	39 (85)	46 (100)		
Model 4	20 (44)	34 (74)	37 (80)	46 (100)	
Model 5	23 (50)	37 (80)	42 (91)	35 (76)	73 (100)

Appendix 5. Changes in the group of 15% most and least efficient schools in models 2, 3, or 4 assuming variable returns to scale (Grouping 1). Survey respondents. Percentage share in parentheses

Group of inefficient schools					
	Model 1	Model 2	Model 3	Model 4	Model 5
Model 1	15 (100)				
Model 2	6 (40)	17 (100)			
Model 3	5 (33)	14 (82)	18 (100)		
Model 4	6 (40)	9 (53)	10 (56)	22 (100)	
Model 5	2 (13)	8 (47)	8 (44)	9 (41)	11 (100)
Group of efficient schools					
	Model 1	Model 2	Model 3	Model 4	Model 5
Model 1	16 (100)				
Model 2	15 (94)	34 (100)			
Model 3	15 (94)	28 (82)	32 (100)		
Model 4	14 (88)	26 (77)	27 (84)	32 (100)	
Model 5	14 (88)	28 (82)	30 (94)	26 (94)	34 (100)

Appendix 6. Grouping of schools based on their efficiency scores in models 2, 3, 4 and 5

Grouping 1	Efficient: School is in the group of 15% most efficient ones in at least one of the models 2, 3 or 4 assuming variable returns to scale. Inefficient: School is in the group of 15% least efficient schools in at least one of the models 2, 3 or 4 assuming variable returns to scale. This grouping correspond the grouping used in probit models.
Grouping 2	Efficient: School is at the efficiency frontier (efficiency score=100) in model 5 assuming variable returns to scale. Because of the method, schools are also at the efficiency frontier in models 2, 3, and 4. Inefficient: School is not at the efficiency frontier in model 5 assuming variable returns to scale. This grouping corresponds the grouping used in tobit models.
Grouping 3	Efficient: School is in the group of 15% most efficient schools in all models 2, 3, 4 and 5 assuming variable returns to scale. Inefficient: School is not in the group of 15% most efficient schools in models 2, 3, 4 and 5 assuming variable returns to scale.
Grouping 4	Efficient: School is in the group of 15% most efficient schools in all models 2, 3, 4 and 5 assuming variable returns to scale. Inefficient: School is in the group of 15% least efficient schools in all models 2, 3, 4 and 5 assuming variable returns to scale.

Appendix 7. Summary statistics of the variables used in explaining efficiency differences

	N	Mean	Std. dev.	Min.	Max.
	RESOURCES AND ENVIRONMENTAL CHARACTERISTICS				
Student body composition					
Admission level	73	6.92	0.51	5.91	8.50
Heterogeneity	73	0.89	0.12	0.52	1.14
Female (%)	73	60.80	11.43	29.41	87.50
Education of teachers	73	0.83	0.11	0.53	1.00
Experience of teachers	73	3.39	0.70	1.22	4.80
Parental involvement					
Attendance in parent meetings	71	2.52	0.58	1.00	3.00
Facilities	73	1.47	0.64	0.33	3.00
Municipality type					
Rural	73	0.479	0.50	0.00	1.00
Dense	73	0.164	0.37	0.00	1.00
	ORGANIZATIONAL CHARACTERISTICS				
Organizational structure	62	0.37	0.49	0.00	1.00
Mission, goals and strategy	67	0.72	0.31	0.00	1.00
Consensus about the goals	70	1.54	0.70	1.00	4.00
Decision making					
Influence of teachers	61	0.62	0.31	0.00	1.00
Bottom-up planning	65	0.69	0.47	0.00	1.00
Principal proposal, educ. committee decisions in hiring new teachers	62	0.50	0.50	0.00	1.00
Educ. committee proposals & decisions in hiring new teachers	62	0.31	0.46	0.00	1.00
Staff development and monitoring					
Plans	66	0.42	0.50	0.00	1.00
Participation in training	70	3.10	1.12	1.00	5.00
Monitoring of teachers	69	0.81	1.02	0.00	3.00
Monitoring of students	70	5.14	1.65	1.00	12.00
Development of instruction					
Participation in projects	71	0.21	0.41	0.00	1.00
Joint teaching projects	68	1.47	1.41	0.00	5.00
Policy of parent-school relations	64	0.38	0.49	0.00	1.00
No. of active clubs	71	3.46	2.68	0.00	13.00
Overall satisfaction	68	8.41	0.56	7.13	9.50
Satisfaction with teachers preparation	67	8.63	0.61	7.00	10.00
Getting things done with teachers	69	1.61	0.52	1.00	3.00

Appendix 8. Mean differences in student and teacher characteristics, and school facilities between efficient and inefficient schools based on different groupings of schools. Number of schools in parentheses

	Grouping 1		Grouping 2		Grouping 3		Grouping 4	
	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient
Student body composition								
Admission level	7.02** (40)	6.79** (33)						
Heterogeneity	0.86*** (40)	0.93*** (33)	0.86*** (34)	0.92*** (39)	0.84*** (22)	0.92*** (51)	0.84** (22)	0.96** (6)
Females	63** (40)	58** (33)	63* (34)	59* (39)	64* (22)	60* (51)	64 (22)	61 (6)
Education of teachers	81.40* (40)	85.34* (33)						
Experience of teachers	3.40 (40)	3.39 (33)						
Physical facilities	1.64*** (40)	1.26*** (33)	1.67*** (34)	1.29*** (39)	1.65 (22)	1.39 (51)	1.64 (22)	1.39 (6)
Dense (%)	23 (40)	9 (33)	24 (34)	10 (39)	27 (22)	12 (51)	27 (22)	17 (6)
Rural (%)	45 (40)	52 (33)	44 (34)	51 (39)	41 (22)	51 (51)	41 (22)	17 (6)

*t-test significant at 10 %, **t-test significant at 5 %, ***t-test significant at 1 %

Appendix 9. Differences in organization structure, decision making and goals in efficient and inefficient schools based on different groupings. Number of schools in parentheses

	Grouping 1		Grouping 2		Grouping 3		Grouping 4	
	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient
Organizational structure								
Organization with no management or other assisting group (%)	56 (20)	73 (19)	61 (19)	65 (20)	57 (12)	66 (27)	57 (12)	80 (4)
Organization with management group or other similar body (%)	44 (16)	27 (7)	39 (12)	35 (11)	43 (9)	34 (14)	43 (9)	20 (1)
Decision making								
Top-down (%)	39 (14)	21 (6)	31 (10)	30 (10)	24 (5)	34 (15)	24 (5)	17 (1)
Bottom-up (%)	61 (22)	79 (23)	69 (22)	70 (23)	76 (16)	66 (29)	76 (16)	83 (5)
Teachers' influence in decision making								
Low (%)	9 (3)	4 (1)	7 (2)	6 (2)	11 (2)	5 (2)	11 (2)	17 (1)
Quite low (%)	27 (9)	36 (10)	29 (8)	33 (11)	28 (5)	33 (14)	28 (5)	33 (2)
Quite high (%)	24 (8)	39 (11)	25 (7)	36 (12)	17 (3)	37 (16)	17 (3)	50 (3)
High (%)	39 (13)	21 (6)	39 (11)	24 (8)	44 (8)	26 (11)	44 (8)	0 (0)
Autonomy of schools in hiring new teachers								
Principal makes proposals and principal/local school board decisions (%)	15 (5)	24 (7)	22 (4)	18 (8)	18 (5)	21 (7)	22 (4)	17 (1)
Principal makes proposals and local school board decisions (%)	61 (20)	38 (11)	44 (8)	52 (52)	54 (15)	47 (16)	44 (8)	33 (2)
Local school board makes proposals and decisions (%)	24 (8)	38 (11)	33 (6)	30 (13)	29 (8)	32 (11)	33 (6)	50 (3)
Consensus about the goals								
Very unanemous (%)	47 (18)	66 (21)	45 (15)	65 (24)	48 (10)	59 (29)	48 (10)	67 (4)
Unanemous (%)	42 (16)	28 (28)	42 (14)	30 (11)	38 (8)	35 (17)	38 (8)	33 (2)
Disagreeing (%)	8 (3)	6 (2)	9 (3)	5 (2)	10 (2)	6 (3)	10 (2)	0 (0)
Very disagreeing (%)	3 (1)	0 (0)	3 (1)	0 (0)	5 (1)	0 (0)	5 (1)	0 (0)
Mission, goals and strategy	0.67 (37)	0.78 (30)	0.70 (32)	0.73 (35)	0.65 (21)	0.75 (46)	0.65* (21)	0.94* (6)
Clearly defined goals (%)	87 (33)	91 (29)	88 (28)	89 (34)	86 (18)	90 (44)	86 (18)	90 (5)

*t-test significant at 10 %

Appendix 10. *Differences in staff development, teacher training and monitoring of students and teachers in efficient and inefficient schools based on different groupings. Number of schools in parentheses*

	Grouping 1		Grouping 2		Grouping 3		Grouping 4	
	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient
Staff development plans								
No plans (%)	54 (19)	61 (19)	52 (16)	63 (22)	50 (10)	61 (28)	50 (10)	67 (4)
Plans (%)	46 (16)	39 (12)	48 (15)	37 (13)	50 (10)	39 (18)	50 (10)	33 (2)
Participation in training								
Teachers did not participate in training (%)	**3** (1)	**0** (0)	3 (1)	0 (0)	5 (1)	0 (0)	5 (1)	0 (0)
Less than 25% of teachers participated (%)	**47** (18)	**25** (8)	41 (13)	34 (13)	50 (10)	32 (16)	50 (10)	33 (2)
25-50% of teachers participated (%)	**32** (12)	**22** (7)	34 (11)	21 (8)	25 (5)	28 (14)	25 (5)	33 (2)
50-75% of teachers participated (%)	**3** (1)	**38** (12)	9 (3)	26 (10)	5 (1)	24 (12)	5 (1)	33 (2)
Over 75% of teachers participated (%)	**16** (11)	**16** (5)	13 (4)	18 (7)	15 (3)	16 (8)	15 (3)	0 (0)
Monitoring of teachers								
No monitoring sessions (%)	51 (19)	50 (16)	47 (15)	54 (20)	57 (12)	48 (23)	57 (12)	50 (3)
Session once a year (%)	27 (10)	31 (10)	28 (9)	30 (11)	24 (5)	31 (15)	24 (5)	33 (2)
Session once a semester (%)	14 (5)	3 (1)	16 (5)	3 (1)	10 (2)	8 (4)	10 (2)	0 (0)
Session more often (%)	8 (3)	16 (5)	9 (3)	14 (5)	10 (2)	13 (6)	10 (2)	17 (1)
Participation in teaching projects								
Yes (%)	23 (9)	19 (6)	18 (6)	24 (9)	23 (5)	20 (10)	23 (5)	50 (3)
Ave. number of teaching projects across subjects	1.2 (37)	1.8 (31)	1.2 (32)	1.7 (36)	1.1 (20)	1.6 (48)	**1.1** (20)	**1.8** (6)

Bold = likelihood ratio chi2 significant at 10 % level

Appendix 11. Differences in cooperation, parent-school policies, and principals' satisfaction in efficient and inefficient schools based on different groupings. Number of schools in parentheses

	Grouping 1		Grouping 2		Grouping 3		Grouping 4	
	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient	Efficient	Inefficient
Cooperation with teachers								
Easy (%)	35 (13)	47 (15)	41 (13)	41 (15)	33 (7)	44 (21)	33 (7)	67 (4)
Fairly easy (%)	62 (23)	53 (17)	56 (18)	59 (22)	62 (13)	56 (27)	62 (13)	33 (2)
Fairly difficult (%)	3 (1)	0 (0)	3 (1)	0 (0)	5 (1)	0 (0)	5 (1)	0 (0)
Difficult (%)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Policy of parents-school relations								
Parent meetings (%)	**53** (19)	**75** (21)	***45*** (13)	*77* (27)	*42* (8)	*71* (32)	42 (8)	50 (3)
Meetings, bulletins, parents' association etc. (%)	**47** (17)	**25** (7)	***55*** (16)	***23*** (8)	*58* (11)	*29* (13)	58 (11)	50 (3)
Number of clubs in the school	4.0* (39)	2.8* (32)	3.7 (33)	3.3 (38)	3.2 (21)	3.6 (50)	3.2 (21)	1.9 (6)
Overall satisfaction (mean)	8.4 (37)	8.5 (31)	8.3(31)	8.5 (37)	8.4 (21)	8.4 (47)	8.4 (21)	8.4 (6)
Satisfaction of principals with the preparation of classroom work of teachers (mean)	8.5 (36)	8.8 (32)	8.5 (30)	8.7 (38)	8.4 (19)	8.7 (49)	8.4 (19)	8.7 (6)
Satisfaction of principals with the working together with the teachers (mean)	8.5 (37)	8.7 (30)	8.6 (31)	8.6 (36)	8.6 (21)	8.6 (46)	8.6 (21)	8.8 (6)

Bold = likelihood ratio chi2 significant at 10 % level; *italics* = likelihood ratio chi2 significant at 5 % level; ***bold and italics*** = likelihood ratio chi2 significant at 1 % level: * = t-test significant at 10 % level

Article III

EFFICIENCY OF FINNISH GENERAL UPPER SECONDARY SCHOOLS: AN APPLICATION OF STOCHASTIC FRONTIER ANALYSIS WITH PANEL DATA

Tanja Kirjavainen

Abstract

In this study the efficiency of Finnish general upper secondary schools is evaluated with stochastic frontier analysis. Different stochastic frontier models for panel data are used to estimate education production functions. Grades in the matriculation examination are used as output and explained with the comprehensive school grade point average, parental socioeconomic background, school resources, the length of studies and decentralization of test taking. Controls for the schools with specialized curriculum are also included. Heterogeneity across schools is allowed for by estimating both true random and true fixed effects models. The results show that the effect of teaching resources on examination results is negative when the heterogeneity across schools is taken into account. The length of studies and decentralization of test taking negatively affect student achievement. The inefficiency and the rankings of schools based on their inefficiency scores varies considerably depending on the type of stochastic frontier model applied. The lowest estimates for inefficiency were obtained with true random and true fixed effects models which separate time constant random or fixed effects from inefficiency.

Key words: Efficiency, Stochastic frontier analysis, Secondary schools (JEL I21)

Contents

1. Introduction

In industrialized countries the importance of high quality education as a source of well-being and economic growth is widely recognized. In most of these countries, education is publicly provided. Rising educational expenditures are another common feature for them and because of the aging population and increased global competition, governments are facing additional pressures in financing. Education, along with other sectors, is competing for resources to an increasing extent. In such a situation, enhancing the productivity and efficiency of the schooling system and individual schools is one way to maintain or even improve the provision of good quality education. The importance of the topic has also been widely recognized within the OECD and EU (see e.g. Gonand *et al*., 2007, Sutherland *et al*., 2007 and Wössmann and Schütz, 2006).

Measurement of efficiency in education is by no means straight forward. There are several features that make it controversial or at least complicated. Since education is often publicly funded, information on input and output prices is usually missing. In some cases, there is no clear consensus (at least amongst the practitioners) on what the 'real' outputs are and how they should be measured. The same applies also to schooling inputs. Even if the understanding is reached, the process involves several inputs and outputs. In addition, some of the inputs are not controllable by schooling institutions even though their influence on outputs is evident. Despite these difficulties, there is plenty of research on educational production functions.[1]

Several methodological approaches have been used to overcome problems in educational efficiency measurement. They all have their advantages and shortcomings. The early studies of the educational production function mostly used least-squares regression techniques, but since the 1980's the use of non-stochastic Data Envelopment Analysis (DEA) has become quite common. The wide application of DEA in this context is mostly due to its flexibility. DEA easily allows the use of several inputs and outputs and no information on

[1] See for reviews e.g. Hanushek (2003), Krueger (2003) and Worthington (2001).

prices is needed. In addition, as it is a linear optimization technique no assumption about the exact functional form is required.

As a linear programming method, DEA does not allow statistical interference. In addition, being a deterministic approach, it does not distinguish inefficiency from statistical noise. As a consequence, inefficiency may be overestimated. For policy purposes, both of these factors may cause problems and uncertainty. In stochastic frontier analysis (SFA), these shortcomings are avoided and for this reason it is an interesting alternative to DEA. In addition to inefficiency differences, information on the on estimated parameters, i.e. the effect of quantitative inputs (such as class size, teachers' salaries, education and experience as well as environmental variables) on outputs is obtained. The possibility to use panel data to control for unobserved heterogeneity further increases the attractivness of the method.

Studies using SFA are in a minority compared to applications of DEA[2] in the context of measuring the efficiency of schooling institutions. Most of these studies are using cross section data. Some studies compare the results of SFA and DEA (Sengupta and Sfeir, 1986 and Mizala *et al*., 2002). Others concentrate on inefficiency differences and testing the relationship between test scores and spending on instruction (Deller and Rudnicki, 1993) or teachers' merit pay (Cooper and Cohn, 1997). Heshmati and Kumbhakar (1997) used a more sophisticated model introduced by Battese and Coelli (1995) which assumes that inefficiency has a truncated-normal distribution and is dependent on some e.g. environmental factors.

However, few studies have made use of panel data. Barrow (1991) assessed the efficiency of local education authorities using both cross-section and panel data with stochastic and deterministic methods. Johnes and Johnes (2009) analysed the cost efficiency of British universities using SFA allowing for heterogeneity between universities with random parameters i.e. using true random effects model introduced by Greene (2005a, b).

---

[2] See for reviews e.g. Worthington (2001) and Johnes (2004).

In this study, the technical efficiency of Finnish general upper secondary schools is studied using different stochastic frontier models. A rich five-year set of panel data is used for estimating educational production functions. School level differences in scores in the matriculation examination are explained with students' prior achievement, family background characteristics, and school resources, school size and some environmental factors. Both teaching expenditures per student and student-teacher ratio (with a smaller sample) are tested as the measures of teaching resources.

Educational production function studies using register data, such as the present study, often suffer from two weaknesses. Variables measuring the school resources such as the student-teacher ratio may produce biased results because of non-random sorting of students between schools and within schools between teaching groups. Another source of bias is produced by the omitted variables (see e.g. Averett and McLennan, 2004). Best way to overcome the non-random sorting of students between schools is to use randomized experiments (see e.g. Krueger, 1999). Other methods include quasi-experimental designs and the instrumental variables approach (see e.g. Hoxby, 2000; Angrist and Lavy, 1997; Akerhielm, 1995; Goldhaber and Brewer, 1997). In the context of efficiency measurement, however, some of these techniques are inappropriate and others are difficult to carry out.

The new variants of stochastic frontier models are able to take into account the bias caused by omitted variables and the non-random assignment of students to schools. While traditional random and fixed effects stochastic frontier models interpret all unobserved random or fixed effects as inefficiency (Pitt and Lee, 1981 and Cornwell *et al*., 1990), new true random and true fixed effects models (Greene, 2005a, b) allow the decomposition of the inefficiency term into time-constant random or fixed effects and time varying inefficiency. Therefore, the unobserved time-constant school-specific factors are controlled for by the fixed effects and separated from inefficiency.

In Finnish general upper secondary schools students are not randomly assigned to different schools since students are admitted based on their comprehensive school grade point average (GPA). The number of general upper secondary schools varies from no school in small municipalities to tens of schools in larger cities. Competition for the places in best schools is intense, especially in larger cities. Students and parents choose the school based on school's curriculum (some of the schools have a specialized curriculum) and course offerings. Since larger schools are able to offer more extensive selection of courses they may attract better performing students. In this study, the comprehensive school GPA is used to control for the non-random assignment of students between schools. In addition, true fixed effects models, estimated with a five year panel capture the selection bias to the extent that it is time-constant and not captured by the GPA. Bias caused by omitted variables can also be controlled for by the true random effects models which are additionally tested in this study.

In 1996 two different reforms took place in Finnish general upper secondary schools. First, the grade system with fixed classes was removed providing students the possibilities to formulate their individual study plans and determine their own pace of learning. Students were allowed to complete their studies in 2 to 4 years. Before, students followed year classes and usually graduated in three years. As a consequence of the reform, the share of students using more than three years for their studies rapidly increased to over 15%. In this study, it is possible to test whether the length of studies affects performance in the matriculation examination.

In addition to the individual study plans, students were able to take the tests in the matriculation examination in up to three consecutive examination periods instead of an earlier one period. As a result, students have increasingly decentralized their test taking in recent years so that on average students participate in two test taking periods. The effect of decentralization of test taking on matriculation examination is also tested in this paper.

This study aims at, in other words, by using a fairly long panel data estimating efficiency of schools with education production functions.

It uses new variants of stochastic frontier models that allow the separation of time constant factors from inefficiency and tests the differences in results. These methods have not been previously applied to such a long and rich data set used in this study. In addition to school resource variables, it tests the effect of average length of studies and average participation in test taking periods on student achievement.

The results of this study demonstrate that in panel data models the selection of stochastic frontier model matters since average inefficiency varies depending on the model. It is lowest in true random and true fixed effects models. The rankings of schools based on inefficiency scores also vary substantially between different stochastic frontier models. True random and true fixed effects models provide more information about the structure of the inefficiency components, since they separate the time constant factors from inefficiency. The inefficiency term in these models captures smaller yearly fluctuations in input and output variables and the time-constant factor depicts some structural differences between the schools. The longer on average students stay in general upper secondary school and the more they decentralize their test taking, the worse they perform in the matriculation examination.

The paper continues as follows. In section 2 the Finnish general upper secondary school system is described. The models and estimation methods are discussed in section 3. In section 4, data and variables used in the study are described. Section 5 presents the estimation results and section 6 concludes.

2. Finnish general upper secondary schooling

Finnish general upper secondary schools provide a post-comprehensive education for students aged 16 to 19 years. The general upper secondary school certificate together with the matriculation examination certificate provides eligibility for university or tertiary level vocational education. Studies in general upper secondary schools can be completed in 2 to 4 years. Most of the students complete the general upper secondary schooling in three years. For approximately 17 percent of the students it takes more than three years to complete the studies.

There are no year classes in general upper secondary schools i.e. schools are non-graded. Studies have been divided into courses. The school year is usually divided into five or six periods. A schedule is devised for each period focusing on certain subjects. The students' progress and the composition of teaching groups thus depends on the students' choice of courses.

General upper secondary schooling concludes with a matriculation examination which is a compulsory nationwide test. The purpose of the examination is to determine whether students have assimilated the knowledge and skills required by the curriculum for the general upper secondary school. The examination is arranged in general upper secondary schools throughout Finland. The Matriculation Examination Board is responsible for administering the examination, for preparing the tests and for the final assessment of the answer papers[3]. The results of each individual test are normalized to be comparable between years.

Matriculation examinations are arranged in the autumn and spring during a two-week examination period. Students can take individual tests in up to three consecutive examination periods. The examination consists of at least four tests; one of these, the test in the candidate's mother tongue, is compulsory for all candidates. The candidate then chooses three other compulsory tests from among the following four tests: the test in the second domestic language, a foreign language test, the mathematics test, and the general studies test. As part of his or her examination, the candidate may additionally include one or more optional tests (foreign language test, mathematics test or general studies test).

Admission to general upper secondary schools is selective based on the grade point average (GPA) in comprehensive school. Application

[3] The preliminary assessment is carried out by the teachers in each school.

takes place through a national joint application procedure. Some 55% of graduates in each age cohort continue their studies at general upper secondary schools each year. The competition for places in the best schools is intense especially in large cities. Consequently, the average GPA of the students varies considerably across schools.

General upper secondary schools are mostly maintained by municipalities. In addition, there are some private schools, schools maintained by the joint organization of municipalities and some state owned schools. No municipal general upper secondary schools charge school fees. Municipalities cover school expenditures from their general revenue services which consist of local income tax, property tax and non-earmarked grants. Private schools get a state grant on a per student basis, and they also obtain funding from municipalities. Private schools may charge minor fees.

3. Models and estimation methods

Stochastic production frontier models were first introduced by Aigner *et al*. (1977) and Meeusen and van den Broeck (1977). In these models, variation unexplained by the input variables is not completely interpreted as technical inefficiency but statistical noise and technical inefficiency are separated. In addition, as parametric statistical models, they also provide information on the effect of inputs on the output.

In the following, it is assumed that schools I maximize their output production and use N inputs to produce a single output. In addition, cross-sectional data are assumed. A stochastic production frontier model can be written as

$$y_i = f(x_i;\beta)\exp\{v_i\}TE_i\,. \qquad (1)$$

where $\left[f(x_i;\beta)\exp\{v_i\}\right]$ is the stochastic production frontier, y_i is a scalar output produced by school *i*, *i=1,...,I*, x_i is a vector of *N* inputs by school *i*, $f(x_i;\beta)$ is the production frontier and β is a vector of technology parameters to be estimated. The stochastic frontier consists of two parts: a deterministic part $f(x_i;\beta)$, that is common to all schools and a school specific random part, $\exp\{v_i\}$, that captures the effect of random shocks on each school. TE_i is the output-oriented technical efficiency of school *i*. The equation (1) can be written as

$$TE_i = \frac{y_i}{f(x_i;\beta)\exp\{v_i\}}. \tag{2}$$

This equation defines technical efficiency as the ratio of the observed output to the maximum feasible output. Now y_i achieves its maximum feasible value of $\left[f(x_i;\beta)\exp\{v_i\}\right]$ if, and only if $TE_i = 1$. Otherwise, $TE_i < 1$ provides a measure of the shortfall of the observed output compared to the maximum feasible output in an environment characterized as $\exp\{v_i\}$, which is allowed to vary across schools.

Next, if we assume that $f(x_i;\beta)$ takes the usual log-linear Cobb-Douglas form, the basic cross-section model in (1) is written as

$$\ln y_i = \beta_0 + \sum_n \beta_n \ln x_{ni} + v_i - u_i\ , \tag{3}$$

where y_i depicts the output of school *i*, β are parameters to be estimated, x_{ni} are the explanatory variables, v_i is the idiosyncratic error term distributed independently of u_i and as iid $N(0, \sigma_v^2)$ and u_i is the nonnegative inefficiency term for school *i* distributed as iid $N^+(0, \sigma_u^2)$. The error terms in (3) form the composed error term $e_i = v_i - u_i$ which is asymmetric since $u_i \geq 0$.

The error term e_i is positively skewed if $u_i \geq 0$. If $u_i = 0$, then $e_i = v_i$, the error term is symmetric and there is no inefficiency in the data. In such a case equation (3) can be estimated with ordinary least squares without inefficiency. Before proceeding the skewness of inefficiency term must be tested. Coelli (1995) has introduced one test but there are also other alternatives (see e.g. Kumbhakar & Lovell, 2000).

If there is inefficiency in the error term, a distributional assumption for it has to be made. The assumption of a half-normal distribution is the most common. Other possibilities are exponential, truncated-normal and gamma distributions. Selection of the distribution is aided by certain statistical tests (see Kumbhakar & Lovell, 2000). Sample means of inefficiencies may be sensitive to the distribution assumption, whereas the rankings and the top and bottom deciles of efficiency scores are more likely to remain unaffected by it (Kumbhakar and Lovell, 2000 p. 90).

In this study, the one-sided likelihood ratio test introduced by Coelli (1995) is used to test the presence of a half-normal distribution for the inefficiency term. The test is based on the λ coefficient, which provides the relative contributions of u_i and v_i to e_i. The hypothesis tested is that $\lambda = \sigma_u/\sigma_v = 0$. As $\lambda \rightarrow 0$, the symmetric error component dominates the one-sided error component in the determination of e_i. In the case where $\lambda = 0$, there is no technical inefficiency and the model returns to the OLS. The assumption of truncated normal distribution, (u_i~iid, $N^+(\mu,\sigma_u^2)$) is tested using the assumption of H_0: $\mu = 0$. If $\mu = 0$ the density function returns to a normal density function.

The use of panel data slightly changes the equation in (3), since the time dimension is added.

$$\ln y_{it} = \beta_0 + \sum_n \beta_n \ln x_{nit} + v_{it} - u_{it} \,, \qquad (4)$$

where y_{it} is the output in period t, β are parameters to be estimated, x_{nit} are the explanatory variables in period t, v_{it} is the idiosyncratic error term for each period t and u_{it} are the inefficiency terms for each school in each period.

In a random effects models (Pitt & Lee, 1981), both between-schools and within-school variation is taken into account. The model assumes that the inefficiency term for each school is invariant through time and it is not dependent on other regressors. The estimated model becomes the following

$$\ln y_{it} = \beta_0 + \sum_n \beta_n \ln x_{nit} + v_{it} - u_i \,. \tag{5}$$

The weakness of model (5) is the assumption of inefficiency being constant through time as in longer panels inefficiency is more likely to vary through time. In true random and true fixed effects models, more heterogeneity is allowed for by dividing the school-specific inefficiency term into unmeasured heterogeneity that is constant through time and inefficiency that varies through time (Greene, 2005a, b). The true random effects model is as follows:

$$\ln y_{it} = \beta_i + \sum_n \beta_n \ln x_{nit} + v_{it} - u_{it} \,, \tag{6}$$

where β_i is a school-specific random term defined as $\beta_i = \beta + w_i$ and in which w_i is distributed as iid $N(0, \sigma_w^2)$, v_{it} is the school-specific error term and u_{it} is the inefficiency that varies through time.

If the familiar fixed effects model is used in efficiency measurement, the school-specific fixed effect that is constant through time is interpreted as inefficiency. The model is not in fact stochastic but "deterministic". The u_i of the most efficient unit equals zero. Thus, the model does not measure absolute inefficiency but the inefficiency of school *i* relative to the other schools in the sample. The advantage of this model is that it is distribution free. However, because it interprets the whole fixed effect as inefficiency, it most likely overestimates the inefficiency. The fixed effects model by Cornwell *et al*. (1990) can be written as

$$\begin{aligned} \ln y_{it} &= \beta_i + \sum_n \beta_n \ln x_{nit} + v_{it} \\ &= \max(\beta_i) + \sum_n \beta_n \ln x_{nit} + v_{it} + \left[\beta_i - \max(\beta_i)\right] \\ &= \beta_0 + \sum_n \beta_n \ln x_{nit} + v_{it} - u_i \,, \end{aligned} \tag{7}$$

where $u_i = \max(\beta_i) - \beta_i > 0$.

In a true fixed effects model (Greene, 2005a, b) the fixed effects and inefficiency effects are separated. The model allows inefficiency to vary through time and takes into account the omitted variable bias. The model is written as

$$\ln y_{it} = \beta_i + \sum_n \beta_n \ln x_{nit} + v_{it} - u_{it}. \qquad (8)$$

where β_i is a school-specific fixed term, v_{it} is the school-specific error term and u_{it} is the inefficiency that varies through time. In this model, these effects can be correlated with the included variables.

Heteroskedasticity is potentially a problem in stochastic frontier models (see e.g. Kumbhakar and Lovell, 2000). In cross-section models, it can appear in either of the error terms, and can affect interferences between the production technology parameters and the error components. Thus, it can also have effects on the inefficiency estimates. If heteroskedasticity appears in the inefficiency term, the problem is more severe, since both the estimates of production technology and inefficiency are biased. If v_i is heteroskedastic, only the inefficiency estimates are affected. If both error terms are heteroskedastic the effect is not clear, since the unmodelled heteroskedasticity causes biases in opposite directions. In such a case, the overall bias can be small.

In random and fixed effects panel data models, only v_{it} can be heteroskedastic, and even if it is ignored it does not cause serious problems with the results. In the case of time varying inefficiency, heteroskedasticity may only appear in v_{it} in the random effects model. In such a model, a time effect for the u_{it} is assumed and there are several alternatives to model it (see e.g. Cornwell *et al*., 1990; Lee and Schmidt, 1993; Kumbhakar, 1990 and Battese and Coelli, 1992).

There are also other alternatives to model the heterogeneity in the inefficiency term. Examples, such as those presented by Battese and Coelli (1995) and Coelli *et al*. (1999) express u_{it} as a parametric function of some explanatory variables z_{it} that characterize, for instance, the environment or the organization of the producer. In this study, no such variables were available. Thus, throughout the paper it is assumed that all explanatory variables directly depend on the dependent variable rather than on inefficiency.

4. Data and descriptive statistics

This paper examines the efficiency of Finnish general upper secondary schools using a nationally representative school-level panel data from the years 2000–2004. Data includes most of the Finnish general upper secondary schools. Some language schools that do not participate in the Finnish Matriculation Examination were excluded from the data. Schools providing general upper secondary education for adults were also omitted since their curriculum differs from that for youths. Some schools were excluded because of data problems. The data are unbalanced consisting of 436 schools. The number of schools varies from 424 to 427 depending on the year.[4]

The data were compiled from several different official registers. Matriculation examination grades and information on test taking as well as the graduates' mother tongue and sex were obtained from the Matriculation Examination register. Only information on students completing their studies and matriculating was used. Students retaking individual tests in later years in order to improve their grades after completing their studies were excluded from the data. National joint application register provided the information on the grade point average (GPA) of comprehensive school reports. Information on students' socio-economic status and length of studies, as well as size, location and type of owner of the school was obtained from Statistics Finland. Information on expenditures was obtained from the VALOS register maintained by the Finnish National Board of Education.

School output is measured by the *scores in compulsory tests in the matriculation examination*. As mentioned earlier, there are four

[4] There are 413 schools in the data which are observed every year. For 10 schools, there are observations in four consecutive years, for two schools three consecutive years and for 11 schools two consecutive years.

compulsory tests in the matriculation examination. The grades in each test range from improbatur (failed) to laudatur (excellent) and they are converted to scores using a scale from 0 to 7.[5] The maximum scores in compulsory subjects is therefore 28. In addition to compulsory subjects, students can take optional tests, and they usually take one or two of them.[6] As an alternative output measure, *scores in all tests in the matriculation examination* were also used. This measure includes grades in both compulsory and optional tests.

The average score in the compulsory tests was 16.7 but the variation across schools was quite large. In the top schools the average score was 23.4 whereas in the bottom schools the average score was 10.7. The variation in scores in all tests was even larger ranging from 11.3 to 34.7. Students' prior achievement is controlled for with the comprehensive school grade point average (GPA). These grades are awarded by teachers and provide the best information available, since students are not tested on a national level at the end of comprehensive schooling. Information on the GPA was linked at the student level before averaging it to the school level.

The socio-economic status of students is measured with three variables: the e*ducational level of their parents*, *the proportion of white-collar workers*, and *the proportion of single parents*. This information was also linked to matriculation examination results at the student level before averaging to the school level. The educational level of the parents is provided as an index by Statistics Finland and is based on the number of years of schooling. Other controls for the students of each school include the *proportion of female* and *Swedish-speaking students*. The latter is included because the matriculation examination is quite language oriented. Swedish-speaking students may have some advantage because of their language background.

[5] The grades are converted into scores so that improbatur=0, approbatur=2, lubenter approbatur=3, cum laude approbatur=4, magna cum laude approbatur=5, eximia cum laude approbatur=6 and laudatur=7.

[6] The average number of tests taken was 5.1 during the research period.

Heterogeneity of the student body may affect instruction and instructional methods. Some teachers often claim that it's difficult to adjust teaching methods if the skills of the students are very heterogeneous. This variable tests whether this has some effect on student performance. Heterogeneity of the student body is measured with the within-school standard deviation of scores in compulsory subjects in the matriculation examination.

School resources are measured with two variables: *teaching expenditures per student* and *other current expenditures per student*. The effect of *average student teacher ratio* is also tested. Teaching expenditures consist of teachers' and principals' salaries, teaching materials and other costs that can be directly attributed to teaching. They account for some 75% of the total expenditures. Other expenditures consist of the costs of meals, health care and counselling, administration and rents (pure or calculatory) for the school properties.

The expenditure information is averaged over the three years that students usually enrol for general upper secondary school. Since most of the students complete their studies in the spring term the average is taken over the three previous years. It means that e.g. the expenditures for the year 2004 are an average over the years of 2001 to 2003. The costs are deflated to prices for the year 2003 using a price index for public spending in education. The expenditures have risen partly because there was a reform in the pension insurance system that was implemented gradually since 1998. In order to take the effect of the reform into account, the expenditures are deflated with chained deflator. The base year in the first deflator was 1995 and in the latter one 2000. Regional differences in expenditure information (mainly salaries and rents on properties) are not taken into account. There are quite considerable differences in teaching expenditures per pupil across schools. For students matriculating in 2000 the average teaching expenditures were some 3 300 euros. However, they varied between 1 800 and 13 000 euros. For students completing their studies in 2004 the average teaching expenditures per student were some

3 700 euros. Differences in expenditures have somewhat diminished, since they varied between 2 200 and 9 200 euros in 2004.

The VALOS register provides the expenditure information at the level of the service provider. If the service provider (usually the municipality) maintains more than one general upper secondary school, the information cannot be attributed to individual schools. However, in most cases the service provider maintains only one general upper secondary school. Larger cities and some bilingual areas are an exception. The data include school-level expenditure information for 254 to 259 schools (depending on the year) and for 168 to 172 schools a municipality level average is used.[7]

The average student-teacher ratio is only obtained for the municipal general upper secondary schools. The models including this variable are therefore estimated with a smaller data containing 343–369 schools depending on the year. This information is obtained at the school level. Thus it is a more accurate measure than the teaching expenditures per student. The number of teachers and principals are in full time equivalents. Due to absence of data the variable is averaged over the two previous years instead of three years.

[7] To obtain school-level expenditure information the Government Institute for Economic Research (VATT) also carried out a survey of those municipalities that had more than one general upper secondary school. There were 53–54 such municipalities in the data. In the survey, municipalities were asked to report their school-level expenditures and teaching hours as they are reported in VALOS register. This information was obtained from 23 municipalities maintaining 70 general upper secondary schools. The three biggest cities namely Helsinki, Espoo and Vantaa did not submit the information. Based on information reported by the municipalities, the percentage share of the total teaching and other expenditures were calculated for each school. Using this percentage the expenditures in the VALOS register were divided between schools. For those schools that did not report school-level information, the expenditures were divided using the relative proportion of students from the total number of students in the municipality. Models were also estimated using this expenditure information, but the parameter and inefficiency estimates were unaffected. The results did not change, either, when the models were estimated using the expenditure information from VALOS register and leaving out the schools in Helsinki, Espoo and Vantaa. For this reason, only the expenditure information from the VALOS register is used in this paper with a municipal-level average for schools in municipalities having more than one general upper secondary school.

The general upper secondary schooling can be completed in two to four years, depending on the study plan of the student. For some 80% of students the studies take three years. A greater *length of studies* may require more resources, but it is not clear how it affects student achievement, since students may either study more or they may use their time for other activities, such as working.[8] In 2004 the average length of studies was 3.1 years and it varied between 3 to 3.8 years. The length of studies is a school-level average and it is calculated from student-level data as the difference between the semester of starting and completing the studies.

The tests in the matriculation examination can be taken in up to three consecutive examination periods. On average, students participate in 1.8 examination periods but this ratio has steadily increased. The advantage of taking tests in more than one examination period is that students can concentrate more fully for each test. However, whether this is really the case and whether the decentralization of test taking affects the results has not been tested. The average *decentralization of tests* is calculated from individual-level information on the number of examination periods a graduate has participated in.

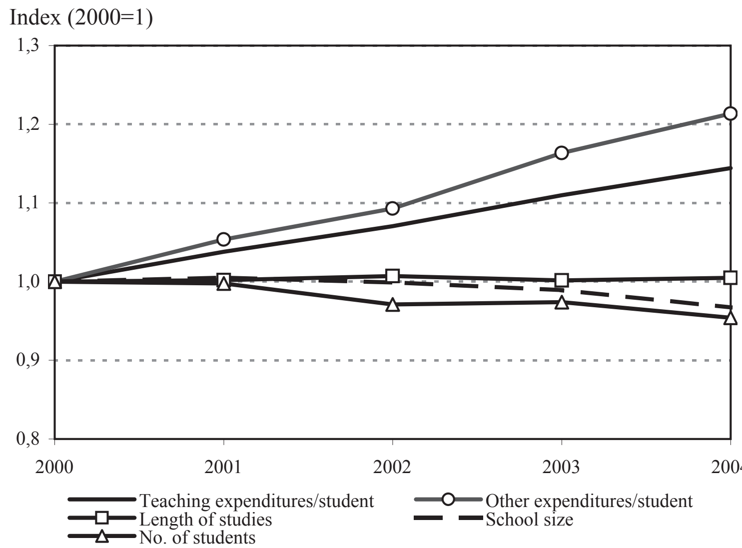
Some 13% of the general upper secondary schools have a specialized curriculum. In these schools, students can specialize in music, arts, sports, languages, natural sciences or mathematics. It is evident that specialization may have some effect on student performance. Students specializing, for instance, in languages may have some advantage over other students because of the relatively large emphasis on languages in the matriculation examination. Separate dummies are used for schools having a curriculum specialized in *sports*, *languages*, *mathematics and science*, and *music and arts*. However, it must be noted that most of the schools with a specialized curriculum also have the general track. Unfortunately, it was not possible to separate the proportion of students following the general track from those following the specialized track.

[8] Some 20% of the students in general upper secondary schooling worked during their studies in 2004. Among the younger students the share is around 10% and it increases with the age. Unfortunately, it was not possible to control the working of students in this study.

The *average size of general upper secondary schools* is quite small, being some 260 students. The smallest schools have only some 30 to 40 students and largest about 850 students. Most of the schools are maintained by the municipalities while only some two per cent of the schools are *private* and one per cent are *state-maintained* schools. The latter serve as training schools for new teachers. Concerning the *location of the school*, about half of the schools are located in *urban* areas and one third in *rural* areas. The summary statistics are reported in Appendix 1.

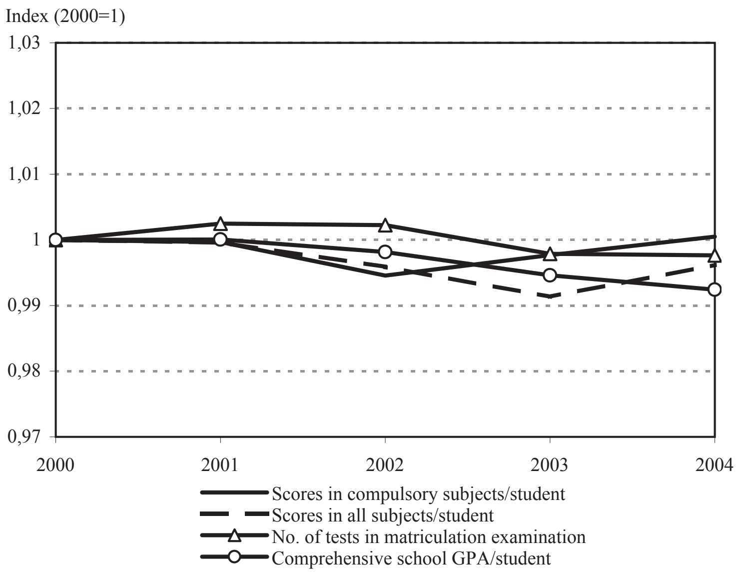
There are clear trends in some of the input variables (Figure 1). Expenditures in general upper secondary schooling have risen quite fast during the period. Teaching expenditures per student increased 14% and other expenditures per student 21% in five years. At the same time average school size has dropped some five percent. This is mainly because the number of students has decreased especially in rural areas.

Figure 1. Trends in some input variables in 2000-2004



There is no trend in the matriculation examination scores per student (see Figure 2). This is because they are standardized to yield the same distribution every year. Even if the scores are standardized at the national level, the performance of single schools can vary much from year to year. The number of tests taken in matriculation examination has remained stable. The same applies to the comprehensive school GPA.

Figure 2. Trends in matriculation examination scores and comprehensive school GPA in 2000-2004



5. Results

Five stochastic frontier models were estimated using an unbalanced panel data.[9] They were a pooled stochastic frontier model, a random effects (RE) model, a fixed effects (FE) model and true random effects (TRE) and true fixed effects (TFE) models (Greene, 2005a, b).[10] In addition, two alternative specifications were used. In Specification A, teaching resources were measured with *the teaching expenditures per student*. In Specification B this variable was replaced with *the student-teacher ratio*. The latter specification was estimated with a smaller data.

All the models and specifications were estimated assuming a half-normal distribution for the inefficiency term. For the pooled panel data model and the RE model the truncated normal assumption was also tested, but the μ term turned out to be statistically insignificant. In most cases an exponential distributional assumption did not converge. There was heteroskedasticity related to school size in the idiosyncratic error term v_{it} for the pooled panel data model and random effects models. The results are therefore heteroskedasticity-corrected for these models. The correction had only minor effects on the results.

Concerning the choice between random and fixed effects models, the Hausman specification test was performed. The results of the test supported fixed effects models. Since the panel is fairly short (five years) and some of the explanatory variables of interest remain constant through time, the results of random effects models are also presented and discussed.

---

[9] The production function was first estimated with cross section data separately for each year. A Chow-test (H_0: $\beta_{i,2000} = \beta_{i,2001} = \beta_{i,2002} = \beta_{i,2003} = \beta_{i,2004}$, where $i=1,\dots,n$ depicts explanatory variables) was performed to test whether the parameter estimates differed statistically significantly across years. According to results, the H_0-hypothesis could not be rejected, supporting the pooling of the data. In the following only the results of panel data models are presented. The results of cross section estimations are available upon request.

[10] Models were estimated with Nlogit 4.0/Limdep 10.0.

Since the grades in each individual test in matriculation examination are normalized each year, the output measure is not genuinely increasing. The scores of each school are rather fluctuating around their yearly overall mean. To test this kind of bias all the explanatory variables were centered on their yearly overall mean and models were estimated with these variables. The centering did not affect the results. Therefore, they are not reported here but are available upon request.

The stochastic frontier model is appropriate for the description of the production technology since the lambda coefficient is statistically significant in all the models. There is, in other words, inefficiency that is captured with the term u_{it}. The variables explain a high proportion of the variation in matriculation examination results, since the R-squared in normal OLS is 0.70[11]. In most of the models, the size of the parameter estimates is quite similar. A similar pattern concerning the size of the coefficients has also been reported by Greene (2005b), among others. There are, however, some exceptions. The size of the effect of *parental education*, *heterogeneity of the students* and *the average length of the studies* can double in some models.

Most of the explanatory variables in these models have the expected sign. *Comprehensive school GPA* positively affects matriculation examination scores. The effect is quite large and a one-tenth growth in a school's GPA gives 0.3 increase in scores. As for the students' socio-economic status, *parental educational level* and the *proportion of white collar workers* increase achievement in the matriculation examination whereas the *proportion of single parents* decreases it. Their effect is smaller than the effect of GPA. Schools with a higher proportion of female and Swedish-speaking students perform better in the matriculation examination. The *heterogeneity of the student body* negatively affects student achievement.

Both *teaching expenditures* and *other expenditures per student* have statistically significant coefficients in the TRE model. Teaching expenditures are also statistically significant in the TFE model. It

[11] The results can be obtained upon request.

appears that as soon as the unmeasured heterogeneity across schools is captured in the model, teaching expenditures become statistically significant.[12] Interestingly, the coefficient is negative in all models, indicating that schools with a higher level of teaching resources perform worse. The size of the effect is, however, quite small.

The average length of the studies negatively affects the matriculation examination results in all models and the coefficient is statistically significant in TRE and FE models. On average, a longer duration of studies does not appear to contribute to increased performance in the matriculation examination. This is an interesting result, since the length of the studies clearly increased after the year classes were removed from general upper secondary schools and students were able to complete their studies in 2 to 4 years. It is also contrary to normal argumentation according to which increased time would enhance the results.

As mentioned earlier, in matriculation examination students can take tests in up to three consecutive test periods. According to results, the more students decentralize their test taking into separate test periods, the worse they perform. This is a robust result. One reason behind the result might be that students more easily take tests to “try their luck” and do not prepare themselves for the test as carefully as would be the case when there is no possibility to retake it. The incentive to upgrade the test result in a later examination period is perhaps small.

There are no systematic differences in achievement between students in schools with a specialized curriculum compared to those in non-specialized general upper secondary schools. In some of the models, students in schools having specialized track in *mathematics and sciences* score two percentage points lower than those in non-specialized schools. In TRE model, students in general upper secondary schools having specialized track in *sports* score somewhat

[12] In TRE and TFE models it was also tested whether the sign of the resource variables was affected by some variables that may affect the size of the teaching and other expenditures, namely the school size and location. If these variables were omitted from the models the expenditure variables were statistically insignificant. Hence, these variables provide some additional information and their exclusion from the model bias the results.

higher. The same applies to students in schools having specialized track in *languages*. However, the results in schools specializing in *music and arts* do not differ from those in general schools.

In pooled panel and TRE models, student achievement in *private schools* is lower than the achievement in schools maintained by municipalities. However, *state maintained schools* do not differ from municipal schools. Student achievement does not differ statistically significantly between schools in *rural* and *urban areas*. Schools in *densely populated areas* score somewhat higher relative to schools in urban areas in pooled panel data and TRE models.

School size has a small positive effect on student achievement, and the coefficient is statistically significant in pooled panel data, RE and TFE models. The size of the coefficient is, however, quite small. The nonlinearities in school size were also tested by adding a second order term, but it turned out to be statistically insignificant.

Table 1. The results of stochastic frontier models for panel data from 2000-2004. Dependent variable: average score in compulsory tests in the matriculation examination. Specification A

	Pooled panel data model	Random effects model	True random effects model	Fixed effects model	True fixed effects model
GPA	1.622 (39.37)**	1.596 (38.86)**	1.637 (57.18)**	1.611 (17.99)**	1.578 (39.74)**
Parental education	0.141 (9.20)**	0.071 (5.12)**	0.063 (6.78)**	0.039 (2.18)*	0.135 (12.19)**
% White collar workers	0.001 (2.57)*	0.001 (2.94)**	0.001 (6.20)**	0.001 (3.01)**	0.001 (3.46)**
% Single parents	-0.001 (-4.26)**	-0.0002 (-1.09)	-0.0002 (-1.10)	-0.000 (-0.01)	-0.001 (-6.13)**
Teaching expenditures/student	-0.016 (-1.55)	-0.019 (-1. 66)	-0.028 (-4.38)**	-0.000 (-0.01)	-0.021 (-3.68)**
Other expenditures/student	-0.005 (-0.88)	0.010 (1.34)	0.014 (3.73)**	0.029 (2.48)*	0.004 (1.09)
SD of matriculation exam score	-0.100 (-9.66)**	-0.080 (-9.19)**	-0.069 (-10.68)**	-0.055 (-5.36)**	-0.109 (-13.33)**
% Female	0.001 (8.11)**	0.001 (4.94)**	0.001 (6.93)**	0.001 (3.75)**	0.001 (10.07)**
% Swedish speaking	0.001 (9.20)**	0.001 (7.26)**	0.001 (16.82)**	0.0004 (0.64)	0.001 (15.83)**
Mean length of studies	-0.060 (-1.58)	-0.041 (-1.25)	-0.108 (-4.03)**	-0.087 (-1.60)	-0.111 (-3.72)**
Mean participation in exam periods	-0.102 (-11.34)**	-0.092 (-10.35)**	-0.100 (-17.89)**	-0.104 (-8.17)**	-0.108 (-17.96)**
School size	-0.011 (-2.85)**	-0.013 (-3.02)**	-0.007 (-2.51)*	-0.019 (-1.13)	-0.016 (-6.64)**
Languages	0.019 (1.24)	0.009 (1.11)	0.017 (2.01)*	-0.017 (-0.53)	0.016 (1.74)
Mathematics and science	-0.018 (-2.88)**	-0.020 (-1.78)	-0.020 (-3.51)**		
Music and arts	-0.004 (-0.55)	-0.001 (-0.10)	0.008 (1.60)	0.006 (0.25)	-0.004 (-0.97)
Sports	0.005 (0.07)	0.012 (1.43)	0.014 (2.35)*		
Private	-0.003 (-0.65)	0.012 (1.45)	-0.012 (-3.12)**		
State owned	0.017 (1.30)	0.015 (1.06)	0.018 (2.00)*		

Table 1. continues

Densely populated area	0.013	-0.003	0.009		
	(3.55)**	(-0.43)	(3.39)**		
Rural area	0.008	-0.013	0.003		
	(1.46)	(-2.17)*	(1.02)		
Year	0.003	0.004	0.004	0.003	0.004
	(3.29)**	(4.77)**	(6.38)**	(2.12)*	(77.89)**
Constant	-7.567	-7.960	-8.413	-6.261	
	(-3.86)**	(-5.36)**	(-6.99)**	(-2.50)*	
Lambda		0.207	1.008		1.349
		(16.98)**	(8.14)**		(19.41)**
Sigma(v)	0.052	0.377	0.034		0.059
Sigma(u)	0.046	0.078	0.035		0.080
Het: Constant	-1.627				
	(-4.46)**				
Het: School size	-0.832	-0.870			
	(-10.82)**	(-21.36)**			
Log-L	3117.66	3542.45	-3417.02		3301.12
No. of obs.	2133	2133	2133	2133	2133
No. of schools	436	436	436	436	436

Note. T-values in parentheses, robust t-values in the fixed effects model. *Significant at 5%, **significant at 1%. All explanatory variables are in a logarithmic form except ratios. Schools maintained by the municipalities are the reference group for private and state own schools. Schools with general curriculum are the reference group for schools with specialized curriculum. Schools in urban areas are the reference group for schools in densely populated or rural areas.

The results are very similar when teaching expenditures per student are replaced with *student-teacher ratio* and the model is estimated with a smaller data consisting only of municipal general upper secondary schools (Specification B, see Table 2).[13] The change has only minor influence on the coefficients of other variables. Interestingly, the coefficient of student-teacher ratio still has the unexpected sign in all the models and it is statistically significant in pooled panel data, RE and TFE models. In other words, the larger the group size the better is the results in the matriculation examination. As discussed earlier, it is possible that the selection bias is only partly controlled in this model.

[13] Models in Table 1 were also estimated with the smaller data (only municipal general upper secondary schools) to test the robustness of results. The results were very similar. Hence the size of the data did not affect the results.

Table 2. The results of stochastic frontier models for panel data from 2000–2004. Dependent variable: average score in compulsory tests in the matriculation examination. Specification B

	Pooled panel data model	Random effects model	True random effects model	Fixed effects model	True fixed effects model
GPA	1.673	1.570	1.614	1.570	1.611
	(36.87)**	(35.02)**	(53.24)**	(15.91)**	(36.15)**
Parental education	0.126	0.069	0.057	0.040	0.105
	(7.27)**	(4.34)**	(5.58)**	(2.09)*	(8.70)**
% White collar workers	0.001	0.001	0.001	0.001	0.001
	(2.53)*	(2.34)*	(5.50)**	(2.65)**	(3.57)**
% Single parents	-0.001	-0.0001	0.000	0.000	-0.001
	(-3.00)**	(-0.62)	(0.02)	(0.76)	(-3.37)**
Student-teacher ratio	0.017	0.030	0.010	0.018	0.019
	(2.50)*	(4.10)**	(2.03)*	(1.16)	(4.30)**
Other expenditures/student	-0.003	0.014	0.010	0.033	0.000
	(-0.47)	(1.64)	(2.50)*	(2.51)*	(0.10)
SD of matriculation exam score	-0.091	-0.073	-0.060	-0.044	-0.105
	(-7.82)**	(-7.22)**	(-7.83)**	(-3.95)**	(-11.68)**
% Female	0.001	0.001	0.001	0.001	0.001
	(7.06)**	(4.65)**	(6.71)**	(3.82)**	(8.61)**
% Swedish speaking	0.001	0.001	0.001	-0.001	0.001
	(8.57)**	(7.04)**	(14.37)**	(-1.09)	(14.18)**
Mean length of studies	-0.059	-0.041	-0.139	-0.132	-0.094
	(-1.32)	(-1.15)	(-4.75)**	(-2.28)*	(-2.83)**
Mean participation in exam periods	-0.104	-0.091	-0.102	-0.106	-0.111
	(-10.60)**	(-9.43)**	(-16.97)**	(-7.86)**	(-16.78)**
School size	-0.006	-0.013	-0.004	-0.041	-0.010
	(-1.52)	(-2.44)*	(-1.29)	(-2.17)*	(-4.24)**
Languages	0.015	-0.033	-0.030	-0.063	0.012
	(0.51)	(-2.22)*	(-1.74)	(-4.05)**	(0.77)
Mathematics and science	-0.023	-0.012	-0.007		
	(-3.34)**	(-0.91)	(-1.17)		
Music and arts	-0.005	-0.0004	0.010		
	(-0.62)	(-0.04)	(1.89)		
Sports	0.003	0.018	0.015		
	(0.44)	(1.88)	(2.38)*		
Densely populated area	0.011	-0.004	0.006		
	(2.75)**	(-0.61)	(2.02)*		
Rural area	0.006	-0.014	-0.002		
	(1.15)	(-2.01)*	(-0.46)		

Table 2. continues

Year	0.003 (3.03)**	0.003 (4.25)**	0.004 (5.34)**	0.003 (2.39)*	0.004 (77.98)**
Constant	-7.792 (-3.65)**	-7.086 (-4.85)**	-7.810 (-5.86)**	-5.856 (-2.59)**	
Lambda		0.201 (15.35)**	1.181 (8.40)**		1.381 (18.12)**
Sigma(v)	0.055	0.393	0.032		0.059
Sigma(u)	0.033	0.079	0.038		0.081
Het: Constant	-1.752 (-4.23)**				
Het: School size	-0.776 (-8.54)**	-0. 890 (-19.37)**			
Log-L	2627.41	2997.19	-2898.39		2794.03
No. of obs.	1798	1798	1798	1798	1798
No. of schools	376	376	376	376	376

Note. T-values in parentheses, robust t-values in the fixed effects model. *Significant at 5%, **significant at 1%. All explanatory variables are in a logarithmic form except ratios. Schools maintained by the municipalities are the reference group for private and state own schools. Schools with general curriculum are the reference group for schools with specialized curriculum. Schools in urban areas are the reference group for schools in densely populated or rural areas.

All the models were also estimated using scores all tests in the matriculation examination as the dependent variable. These results are reported in Appendix 3 and Appendix 4. They were mostly similar with those presented above. The lambda coefficient in TRE model was statistically insignificant indicating that there is no inefficiency in the model and traditional RE model would be appropriate to describe the production. This result applied to Specification A. In Specification B in which the *teaching expenditures per student* was replaced with the *student-teacher ratio*, lambda coefficient was statistically significant indicating that there was some inefficiency.

The size of the effect of comprehensive school GPA was even larger in these models. Concerning the resource variables, *teaching expenditures per student* was statistically insignificant in all models. *Other expenditures per student* had positive and statistically significant coefficients in TRE, FE and TFE models. *Student-teacher ratio* had a positive and statistically significant coefficient only in RE model.

5.1 Inefficiency differences

There are clear differences in the inefficiency of schools between the separate stochastic frontier models (see Table 3, Figure 3 and Appendix 5).[14] The traditional FE model, by labelling all school-specific fixed effects as inefficiency, produces clearly the highest average inefficiency of some 15% in Specification A. As a consequence of taking into account school-specific fixed effects and separating them from inefficiency, average inefficiency decreases to 6%. Therefore, interpreting the whole school-specific fixed effect inefficiency most likely overestimates its magnitude. The same pattern applies to RE and TRE models, although the average inefficiency is lower in both cases compared to fixed effects models. The clear difference in the average inefficiency between random and fixed effects models highlights the importance of the choice of an appropriate model.

The variation in inefficiency scores across schools is also highest in FE model, whereas the TRE model produces the lowest variation each year. The results of pooled panel data and the TFE model are also very close to that of TRE model.

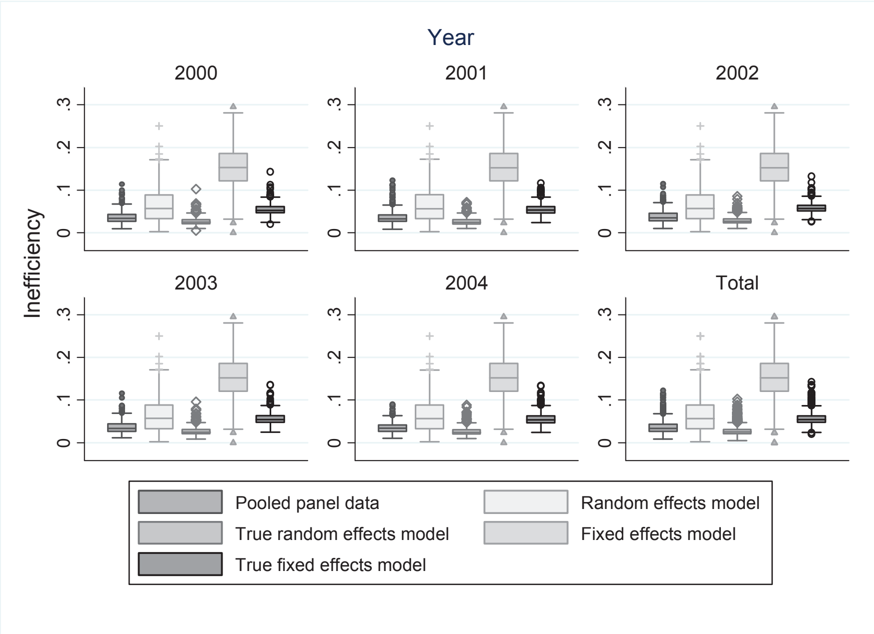
[14] The inefficiency score (1-TE_i) varies between 0 and 1 and the larger the figure the more inefficient the school.

Table 3. Average inefficiency, standard deviation, minimum and maximum in panel data models for the years 2000-2004. Scores in the compulsory tests in the matriculation examination as the dependent variable

	Pooled panel	Random effects	True random effects	Fixed effects	True fixed effects
Specification A					
Mean	0.037	0.064	0.027	0.154	0.056
Standard deviation	0.015	0.041	0.010	0.046	0.014
Min.	0.008	0.002	0.005	0.000	0.020
Max.	0.123	0.250	0.103	0.295	0.143
Specification B					
Mean	0.026	0.065	0.030	0.179	0.057
Standard deviation	0.008	0.042	0.013	0.063	0.014
Min.	0.009	0.002	0.006	0.000	0.020
Max.	0.068	0.209	0.121	0.341	0.146

A change in the output variable from scores in compulsory tests to scores in all tests in the matriculation examination had some effect on inefficiency (see Appendix 6). Especially in random and fixed effects models the average inefficiency was higher. In the true random effects model, the lambda coefficient showed that there was no inefficiency. The replacement of teaching expenditures per student with student-teacher ratio and the use of smaller sample (Specification B) had only minor effect on average inefficiency (see Appendix 7 and Appendix 8).

Figure 3. Average technical inefficiency in panel data models for the years 2000-2004. Scores in the compulsory tests in the matriculation examination as the dependent variable. Specification A



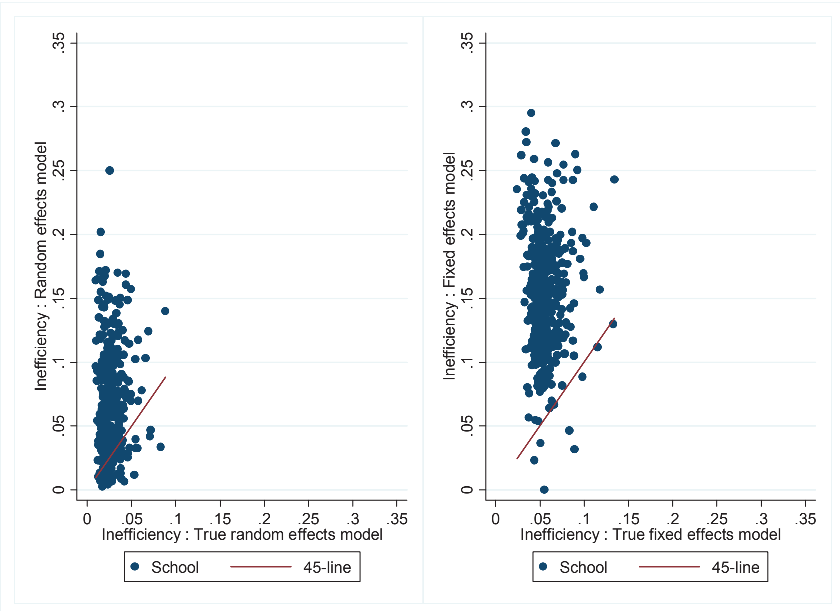
There are clear differences in the rankings based on inefficiency scores between the different stochastic frontier models (see Table 4). The rankings of RE and FE models are quite similar as are those of TRE and TFE models. However, there is practically no relation between the rankings of RE and TRE models or FE and true TFE models. RE and TFE models also produce quite different rankings. The replacement of scores in compulsory tests with scores in all tests and teaching expenditures per student with the student-teacher ratio did not cause any real differences in Spearman rank correlation coefficients. Thus, the results are not reported here.

Table 4. Spearman rank correlation coefficients between different stochastic frontier models. Scores in compulsory tests in the matriculation examination as the dependent variable. Specification A

	Pooled panel data model	Random effects model	True random effects model	Fixed effects model	True fixed effects model
Pooled panel data model	1				
Random effects model	0.7465	1			
True random effects model	0.6804	0.1838	1		
Fixed effects model	0.6892	0.9031	0.1775	1	
True fixed effects model	0.5520	0.0069	0.9452	0.0163	1

The difference in inefficiency scores between the RE and TRE models and the FE and TFE models is also depicted in Figure 4. This demonstrates the same pattern as the Spearman correlation coefficients. There is no systematic relationship between the RE and TRE models or FE and TFE models. RE and FE models produce high inefficiency estimates by interpreting all random and fixed effects as inefficiency. According to TRE and TFE models, these effects are only partly due to inefficiency.

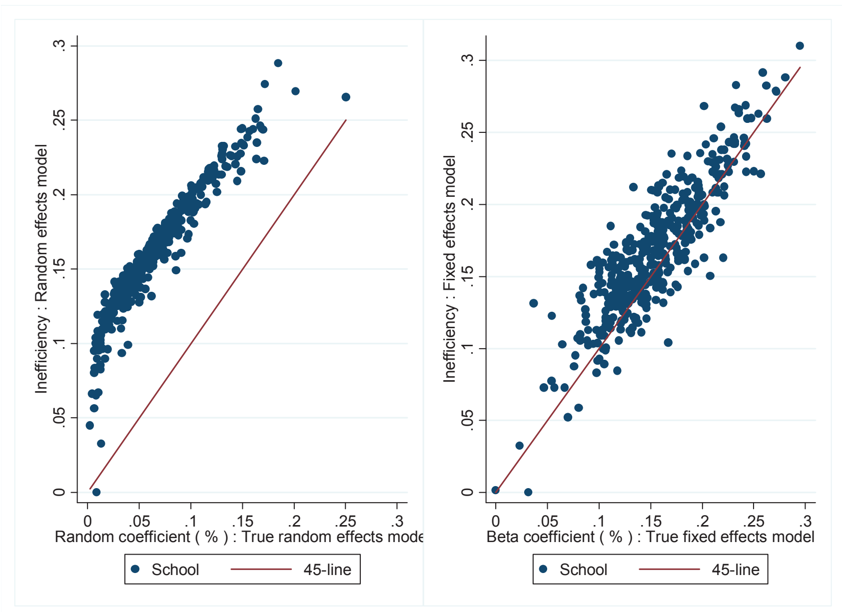
Figure 4. *Scatter plots illustrating the association between inefficiency scores for individual schools produced by RE and TRE models (left panel) and FE and TFE models (right panel) using scores in the compulsory tests in the matriculation examination as the dependent variable. Data are from 2004. Specification A*



To further illustrate the differences between the models, the inefficiency scores of the RE model are compared with random coefficients (as percentage deviation from the maximum value) of the TRE model. The same comparison is also carried out between the inefficiency scores in the FE model and the beta coefficients (also as percentage deviation from the maximum value) in TFE model (see Figure 5). In the TRE model, the random coefficient depicts the school-specific random effect. In TFE model the school-specific fixed effects are depicted by the beta coefficient.

The figure shows that the time-constant random effect of schools is very similar in RE and TRE models. As expected, only the magnitude is smaller in the TRE model. The same pattern is observed between FE and TFE models. The difference in magnitude is, however, smaller than between RE and TRE models. Unfortunately, the problem of what should be counted as inefficiency remains unsolved. The advantage of TRE and TFE models is, however, that they allow the separate identification and investigation of time-constant effects and time-varying inefficiency.

Figure 5. *Scatter plots illustrating the association between the inefficiency score of the RE model and random coefficient of the TRE model(left panel) and between the inefficiency score of the FE model and the beta coefficient of the TFE model (right panel) using scores in the compulsory tests in the matriculation examination as the dependent variable. Data are from 2004. Specification A*



6. Conclusions

In this study, new variants of stochastic frontier models for panel data were used to evaluate the efficiency of Finnish general upper secondary schools. These models allow the separation of random and fixed effects from inefficiency. Hence, they take the school-specific time-constant heterogeneity into account and allow omitted variables bias, which is quite common in education production function studies, to be controlled for. True fixed effects model also to some extent controls for the non-random selection of students into schools.

The results of both random and fixed effects models were presented. The choice between these two alternatives is complicated and influenced by several factors. The results of the Hausman specification test supported the fixed effects models as did the less restrictive assumption of fixed effects models that unmeasured heterogeneity can be correlated with the included variables. True fixed effects models, however, only take into account the within-school variation. Since the panel used in the study was fairly short, there was no variation through time in some of the variables, and as the variation across schools was also considered important, the results of random effects models were additionally presented.

The estimation results were very similar to previous studies. Variables related to students' earlier school success (comprehensive school GPA) and family background were the strongest predictors of performance in the matriculation examination. In most of the models, the effect of school resources (teaching expenditures per student) was not statistically significant. When the school-specific heterogeneity was taken into account with the true random and true fixed effects models, the coefficient became significant. The effect was, however, small and negative. The effect of the student-teacher ratio estimated with a smaller sample had a small, statistically significant positive coefficient in random effects and true fixed effects models.

The reforms which took place in the mid 1990s in general upper secondary schools did not turn out to be beneficial. A longer stay in general upper secondary school was not improving the scores in the matriculation examination since schools with longer length of studies performed worse. Neither did the decentralization of test taking. The more students decentralized their test taking in the matriculation examination the worse they performed. To gain more certainty on these two matters, they should be further studied with student level data.

The estimated inefficiency varied depending on the stochastic frontier model. The average inefficiency was between 3–17% depending on the model and the year. Taking into account the unmeasured heterogeneity with the true random and true fixed effects models reduced the inefficiency. Fixed effects models produced highest estimates for inefficiency as well as for the variation in inefficiency scores across schools.

The ranking of schools based on their inefficiency score also considerably changed in true random and true fixed effects models compared to other models. However, the rankings remained stable between random effects and true random effects models. Only the size of the effects was somewhat reduced in the true random effects model. They same pattern emerged between fixed and true fixed effects models.

Different stochastic frontier models were robust to variable specifications. The replacement of teaching expenditures per student with student teacher ratio and the use of a smaller data had only minor influence on the results.

To conclude, the choice of the stochastic frontier model matters. In random and fixed effects models, the identified inefficiency mostly depicts permanent differences between the operations of schools. In true random and true fixed effects models the inefficiency term captures smaller yearly fluctuations in input and output variables and the random or fixed effects terms depict permanent differences across schools. Permanent effects are smaller in magnitude in the latter cases.

From the policy point of view, if the policymaker is interested in identifying short term efficiency improvements, calculation of inefficiency scores are more appropriate to base on true random or true fixed effects models. If structural or permanent changes are pursued, random and fixed effects should be considered as identifiers of differences. A closer analysis of these terms may reveal some omitted factors but also potential for efficiency improvement. However, the advantage of using true random and true fixed effects models is that they allow the separation and investigation of both of these components.

References

Aigner, D. J. – Lovell, C. A. K. – Schmidt, P. (1977): Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics*, 6, 21–37.

Akerhielm, K. (1995): Does Class Size Matter? *Economics of Education Review* 14, 229–241.

Angrist, J. D. – Lavy, V. (1999): Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics* 114, 533–575.

Averett, S. – McLennan, M. C. (2004): Exploring the Effect of Class Size on Student Achievement: What Have We Learned Over the Past Two Decades. In Johnes, G. and Johnes, J. (eds.) *International Handbook on the Economics of Education*. Northampton: Edward Elgar Publishing.

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

Battese, G. – Coelli, T. (1992): Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis* 3, 153–169.

Battese, G. – Coelli, T. (1995): A Model for Technical Efficiency Effects in a Stochastic Frontier Production Function for Panel Data. *Empirical Economics* 20, 325–332.

Coelli, T. (1995): Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis. *Journal of Productivity Analysis* 6, 247–268.

Coelli, T. – Rao, P.D.S. – Battese, G. E. (1999): *An Introduction to Efficiency and Productivity Analyses.* Kluwer Academic Publishers. Boston.

Cooper, S. T. – Cohn, E. (1997): Estimation of a Frontier Production Function for the South Carolina Educational Process. *Economics of Education Review* 16, 313–327.

Cornwell, C. – Schmidt, P. – Sickles, R. (1990): Production Frontiers with Cross Sectional and Time Series Variation in Efficiency Levels. *Journal of Econometrics* 46, 185–200.

Deller, S. C. – Rudnicki, E. (1993): Production Efficiency in Elementary Education: The Case of Maine Public Schools. *Economics of Education Review* 12, 45–57.

Goldhaber, D. D. – Brewer, D. J. (1997): Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity. *Journal of Human Resources* 32, 505–523.

Gonand, F. – Joumard, I. – Price, R. (2007): Public Spending Efficiency: Institutional Indicators in Primary and Secondary Education. *OECD Economics Department Working Papers* No. 543. Paris: OECD. www.oecd.org/eco/working_papers.

Greene, W. (2005a): Fixed and Random Effects in Stochastic Frontier Models. *Journal of Productivity Analysis* 23, 7–32.

Greene, W. (2005b): Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model. *Journal of Econometrics* 126, 269–303.

Hanushek, E. A. (2003): The Failure of Input-Based Schooling Policies. *The Economic Journal* 113, F64–F98.

Heshmati, A. – Kumbhakar, S. C. (1997): Efficiency of the Primary and Secondary Schools in Sweden. *Scandinavian Journal of Educational Research* 41, 33–52.

Hoxby, C. M. (2000): The Effects of Class Size and Composition of Student Achievement: New Evidence from Natural Population Variation. *Quarterly Journal of Economics* 115, 1239–1285.

Johnes, G. – Johnes, J. (2009): Higher Education Institutions' Costs and Efficiency: Taking the Decomposition a Further Step. *Economics of Education Review* 28, 107-113.

Johnes, J. (2004): Efficiency Measurement. In Johnes, G. and Johnes, J. (eds.) *International Handbook on the Economics of Education*. Northampton: Edward Elgar Publishing.

Krueger, A. (2003): Economic Considerations and Class Size. *The Economic Journal* 113, F34–F63.

Kumbhakar, S. C. (1990): Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency. *Journal of Econometrics* 46, 201–212.

Kumbhakar, S. C. – Lovell, K. C. A. (2000): *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

Lee, Y. H. – Schmidt, P. (1993): A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency. In Fried H. O. – Lovell, C. A. K. – Schmidt, S. S., eds. *The Measurement of Productive Efficiency: Techniques and Applications.* New York: Oxford University Press.

Meeusen, W. – van den Broeck, J. (1977): Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review* 18, 435–444.

Mizala, A., Romaguera, P. and Farren, D. (2002): The Technical Efficiency of Schools in Chile. *Applied Economics* 34, 1533–1552.

Pitt, M. – Lee, L. F. (1981): The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics* 9, 43–64.

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

Sutherland, D. – Price, R. – Joumard, I. – Nicq, C. (2007): Performance Indicators for Public Spending Efficiency in Primary and Secondary Education. *OECD Economics Department Working papers* No. 546. Paris: OECD. www.oecd.org/eco/working_papers.

Worthington, A. C. (2001): An Empirical Survey of Frontier Efficiency Measurement Techniques in Education. *Education Economics* 9, 245–268.

Wössmann, L. – Schütz, G. (2006): Efficiency and Equity in European Education and Training Systems. Analytical Report for the European Commission. Communation and Staff Working Paper by the European Commission. http://ec.europa.eu/education/policies/2010/doc/eenee.pdf

Appendices

Appendix 1. Summary statistics of the variables in Specification A

		Mean	Std.	Min.	Max.	Observations
Scores in compulsory tests/student	Overall	16.67	1.77	10.71	23.48	N = 2133
	Between		1.62	11.93	23.13	n = 436
	Within		0.76	13.19	19.74	T-bar = 4.89
Scores in all tests/student	Overall	20.95	2.94	11.23	34.69	N = 2133
	Between		2.72	12.73	33.72	n = 436
	Within		1.22	16.28	27.06	T-bar = 4.89
Grade point average (GPA)	Overall	8.31	0.30	7.20	9.55	N = 2133
	Between		0.27	7.48	9.46	n = 436
	Within		0.13	7.27	8.80	T-bar = 4.89
Parental education	Overall	389.30	81.55	186.00	754.50	N = 2133
	Between		72.25	258.30	720.10	n = 436
	Within		30.17	172.89	534.50	T-bar = 4.89
Proportion of white collar workers	Overall	20.69	10.36	0.00	66.20	N = 2133
	Between		9.62	4.28	60.76	n = 436
	Within		4.11	-6.75	46.05	T-bar = 4.89
Proportion of single parents	Overall	13.95	6.35	0.00	66.70	N = 2133
	Between		4.67	2.92	30.06	n = 436
	Within		4.31	-0.79	50.59	T-bar = 4.89
SD of grades in the matriculation examination	Overall	4.69	0.68	2.34	8.19	N = 2133
	Between		0.48	3.04	6.08	n = 436
	Within		0.49	2.57	8.24	T-bar = 4.89
Proportion of females	Overall	0.57	0.10	0.11	1.00	N = 2133
	Between		0.07	0.32	0.81	n = 436
	Within		0.07	0.25	0.89	T-bar = 4.89
Proportion of Swedish speakers	Overall	0.06	0.22	0.00	1.00	N = 2133
	Between		0.22	0.00	0.97	n = 436
	Within		0.02	-0.18	0.37	T-bar = 4.89
Teaching expenditures/students	Overall	3495.08	987.50	1819.50	12991.63	N = 2133
	Between		975.92	1985.74	11120.66	n = 436
	Within		237.79	1395.40	5368.72	T-bar = 4.89
Other expenditues/student	Overall	1085.40	352.38	383.34	3276.37	N = 2133
	Between		360.29	434.89	3118.31	n = 436
	Within		116.08	448.48	1700.16	T-bar = 4.89
Mean length of studies	Overall	3.11	0.11	2.53	3.75	N = 2133
	Between		0.09	2.93	3.59	n = 436
	Within		0.06	2.66	3.51	T-bar = 4.89
Mean participation in test periods	Overall	1.79	0.28	1.00	3.00	N = 2133
	Between		0.22	1.26	2.45	n = 436
	Within		0.17	1.10	2.68	T-bar = 4.89

Appendix 1. continues

School size	Overall	252.76	149.43	23.00	879.00	N = 2133
	Between		147.51	28.00	840.00	n = 436
	Within		17.69	72.16	477.16	T-bar = 4.89
Languages	Overall	0.01	0.09	0.00	1.00	N = 2133
	Between		0.08	0.00	1.00	n = 436
	Within		0.03	-0.79	0.41	T-bar = 4.89
Mathematics and science	Overall	0.03	0.17	0.00	1.00	N = 2133
	Between		0.17	0.00	1.00	n = 436
	Within		0.00	0.03	0.03	T-bar = 4.89
Music and arts	Overall	0.04	0.19	0.00	1.00	N = 2133
	Between		0.19	0.00	1.00	n = 436
	Within		0.03	-0.76	0.64	T-bar = 4.89
Sports	Overall	0.03	0.17	0.00	1.00	N = 2133
	Between		0.16	0.00	0.00	n = 436
	Within		0.00	0.03	0.03	T-bar = 4.89
Private schools	Overall	0.06	0.23	0.00	1.00	N = 2133
	Between		0.23	0.00	1.00	n = 436
	Within		0.02	-0.54	0.46	T-bar = 4.89
State owned	Overall	0.01	0.11	0.00	1.00	N = 2133
	Between		0.15	0.00	1.00	n = 436
	Within		0.00	0.01	0.01	T-bar = 4.89
Densely populated area	Overall	0.20	0.40	0.00	1.00	N = 2133
	Between		0.40	0.00	1.00	n = 436
	Within		0.00	0.20	0.20	T-bar = 4.89
Rural area	Overall	0.33	0.47	0.00	1.00	N = 2133
	Between		0.47	0.00	1.00	n = 436
	Within		0.00	0.33	0.33	T-bar = 4.89

Appendix 2. Summary statistics for variables in Specification B

		Mean	Std. Dev.	Min	Max	Observations
Scores in compulsory tests/student	Overall	16.59	1.66	10.71	23.30	N = 1798
	Between		1.51	11.93	22.68	n = 376
	Within		0.73	13.68	19.37	T-bar = 4.78
Scores in all tests/student	Overall	20.82	2.69	11.23	33.89	N = 1798
	Between		2.47	12.73	32.78	n = 376
	Within		1.18	17.27	25.74	T-bar = 4.78
Grade point average (GPA)	Overall	8.30	0.28	7.20	9.55	N = 1798
	Between		0.25	7.48	9.48	n = 376
	Within		0.13	7.26	8.77	T-bar = 4.78
Parental education	Overall	381.30	74.38	186.00	702.50	N = 1798
	Between		70.17	261.70	673.60	n = 376
	Within		28.39	237.70	514.40	T-bar = 4.78
Proportion of white collar workers	Overall	19.64	9.56	0.00	59.90	N = 1798
	Between		8.97	4.28	55.08	n = 376
	Within		3.81	-0.89	36.86	T-bar = 4.78
Proportion of single parents	Overall	13.34	5.93	0.00	37.80	N = 1798
	Between		4.35	2.92	27.55	n = 376
	Within		4.07	-1.40	30.22	T-bar = 4.78
SD of grades in the matriculation examination	Overall	4.72	0.66	2.71	7.61	N = 1798
	Between		0.47	3.32	6.08	n = 376
	Within		0.47	2.90	6.93	T-bar = 4.78
Proportion of female students	Overall	56.71	9.92	14.29	93.65	N = 1798
	Between		7.19	32.32	80.57	n = 376
	Within		6.99	27.90	81.03	T-bar = 4.78
Proportion Swedish speakers	Overall	6.17	22.02	0.00	100.00	N = 1798
	Between		22.17	0.00	97.06	n = 376
	Within		1.14	-9.39	13.48	T-bar = 4.78
Student-teacher ratio	Overall	17.98	3.85	5.90	29.88	N = 1798
	Between		3.62	6.03	28.43	n = 376
	Within		1.44	10.93	24.93	T-bar = 4.78
Other expenditures/student	Overall	1058.37	307.45	383.34	2971.37	N = 1798
	Between		293.78	449.33	2377.88	n = 376
	Within		105.91	430.74	1673.13	T-bar = 4.78
Mean length of studies	Overall	3.11	0.10	2.53	3.65	N = 1798
	Between		0.09	2.93	3.59	n = 376
	Within		0.06	2.66	3.51	T-bar = 4.78
Mean participation in test periods	Overall	1.78	0.28	1.00	2.78	N = 1798
	Between		0.23	1.18	2.45	n = 376
	Within		0.16	1.17	2.59	T-bar = 4.78

Appendix 2. continues

School size	Overall	253.40	150.43	24.00	879.00	N = 1798
	Between		149.77	27.88	839.90	n = 376
	Within		18.34	72.20	477.70	T-bar = 4.78
Sports	Overall	0.03	0.16	0.00	1.00	N = 1798
	Between		0.16	0.00	1.00	n = 376
	Within		0.00	0.03	0.03	T-bar = 4.78
Languages and communication	Overall	0.00	0.06	0.00	1.00	N = 1798
	Between		0.06	0.00	1.00	n = 376
	Within		0.03	-0.60	0.40	T-bar = 4.78
Music and arts	Overall	0.04	0.20	0.00	1.00	N = 1798
	Between		0.20	0.00	1.00	n = 376
	Within		0.00	0.04	0.04	T-bar = 4.78
Mathematics and science	Overall	0.02	0.15	0.00	1.00	N = 1798
	Between		0.15	0.00	1.00	n = 376
	Within		0.00	0.02	0.02	T-bar = 4.78

Appendix 3. The results of stochastic frontier models for panel data from 2000-2004. Dependent variable: scores in all tests in the matriculation examination. Specification A

	Pooled panel data model	Random effects model	Fixed effects model	True fixed effects model
GPA	2.210	2.296	2.271	2.258
	(35.25)**	(44.54)**	(18.51)**	(39.57)**
Parental education	0.165	0.088	0.047	0.179
	(8.25)**	(5.23)**	(2.09)*	(11.38)**
% white collar workers	0.001	0.001	0.001	0.001
	(3.39)**	(3.14)**	(2.55)*	(3.11)**
% single parents	-0.001	-0.0004	-0.0001	-0.002
	(-4.47)**	(-1.59)	(-0.18)	(-7.59)**
Teaching expenditures/students	-0.013	-0.019	0.025	-0.012
	(-1.07)	(-1.25)	(0.79)	(-1.46)
Other expenditures/student	0.001	0.012	0.030	0.010
	(0.19)	(1.40)	(2.13)*	(2.12)*
SD of matriculation examination	-0.085	-0.053	-0.017	-0.089
	(-6.39)**	(-4.76)**	(-1.35)	(-7.82)**
% female	0.001	0.001	0.0004	0.001
	(5.64)**	(3.58)**	(2.54)*	(7.00)**
% Swedish speaking	0.001	0.001	0.0002	0.001
	(12.10)**	(7.54)**	(0.19)	(17.48)**
Mean length of studies	-0.292	-0.082	-0.070	-0.274
	(-4.81)**	(-1.84)	(-1.12)	(-6.42)**
Mean participation in exam periods	-0.114	-0.090	-0.108	-0.112
	(-9.79)**	(-7.97)**	(-6.62)**	(-13.50)**
School size	-0.014	-0.018	-0.017	-0.018
	(-2.72)**	(-3.05)**	(-0.82)	(-5.50)**
Languages and communication	0.049	0.036	-0.010	0.042
	(2.55)*	(4.22)**	(-0.25)	(3.30)**
Mathematics and sciences	-0.011	-0.018		
	(-1.09)	(-1.25)		
Music and arts	-0.008	-0.001	0.014	-0.013
	(-0.79)	(-0.07)	(0.51)	(-2.10)*
Sports	0.028	0.007		
	(2.50)*	(0.49)		
Private schools	-0.002	0.028		
	(-0.23)	(3.15)**		
State owned schools	0.071	0.136		
	(4.05)**	(11.97)**		
Densely populated area	0.020	-0.006		
	(3.90)**	(-0.67)		

Appendix 3. continues

Rural area	0.017	-0.014		
	(2.57)*	(-1.60)		
Year	0.002	0.003	0.001	0.003
	(1.22)	(3.46)**	(0.75)	(34.69)**
Constant	-5.213	-8.220	-4.712	
	(-1.97)*	(-4.68)**	(-1.48)	
Lambda	0.872	0.234		0.878
	(15.05)**	(17.07)**		(12.00)**
Sigma(v)	0.068	0.489		0.078
Sigma(u)	0.059	0.114		0.088
Het: School size		-0.883		
		(-21.76)**		
Log-L	2451.86	3004.05		2718.55
No. of obs.	2133	2133	2133	2133
No. of schools	436	436	436	436

Note. T-values in parentheses, robust t-values in the fixed effects model. *Significant at 5%, **Significant at 1%. All explanatory variables in a logarithmic form except ratios. Schools maintained by the municipalities are the reference group for the private and state own schools. Schools with general curriculum are the reference group for the schools with specialized curriculum. Schools in urban areas are the reference group for the schools in densely populated or rural areas.

Appendix 4. The results of stochastic frontier models for panel data from 2000-2004. Dependent variable: scores in all tests in the matriculation examination. Specification B

	Pooled panel data model	Random effects model	True random effects model	Fixed effects model	True fixed effects model
GPA	2.220	2.258	2.263	2.208	2.249
	(32.48)**	(39.94)**	(58.17)**	(16.49)**	(36.72)**
Parental education	0.129	0.086	0.068	0.046	0.134
	(5.98)**	(4.59)**	(5.32)**	(1.96)*	(8.21)**
% white collar workers	0.001	0.001	0.001	0.001	0.001
	(3.02)**	(2.51)*	(5.11)**	(2.54)*	(2.94)**
% single parents	-0.001	-0.0004	-0.0002	0.0001	-0.002
	(-3.46)**	(-1.50)	(-0.78)	(0.41)	(-6.16)**
Student-teacher ratio	0.001	0.027	-0.004	0.015	0.007
	(0.14)	(2.72)**	(-0.71)	(0.80)	(1.18)
Other expenditures per student	0.007	0.021	0.015	0.031	0.008
	(0.95)	(2.20)*	(2.87)**	(1.96)*	(1.55)
SD of matriculation examination scores	-0.081	-0.044	-0.022	-0.003	-0.091
	(-5.55)**	(-3.69)**	(-2.38)*	(-0.22)	(-7.67)**
% female	0.001	0.001	0.001	0.001	0.001
	(4.52)**	(3.42)**	(4.70)**	(2.57)**	(5.61)**
% Swedish speaking	0.001	0.001	0.001	0.0003	0.001
	(11.88)**	(7.95)**	(21.47)**	(0.35)	(17.50)**
Mean length of studies	-0.260	-0.075	-0.178	-0.137	-0.217
	(-3.99)**	(-1.60)	(-4.33)**	(-2.08)*	(-4.84)**
Mean participation in exam periods	-0.118	-0.093	-0.103	-0.106	-0.118
	(-9.47)**	(-7.69)**	(-13.53)**	(-6.04)**	(-13.62)**
Densely populated area	0.019	-0.005	0.012		
	(3.48)**	(-0.69)	(3.31)**		
Rural area	0.015	-0.013	0.006		
	(2.16)*	(-1.41)	(1.30)		
School size	-0.002	-0.013	-0.001	-0.050	-0.009
	(-0.39)	(-1.85)	(-0.31)	(-2.16)*	(-2.93)**
Languages and communication	0.015	-0.042	-0.027	-0.066	0.015
	(0.52)	(-1.91)	(-1.33)	(-3.81)**	(0.79)
Mathematics and sciences	-0.016	-0.005	-0.018		
	(-1.36)	(-0.30)	(-2.18)*		
Music and arts	-0.009	0.001	-0.001		
	(-0.87)	(0.10)	(-0.17)		
Sports	0.025	0.015	0.007		
	(2.08)*	(1.14)	(0.85)		
Year	0.002	0.002	0.002	0.001	0.003
	(1.19)	(2.37)*	(2.12)*	(0.84)	(36.79)**

Appendix 4. continues

Constant	-5.344	-6.590		-4.037	
	(-1.91)	(-3.55)**		(-1.42)	
Lambda	1.044	0.208	0.881		1.208
	(15.94)**	(15.57)**	(5.95)**		(15.34)**
Sigma(v)	0.066	0.513	0.044		0.080
Sigma(u)	0.063	0.107	0.038		0.097
Het: School size		-0.907			
		(-19.59)**			
Log-L	2110.30	2574.50	-2480.68		2329.63
No. of obs.	376	376	376	376	376
No. of schools	1798	1798	1798	1798	1798

Note. T-values in parentheses, robust t-values in the fixed effects model. *Significant at 5%, **Significant at 1%. All explanatory variables in a logarithmic form except ratios. Schools maintained by the municipalities are the reference group for the private and state own schools. Schools with general curriculum are the reference group for the schools with specialized curriculum. Schools in urban areas are the reference group for the schools in densely populated or rural areas.

Appendix 5. Average inefficiency, standard deviation, minimum and maximum in panel data models using scores in compulsory tests in the matriculation examination as the dependent variable. Specification A

	Average	Standard deviation	Minimum	Maximum
Pooled panel data model				
2000	0.044	0.021	0.009	0.159
2001	0.044	0.023	0.008	0.143
2002	0.047	0.023	0.010	0.131
2003	0.045	0.023	0.011	0.171
2004	0.044	0.020	0.010	0.142
Random effects model				
2000	0.064	0.042	0.002	0.252
2001	0.064	0.042	0.002	0.252
2002	0.064	0.042	0.002	0.252
2003	0.064	0.041	0.002	0.252
2004	0.064	0.042	0.002	0.252
True random effects model				
2000	0.027	0.010	0.005	0.102
2001	0.027	0.010	0.010	0.071
2002	0.029	0.010	0.010	0.084
2003	0.027	0.010	0.009	0.095
2004	0.026	0.011	0.010	0.087
Fixed effects model				
2000	0.155	0.046	0.000	0.295
2001	0.155	0.046	0.000	0.295
2002	0.155	0.046	0.000	0.295
2003	0.154	0.046	0.000	0.295
2004	0.154	0.046	0.000	0.295
True fixed effects model				
2000	0.055	0.014	0.020	0.143
2001	0.055	0.014	0.023	0.115
2002	0.059	0.014	0.026	0.132
2003	0.056	0.014	0.025	0.136
2004	0.056	0.015	0.024	0.134

Appendix 6. *Average inefficiency, standard deviation, minimum and maximum in panel data models using scores in all tests in the matriculation examination as the dependent variable. Specification A*

	Average	Standard deviation	Minimum	Maximum
Pooled panel data model				
2000	0.048	0.017	0.010	0.125
2001	0.047	0.018	0.015	0.131
2002	0.049	0.018	0.017	0.120
2003	0.049	0.019	0.016	0.187
2004	0.047	0.017	0.017	0.148
Random effects model				
2000	0.094	0.058	0.003	0.347
2001	0.094	0.058	0.003	0.347
2002	0.094	0.058	0.003	0.347
2003	0.096	0.060	0.003	0.347
2004	0.096	0.060	0.003	0.347
Fixed effects model				
2000	0.273	0.054	0.104	0.426
2001	0.273	0.056	0.000	0.426
2002	0.273	0.056	0.000	0.426
2003	0.272	0.055	0.000	0.426
2004	0.272	0.055	0.000	0.426
True fixed effects model				
2000	0.056	0.010	0.024	0.116
2001	0.056	0.009	0.031	0.097
2002	0.058	0.009	0.031	0.100
2003	0.058	0.009	0.037	0.120
2004	0.057	0.011	0.028	0.112

Appendix 7. Average inefficiency, standard deviation, minimum and maximum in panel data models using scores in compulsory tests in the matriculation examination as the dependent variable. Specification B

		Average	Standard deviation	Minimum	Maximum
Pooled panel data model					
	2000	0.026	0.008	0.010	0.066
	2001	0.026	0.008	0.009	0.068
	2002	0.027	0.007	0.010	0.055
	2003	0.026	0.008	0.011	0.061
	2004	0.026	0.007	0.011	0.052
Random effects model					
	2000	0.064	0.041	0.002	0.214
	2001	0.065	0.043	0.002	0.214
	2002	0.065	0.043	0.002	0.214
	2003	0.065	0.042	0.002	0.214
	2004	0.065	0.042	0.002	0.214
True random effects model					
	2000	0.029	0.012	0.006	0.120
	2001	0.030	0.014	0.009	0.091
	2002	0.032	0.012	0.012	0.099
	2003	0.030	0.013	0.008	0.105
	2004	0.029	0.013	0.010	0.103
Fixed effects model					
	2000	0.175	0.054	0.000	0.315
	2001	0.175	0.057	0.000	0.315
	2002	0.176	0.057	0.000	0.315
	2003	0.174	0.057	0.000	0.315
	2004	0.174	0.057	0.000	0.315
True fixed effects model					
	2000	0.055	0.014	0.020	0.146
	2001	0.056	0.014	0.024	0.119
	2002	0.059	0.013	0.027	0.132
	2003	0.057	0.015	0.025	0.138
	2004	0.056	0.015	0.024	0.135

Appendix 8. Average inefficiency, standard deviation, minimum and maximum in panel data models using scores in all tests in the matriculation examination as the dependent variable. Specification B

		Average	Standard deviation	Minimum	Maximum
Pooled panel data model					
	2000	0.053	0.022	0.009	0.154
	2001	0.052	0.023	0.014	0.157
	2002	0.053	0.022	0.018	0.140
	2003	0.054	0.024	0.016	0.220
	2004	0.051	0.021	0.015	0.174
Random effects model					
	2000	0.088	0.053	0.003	0.291
	2001	0.089	0.056	0.003	0.291
	2002	0.089	0.056	0.003	0.291
	2003	0.089	0.055	0.003	0.291
	2004	0.088	0.055	0.003	0.291
True random effects model					
	2000	0.030	0.010	0.007	0.107
	2001	0.030	0.010	0.011	0.079
	2002	0.031	0.009	0.012	0.075
	2003	0.031	0.011	0.013	0.107
	2004	0.029	0.011	0.011	0.096
Fixed effects model					
	2000	0.179	0.061	0.000	0.341
	2001	0.179	0.064	0.000	0.341
	2002	0.179	0.064	0.000	0.341
	2003	0.179	0.064	0.000	0.341
	2004	0.179	0.063	0.000	0.341
True fixed effects model					
	2000	0.068	0.015	0.024	0.167
	2001	0.068	0.015	0.032	0.134
	2002	0.070	0.014	0.032	0.140
	2003	0.070	0.016	0.040	0.175
	2004	0.068	0.016	0.031	0.161

Article IV

UNDERSTANDING EFFICIENCY DIFFERENCES OF SCHOOLS: PRACTITIONERS' VIEWS ON STUDENTS, STAFF RELATIONS, SCHOOL MANAGEMENT AND THE CURRICULUM

Tanja Kirjavainen

Abstract

This study analyses the views of the staff members of nine general upper secondary schools that were mostly in the upper or lower tails of the efficiency distribution measured with stochastic frontier analysis and aims at further describing the schools. Teachers and principals were interviewed on their views about the students, staff relations, school management, curriculum work, parent-school relations, teacher training, and evaluation. In efficient schools, views concerning the students were caring, appreciating all students, including the weaker ones. Respecting views were also present, with students' own initative being respected. In inefficient schools there was more often frustration or disappointment at the low performance of the students. In efficient schools, staff relations were professional, whereas in inefficient schools problems more often occurred. Management and decision making were participative in efficient schools and teachers were happy with their possibilities to influence school matters. In inefficient schools, there were more often disappointments and frustrated views about the management and possibilities to have an influence. Curriculum work was more often seen as way to develop the school and the work in efficient schools. In inefficient schools, it was considered as an administrative measure.

Key words: Efficiency, General upper secondary schools, School management, Staff relations, Stochastic frontier analysis

Contents

1. Introduction

In educational research, school effectiveness research has the aim of identifying school processes that characterize effective schools. Effective schools are determined as those having the highest student performance taking into account their intake of students. The research setting is similar to studies on efficiency differences. As effective school research bases the identification of effective schools mainly on student achievement corrected with earlier test scores and family background, studies on school efficiency also take into account school resources. In addition, the methods for identifing differences are distinct. School effectivenss studies usually apply multilevel modelling based on individual level data.[1] Studies of school efficiency use school or district level data and either parametric or non-parametric methods for efficiency measurement.[2]

In school effectiveness research, both quantitative and qualitative methods are used to examine school processes. Studies have usually concentrated on the effect of school climate, leadership, instructional arrangements, staff development, and monitoring on effectiveness.[3] In school efficiency literature the main emphasis has been on quantitative methods both in determining as well as in explaining efficiency differences.[4] Explanatory factors that have been used relate to school size, governance, and competition (see e.g. Bradley *et al*., 2001; Duncombe *et al*., 1997; Grosskopf *et al*., 2001; Kirjavainen and Loikkanen, 1998). In recent studies, qualitative approaches have also been used to study the role of school leadership, management, staff relations and evaluation practices (Dodd, 2006; Portela and Camanho, 2007).

In this study, I follow earlier case studies of effective schools and school efficiency by investigating the school practices and characteristics of nine Finnish general upper secondary schools that

---

[1] For multilevel modelling see e.g. Goldstein (2003) and Raudenbush and Bryk (2002).

[2] For methods of efficiency measurement see e.g. Coelli *et al*. (1999).

[3] See for reviews Teddlie and Reynolds (2000) and Sammons (1999).

[4] See for reviews Johnes (2004) and Worthington (2001).

were mostly in the upper and lower tails in the efficiency distributions based on earlier analysis (Kirjavainen, 2007a). I concentrate on the views of principals and teachers on students, staff relations, school governance and management and curriculum development. The analysis is mainly based on semistructured interviews of principals and teachers. In analyzing the interview data, I first group the schools into categories based on the themes emerging from the interview texts and then concentrate on describing the characteristics of the emerging categories in more detail. Thus, my approach is different from earlier studies (see e.g. Sammons *et al*., 1998; Dodd, 2006; Portela and Camanho, 2007) that analyze and describe the school practices and views of the practitioners within different categories of school efficiency (schools with high and/or low efficiency). Only at the end of the study, I discuss the relation of the emerging categories with efficiency. With such an approach, I try to illuminate the differences in views and practices within each theme in more detail and attempt to create some new insights into the evaluation of schools and their performance.

The identification of efficiency differences is also different from earlier studies. It is based on stochastic frontier analysis and use of a five-year panel data of Finnish general upper secondary schools and students graduating in 2000-2004 (see Kirjavainen, 2007a). The panel data enabled the identification of school-specific effects that are constant through time. These effects are interpreted as inefficiency in this study.

In efficiency measurement, the concept of technical efficiency was used. It refers to the ratio of observed to maximum potential outputs obtainable from the given inputs. School output was measured with average grades in compulsory subjects in the matriculation examination. The explanatory model controlled for the comprehensive school Grade Point Average (GPA) and family background. School resources inputs are measured with teaching expenditures and other expenditures. In addition, the model controlled for the length of studies, the average decentralization rate in the matriculation examination, school size, whether the school is municipal, private or state run, and the location of the school. With a high number of controls it was possible to exlude the most important student-related factors that affect student achievement.

The findings of school effectiveness research have shown that effective schools are characterized by outstanding leadership, effective instructional arrangements, focus on student acquisition of central learning skills, a productive school climate and culture, high operationalized expectations and requirements for students, appropriate monitoring of student progress, practice-oriented staff development at the school site, and salient parental involvement (see e.g. Reynolds and Teddlie, 2000).

According to results of this study, many of the findings related to staff relations, school governance and management and curriculum work were similar to earlier school effectiveness studies. Hence, efficiency measured with stochastic frontier analysis using panel data seem to produce similar results to earlier studies using different methods. In efficient schools, staff relations were professional and good, school governance and management was characterized as participative and school curriculum work as a way to develop the school and the work.

Views concerning students have not been so often addressed and the analysis of this study showed that they differed between the efficient and inefficient schools. In schools with highest efficiency, staff members' views concerning students were attentive emphasizing the care of all students, especially the weaker ones. The study also showed that some of the findings of school effectiveness studies related to evaluation and monitoring practices, staff development and parent-school relations did not apply to Finnish general upper secondary schools or at least to case schools of this study.

The paper continues as follows. In section 2, previous studies on effective and efficient schools are discussed. The design of the study is introduced in section 3 by first briefly discussing efficiency measurement and then describing inefficiency and other characteristics of the selected case schools. At the end of the section the data and analysis methods are presented. Sections 4 and 5 provide analysis of the interviews and section 6 concludes.

2. Research on effective and efficient schools

In the following, I briefly review some of the most influential and significant European studies on effective schools in order to illustrate their methods and results. A few recent studies on the efficiency of schools that complement the results of efficiency measurement with case studies are also discussed.

One of the early studies in school effectiveness research was a case study by Rutter *et al*. (1979) on English secondary schools. It concentrated on examining how differences in various school outcomes were related to school processes after taking into account the school intake. The data comprised 12 inner city schools in the London area that differed in terms of various outcomes. Rutter *et al*. had measures related to the intake of schools for assessing various characteristics of the students, the process of schooling concentrating on the social organization of the schools, and the outcomes of schooling. In addition, there were also some ecological measures referring to certain context or environmental factors. A large and rich data set was collected from the case schools that included questionnaires, interviews and observations. The report concentrated on discussing the correlates of the various measures.

The results of Rutter *et al*. demonstrated that the outcome measures were fairly stable over time and schools performing well in one measure usually also performed well in other measures. Differences in various outcomes were not, however, wholly accounted for by the background of students, physical factors such as the size of the school, the age of the buildings or the space available, administrative status or organization. Instead, differences were systematically related to the schools' characteristics as social institutions. Factors such as the degree of academic emphasis, teacher actions in lessons, the availability of incentives and rewards, good conditions for pupils, and the extent to which children were able to take responsibility were all significantly related to variation between schools. Outcomes were also influenced by the factors outside the immediate control of the teachers. The socioeconomic status of the students, for example, was positively related to outcomes. It did not, however, influence the functioning of the schools based on measures of school processes. The total pattern of findings by Rutter *et al*. indicated a strong probability that the associations between school process and outcomes partly reflect a causal process. This means that to some extent the behaviour and attitudes of students are shaped and influenced by their experiences at school and especially by the qualities of the school as a social institution.

Mortimore *et al*. (1988) used a quite similar approach to investigate primary schools in England. In their study, students aged 7 to 11 in 50 randomly selected schools were followed up for four years. During the four years, various types of data were gathered describing the intake, student outcomes, cognitive and non-cognitive outcomes, classroom and school environment, school and class organization and policies, teacher strategies, views of parents and school life. Questionnaires, interviews and observations were used to collect information on these matters. The analysis of the data was mostly quantitative and applied various statistical methods. Interviews and observational data were mainly used to complement the statistical analyses and broaden the description of school processes.

According to the findings of Mortimore *et al*. (1988), some schools were better at fostering pupils' cognitive and non-cognitive skills, taking into account the intake of the schools, and the effect of the school was even greater than the effect of socioeconomic background of the students. Mortimore *et al*. identified 12 key factors that differentiated effective from ineffective schools. These were purposeful leadership of the staff by the headteacher, the involvement of the deputy head, the involvement of teachers, consistency amongst teachers, structured sessions, intellectually challenging teaching, work-centered environment, a limited focus within sessions, maximum communication between teachers and pupils, record keeping, parental involvement, and a positive climate.

According to Mortimore *et al*. (1988), effective schools were friendly and supportive environments led by heads who were not afraid to assert their views and yet be able to share management and decision making with staff. Class teachers within effective schools provided a structured learning situation for their pupils but gave them freedom with this framework. By being flexible in their use of whole class, group and individual contacts, they maximized communication with each pupil. Furthermore, through limiting their focus within sessions, their attention was less fragmented. Hence, the opportunities for developing a work-centered environment and for presenting challenging work to pupils increased.

Sammons *et al*. (1998) studied six inner London secondary schools that were outliers based on value added analysis of 94 schools. These included schools with low performance, high performance and schools with mixed effects. Headteachers and deputy heads were questioned about the processes of effectiveness. The results of Sammons *et al*. confirmed many of the earlier results concerning effective schools. According to them, effective schools were characterized as having high expectations and an emphasis on academic achievement, staff consensus and a shared vision of the purpose of the school, great stress on on the headteacher's leadership, a strong senior management team, the high importance of the quality of teaching, the importance of high examination entry and effective homework policies and practicies, encouragement of parental involvement and feedback. Ineffective and mixed schools were found to have problems in pupil behavior and attendance, whereas good behaviour and attendance were seen as necessary conditions for academic effectiveness, allowing the creation of a safe and orderly working environment and contributing to a positive culture.

Similar studies have also been conducted concerning U.S. schools (see e.g. Brookover *et al*., 1979 and Teddlie and Stringfield, 1993). The results of these and numerous other studies have been summarized by various researchers (see e.g. Purkey and Smith, 1983; Levine and Lezotte, 1990; Scheerens and Bosker, 1997; Sammons, 1999; Teddlie and Reynolds, 2000). These summaries have come up with a list of most important characteristics of effective schools. Factors that are in many reviews associated with effective schools include strong educational leadership, high expectations of student achievement, emphasis on basic skills, maximized time for learning, a safe and orderly climate, practice-oriented staff development at the school site, parental involvement and support, monitoring of student progress, and clear goals. Other correlates include school-site management that allows schools to decide on ways to improveme academic performance, instructional leadership that initiates and maintains the development process, staff stability, clear curriculum articulation and organization, schoolwide recognition of academic success, district support, collaborative planning and collegial relationship among staff, and a sense of community.

One of the few studies on school efficiency investigating the school characteristics and their impact on efficiency has concerned English secondary schools (Dodd, 2006). This study differs from the school effectiveness studies in the sense that it used data envelopment analysis (DEA) to identify unusually efficient schools. Several other inputs such as expenditures and teacher qualifications were also used in addition to measures of students' socioeconomic background and earlier school achievement.

The study aimed at verifying the results of DEA, investigating why these schools were efficient, and identifying the good practices that would contribute to school's efficiency. Case study methods were applied to 38 effective schools that were peers for over 100 other secondary schools. Information concerning leadership and governance, people management, policy and strategy, partnership and resources, and processes were gathered during visits to these schools. Staff members and in some schools also the parents participated in structured interviews. In addition, important documents such as the curriculum, strategic plans, and annual budgets were gathered and analysed in the study.

According to Dodd (2006) there were many characteristics that were common to efficient schools. However, the schools did also vary and not all characteristics were present in every school. Dodd identified such characteristics as a school ethos that emphasized learning and achievement, very strong leadership personalized in the head teacher but extending throughout the senior leadership team, monitoring of student performance and goal setting based on earlier results, very strong emphasis on the recruitment, retention and development of high quality staff, the staff's willingness to place extra effort on offerering support measures to students, the development of a curriculum that would reflect the needs of the students, inclusiveness so that every child is offered an opportunity to participate in education independent of their level of skill, proactiveness in seeking additional funding, strong commitment to planning, and extensive use of and investment in ICT. Dodd also found correlates that would be expected to contribute to school effectiveness but did not. They were related to school governance, the setting and class size, financial management, and the learning environment.

A study by Portela and Camanho (2007) has also taken a closer look at some of the secondary schools in Portugal based on their DEA efficiency scores. With the case study material from three benchmark schools and a few other schools, their purpose was to verify the results of efficiency analysis and identify the efficient school practices of benchmark schools. Even though Portela and Camanho did not state how these schools were studied, they came up with a list of characterisitics that differentiated the benchmark schools from other schools. According to their results, the benchmark schools had good resources and infrastructure, a motivated and stable body of teachers, a well defined and inclusive school mission, effective control, self-evaluation and rigorous use of student performance data, a high number of extra-curricular activities with a reasonable involvement of teachers and students, involvement of student' parents, and leadership well adapted to the school context.

Both studies suffered some shortcomings. No description is provided of how the results were obtained. Especially the study by Portela and Camanho (2007) lacks analysis, since there is no description of the case data and methods used for constructing the features of effective schools. Dodd (2006) only examined efficient schools. Since there was no comparison group, it might well be that the same characteristics are also present in the less efficient schools. The need for comparison is also widely established in the school effectiveness literature (see e.g. Purkey and Smith, 1983; Reynolds and Packer, 1992; Gray and Wilcox, 1995; Sammons *et al*. 1998).

School effectiveness research has been subject to much criticism in past years.[5] One focus of the criticism is associated with the fact that the practices of effective schools explain very little of the total variation in student achievement. For example, Scheerens and Bosker (1997) revealed in their meta-analysis that such school organizational factors as a productive climate and culture, pressure for achievement in basic subjects, educational leadership, monitoring/evaluation, cooperation/consensus, parental involvement, staff development, high expectations, and an orderly climate are only weakly related to school effectiveness based on the results of statistical studies.

The idea of stuying and detecting the features of effective schools and transferring their practices directly to ineffective schools has also been criticized. As mentioned by Reynolds and Teddlie (2001), there may be whole areas of schooling that are central to educational life in non-effective schools that simply cannot be seen in effective schools, such as staff groups that possess 'cliques' or interpersonal conflict between staff members, for example. To propose dropping into the context of the ineffective schools those factors that exist in the effective schools may be to generate simply unreachable goals for the ineffective schools, since the distance between the practice of one setting and the practice of another may be too great to be easily bridged. Therefore, Reynolds and Teddlie emphasize the importance of also studying school failure.

Research on failing schools has especially concentrated on the unfavourable context and environment in which these schools are working. Nicolaidou and Ainschow (2005) analysed the experiences

[5] See e.g. Coe and Fizz-Gibbon (1998) and Goldstein and Woodhouse (2000). For a review of the criticism, see Luyten *et al*., 2005.

and cultures of English schools judged as failing by the authorities and therefore being subject to special measures. Their study revealed that where the improvement efforts failed it was because of the assumption that such schools are faced with predictable and straightforward problems. According to Nicolaidou and Ainschow, the actual situation is quite the contrary and the schools are facing complex and simultaneously unique issues. Because of this, no one leadership style fits best, but there is a need for a reflective and shared leadership style that is adaptable to the schools' specific culture. The study also stressed the importance of studying cultural assumptions, since they provide insights into the way schools function and also facilitiate efforts at improvement.

The impact of the school context was also emhapsized in Lupton's (2005) study, which concentrated on examining its connection with the quality of schools. She studied four secondary schools in disadvantaged areas in England. According to Lupton, a high-poverty context exerts downward pressure on quality, and improvement measures concentrating solely on upskilling and motivating the staff will not lead to improved quality. The only way to ensure high quality is to provide additional resources because the schools in high-poverty areas face many additional tasks in addition to teaching. Lupton concluded that a higher level of resources provides the possibility to increase the organizational capacity of schools to better respond to the various problems of students.

3. Design of the study

Measurement of efficiency

Evaluation of the efficiency of the case schools was conducted statistically using Stochastic Frontier Analysis (SFA). Compared to ordinary statistical methods such as OLS, in SFA the error term is divided into a normally distributed error term and inefficiency.[6] As for the unit of observation, the identification of efficiency is based

[6] For the method, see Kumbhakar and Lovell (2000) and Greene (2005a, b).

on school- as opposed to student-level data as used in earlier school effectiveness studies applying multilevel models.[7]

The efficiency evaluation of the case schools was originally based on different stochastic frontier models that vary in how they treat and interpret the school effect, which is constant through time and not captured by the explanatory factors of the models. Here, the results of random and fixed effects models are considered. In these models the time constant effect is interpreted as inefficiency.[8] The data comprised 436 day schools from the years 2000-2004. The stochastic frontier models and the efficiency measurement are discussed in more detail in Kirjavainen (2007a; b).

The output of schools was measured with compulsory grades in matriculation examination. As for the explanatory variables, earlier student achievement was controlled with comprehensive school GPA. Students' socio-economic background was measured with the parents' educational level, the proportion of white-collar workers and the proportion of single parents. Other variables characterizing students were the proportion of female students and the proportion of Swedish-speaking students. The resource inputs were measured in monetary terms with two variables, teaching expenditures per student and other expenditures per student[9]. In addition, the model contained a variable measuring the length of the studies, the average decentralization rate in matriculation examinations[10], and the location (urban, densely populated and rural area) and size of the school.

[7] Efficiency measurement with multilevel models is described in Johnes (2004). For multilevel modeling see e.g. Goldstein (2003) and Raudenbush and Bryk (2002).

[8] Efficiency based on random and fixed effects models is discussed in this study because the study concentrates on organizational characteristics that undergo very little yearly fluctuation. These models depict the time constant variation across schools as inefficiency. The inefficiency term in other stochastic frontier models fluctuates yearly. In true random and true fixed effects models, inefficiency and time constant heterogeneity are separated. The inefficiency term captures small yearly variations in input and output variables which may be very difficult to explain with organizational characteristics.

[9] Other current expenditures consist of the cost of meals, health care and counselling, administration, and rents for the school premises.

[10] Tests can be taken in three consecutive examination periods.

The models also included a dummy for private schools and state-owned schools. The comparison group was schools maintained by the municipalities (the majority of schools).

Selection of case schools

Initally, the case schools were selected based on their inefficiency in an earlier study (see Kirjavainen, 1999). The problem with this setting, however, was that there was a considerable time lag between the efficiency measurement and the case study data collection. The efficiency measurement data was based on a cross-section of students graduating in 1991 and the selection of case schools as well as the data collection took place in 1999.

A few years ago I had an opportunity to conduct an efficiency analysis of general upper secondary schools using good quality data from students and schools that completed their studies in 2000 – 2004. Since some of the stochastic frontier models used in the study clearly divided the case schools into ones with a high and with a low inefficiency, I decided to take this measurement as the reference point. In doing so I was able to avoid the problem arising from the time lag which is quite typical in studies relating statistical data and case studies.[11] With this setting though, the two cohorts graduating in 2003 and 2004 started their studies after the data collection. Since changes in schools usually take place quite slowly these two cohorts most proably faced the same environment as the three earlier cohorts.

The drawback of having collected the case study data based on other efficiency measurement is that among case schools there are no such schools that would stay at the very bottom or very top in every measurement. I would, however, argue that since there is instability in the rankings between different stochastic frontier models it would be difficult to find cases that would robustly maintain their position in different measurements and models. Besides, with the analysis method adopted in this study, it is possible to study if the emerging categorizations have different effect on efficiency in different

[11] See Sammons *et al*. (1998) for problems in retrospective studies.

stochastic frontier models. As such, it gives additional insights to the interpretation of the results of efficiency measurement.

In the selection process, other factors besides efficiency were also considered. All the schools were Finnish speaking in order to make them comparable and because there were no Swedish speaking research assistance availabe. The sample included schools having quite a high comprehensive school GPA as well as schools accepting all the applicants. Some of the schools had been performing poorly and some exceptionally well in matriculation examination. The location of the schools also varied from urban to rural areas.[12] Both large and small schools were included, with the size of the schools ranging from around 100 to 700 students. The selected schools had also participated in the earlier survey in 1995 (see Kirjavainen, 1999), so that additional information was available on their organizational practices some years earlier. In the following, the names that I use to refer to each school are fictitious.

Efficiency of case schools

The efficiency of schools is depicted with an inefficiency score. This varies between 0 and 1 and the larger the figure the more inefficient (less efficient) the school is.[13] The average inefficiency differed to some extent in random and fixed effects models (see upper part of Table 1). The average inefficiency was higher in the fixed effects model, at 15%. According to this model, schools could have increased their output on average by 15%. In the random effects model, the average inefficiency was clearly lower, some 6%. There was only modest variation in the rankings of schools based on their inefficiency score in these two models (for further details, see Kirjavainen 2007a; b).

[12] The schools were located in and around Helsinki and not more than 150 kilometres from the city. This geographical constraint was set for practical reasons, i.e. to limit the amount of travel.

[13] The figures can also be interpreted as percentages indicating how much more output school could have produced to be efficient.

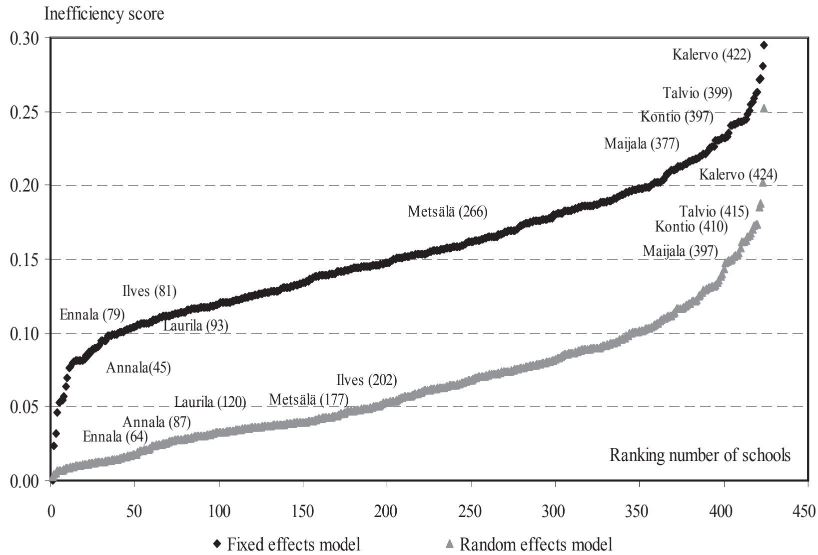
The inefficiency of case schools varied in the random effects model from 2.4% to 20.2% (see lower part of Table 1). In the fixed effects model the inefficiency varied from 10.2% to 27.2%.

Table 1. *Descriptive statistics and inefficiency scores of the case schools in random and fixed effects models in 2000-2004*

		Random effects model	Fixed effects model
All schools			
	Mean	0.064	0.154
	Std. dev.	0.041	0.046
	Min.	0.002	0.000
	Max.	0.250	0.295
Case schools			
	Annala	0.030	0.102
	Ennala	0.024	0.114
	Ilves	0.053	0.115
	Laurila	0.036	0.117
	Metsälä	0.047	0.167
	Maijala	0.134	0.214
	Kontio	0.154	0.231
	Talvio	0.165	0.232
	Kalervo	0.202	0.272

The location of case schools in the inefficiency distribution is quite stable between the two models (see Figure 1). With some exceptions, these models also differentiate the schools into opposite tails of the distribution. Kontio, Kalervo, Maijala and Talvio are among the most inefficient schools. In order to be efficient, these schools should have produced a clearly larger output according to results of both models. Metsälä School is in the middle of the distribution in both models and Annala, Ennala and Laurila schools are among the most efficient schools (i.e. have low inefficiency). Ilves school somewhat changes its rankings in these two models.

Figure 1. *Inefficiency distribution and the location of case schools in the distribution in random and fixed effects models in 2000-2004. Rankings based on inefficiency scores are in parentheses*



Other characteristics of the case schools

Case schools differed from each other in terms of the family background of the students, the comprehensive school GPA, the heterogeneity of the student body, the proportion of female students, and the matriculation examination results. In Figure 2, each school is depicted in terms of these factors as a percentage deviation from the corresponding national average in 2000.[14] These national averages are depicted in the figure by the dotted circles in bold located at the 0%. Circles outside describe performance above the national average (favorable situation for the school) and circles inside performance below the national average. The scales of percentage single parents (*Single parents (%)*) and heterogeneity (*Heterogeneity*) are reversed

[14] In an earlier version of this paper (Kirjavainen, 2008) the dotted circle was erroneously drawn to the 10 per cent circle. The correction of this error caused some modification of the text.

indicating that they usually have the opposite influence on student performance as the other indicators in the figure. Inversing these two indicators provide an opportunity to interpret the figure so that performance outside the dotted circle (national average) is favourable for the school and performance inside the circle is infavourable for the school.

Considering the high efficiency schools, Annala general upper secondary school was located in the Helsinki area in a wealthy and pleasant neighbourhood. It could be characterized as an average Finnish general upper secondary school in only one respect, namely the share of female students was close to national average. As for other factors, it scored clearly above the average in the matriculation examination results, the intake and the family background of students was clearly favourable and the heterogeneity of the student body was lower than on average. It was larger than an average school. The premises were quite old and unpractical and they needed renovation. There was also upper level of comprehensive school (grades 7-9)[15] in the same building. Interviewees were quite satisfied with the location and surroundings of the school.

Ennala general upper secondary school was quite a small school in a small town in southern Finland. Many of its students studied there because it was the nearest general upper secondary school in the area. It was a school with below the national average entrance requirements. Heterogeneity of the student body was also somewhat higher than on average. As for the other measures, they were close to national average except for parents' education which was somewhat higher than the national average. The school was located in a quiet neighbourhood and the premises had recently been partly renovated and extended. The school shared the premises with the upper level of comprehensive school.

[15] In Finland, the comprehensive school provides education for the grades 1 to 9. The education is divided into lower level (grades 1-6) and upper level (grades 7-9). The premises can be joint or separate. In some cases the upper level education is provided in the same premises with general upper secondary school.

Ilves general upper secondary school was close to the average Finnish general upper secondary school. It was located in the surroundings of Helsinki in an area with a high educational level. It attracted students from nearby upper levels of comprehensive schools and had entrance requirement close to national average. Quite a high proportion of the students also came from its own upper level of comprehensive school. The nice neighbourhood was reflected in the background of students, which was more favourable than on average. The results in the matriculation examination had normally been on the national average level or higher. Heterogeneity of the student body was larger than on average. Ilves School was quite large in size.

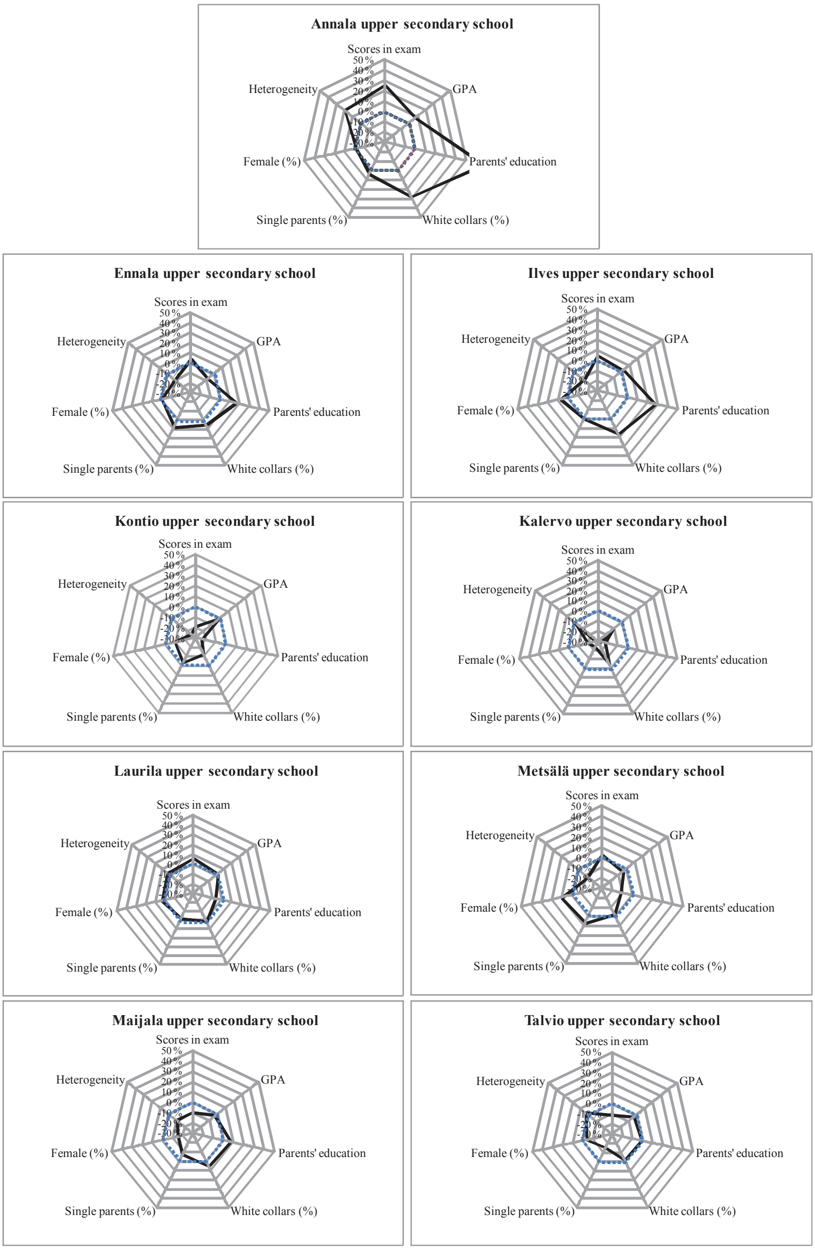
Laurila and Metsälä general upper secondary schools had a very similar profile in the sense that they performed close to the national average in many of the indicators. The performance in matriculation examination and the intake of schools were close to national average. However, parents' education was less favourable. Laurila was a fairly large school in a small town in southern Finland and students came from different upper schools. The school premises were quite old and also somewhat too small and renovation and extension was therefore taking place. Metsälä general upper secondary school was located in a rural area in southern Finland. It was a small school accepting basically all its applicants. The upper level of comprehensive school was also located in the same premises and the majority of the teachers had instruction in both schools. Many of the teachers enjoyed working in these schools because of the location and the surroundings.

As for the low efficiency schools, Kontio and Kalervo general upper secondary schools had the most distinctive profile among the case schools. Kontio general upper secondary school was located outside Helsinki in a small town and provided general upper secondary schooling for this area. Many of its teachers mentioned that they prefered to live in a small town and therefore they chose to work in this school. The students' socioeconomic background was lower and the heterogeneity of the student body was higher than on average in this school. Results in matriculation examination had been clearly below the national average. The size of the school was close to the national average. The school building was old but had recently been renovated.

Kalervo general upper secondary school was a small school in the Helsinki area in a neighbourhood with a somewhat lower socioeconomic status. Its profile was very close to that of Kontio general upper secondary school with a low comprehensive school GPA and a low socio-economic status among the students. The heterogeneity of the student body was close to the national average, reflecting the fact that there were very few high achieving students in the school. The school did not have upper level of comprehensive school in the same premises and the building was quite old. Nevertheless, interviewees felt comfortable in this school because of its cosy atmosphere.

Maijala and Talvio general upper secondary schools were located in the Helsinki area and both were quite large schools. They also shared the premises with an upper level of comprehensive school. Therefore, some of the upper school students continued their studies in the same general upper secondary school. The entrance requirements were close to national average in both schools. Despite that respondents felt that the intake was unfavourable. Concerning the results in the matriculation examination, they had been clearly below the national average. In Maijala School the socioeconomic background of the students was higher than on average whereas the heterogeneity of the student body was lower than on average. In Talvio School they were close to the national average.

Figure 2. *The score in the matriculation examination, GPA, socioeconomic background and heterogeneity of the student body in case schools in 2000*



Note: The scales of *Heterogeneity* and *Single parents (%)* are reversed.

Interview material

The analysis of case schools is based on interview data.[16] The interviews were conducted between 10 May and 16 June 1999. Therefore, they reflect the school practices taking place when students completing their studies in 2000-2002 were studying. Since schools do not change unless something dramatic happens, students completing their studies in 2003 and 2004 most probably faced a similar environment to the three earlier cohorts.

In each school the principal and two teachers were interviewed. The first contact with the school was made with the principal by phone. On the phone, I explained the research project and its aims to the principal and asked for consent to include the school as one of the case schools. The first appointment was also made with the principal. Teachers were selected by asking the principals to name two possible teacher candidates for the interviews. Teachers teaching subjects included in the matriculation examination and with a longer career in the school were given priority. Teachers with a longer career in the school were preferred, since the initial purpose was to use the data from an earlier efficiency measurement. None of the principals or the teachers turned down my request.

The method of selecting the participating teachers was not random. Principals clearly did not name their strongest opponents as candidates. The opposite was more likely to be the case. Thus, teachers may have given a somewhat more favorable view of the school than a randomized sample. In the same way, if some problems in the school were mentioned, they were probably real and not just the complaints of an unhappy teacher. Despite this, the interviewees were horizontally comparable, since the method of selection was the same in all schools.

Another problem created by the research setting was the contrasting of efficient and inefficient schools. Even though the interviewees were not told in which category their school was placed, they may

[16] The interview data was supplemented by school annual reports and school curricula from some of the schools. This material is not, however, analysed in this study.

have given a more positive view of their school compared to the real situation. The concept of efficiency is, however, not very intuitive and school results are most of the time measured by the success in matriculation examination. Therefore, the interviewees were probably most of the time thinking of this measure rather than efficiency.

As for the characterization of interviewees, the majority of the nine principals were male (7), and only two were female. By contrast, 11 of the teachers were females and seven males. As for professional experience, the average employment of teachers in their current school was quite high, being almost 19 years and ranging from 3 to 33 years. Among the principals it was considerably lower, on average 10 years, ranging from one year to 33 years. It should be noted that three of the interviewed principals had been principals in their present school for only one year.

Since the selection of teachers was based on the proposals of principals, I had no prior knowledge of how different subjects would be represented. It turned out that eight teachers taught humanities such as history, biology, geography, religion, and psychology. A foreign language was taught by seven of the teachers. Three of the teachers taught mathematics, physics or chemistry. None of the interviewees taught the mother tongue, Finnish. As for the principals, five had a background as humanities teachers, two in mathematical subjects, one in foreign languages and one in the mother tongue. There was no systematic differences in interviewees background that would be related to efficiency scores of the schools.

As for access to schools, it was surprisingly easy to make appointments. I think that one of the reasons why schools were so willing to participate and cooperate was that the information obtained from the interviews was promised to be kept anonymous. In addition, since the results of earlier efficiency analyses were presented so that individual schools were not identified (even though it would have been possible), it probably helped to gain the trust of the participating schools, despite the research topic itself being sensitive from schools' point of view.

The interviews followed a semi-structured schedule such that the interview guide contained several themes and questions related to each theme. The themes were clearly introduced during the interview and the main questions were always posed for the interviewee. When discussing questions the conversation proceeded loosely, giving each interviewee the freedom to speak without the interruption of interviewer. Even though the interview guide was mainly followed in each interview, exceptions were also made if the interviewee introduced an interesting topic.

The themes discussed during the interviews were mainly the same as in the earlier survey (see Kirjavainen, 1998 and 1999). They were selected based on findings of school effectiveness studies (see e.g. Purkey and Smith, 1993; Levine & Lezotte; Scheerens and Bosker, 1997). The interview guides included questions concerning the students of the school, the school's physical facilities, the role of the parents, and factors related to the school's management such as the goals, teacher participation in school planning and decision making, the possibilities of teachers to influence school decision making and their own work, as well as questions on job training, evaluation practices, cooperation between the teachers, and student counselling. In their responses, the interviewees were asked to describe the current situation in their school and express their own views about it. There were also questions concerning major changes in the past years.

The interview guide used with the principals differed from that of teachers mainly in the sense that it also included some questions concerning the teachers. The similarity of principals' and teachers' interview guide enabled the comparison and validation of the responses, and the identification of areas of disagreement and difference.

When contacting the candidates to set up the appointments I also asked if they would like to see the interview guide beforehand and get to know the interview themes. About half of the candidates became acquinted with the interview guide in advance. Some of the respondents had also prepared some notes for themselves.

The interviews took place in each of the schools and each respondent was interviewed separately. They lasted from approximately one to two and a half hours. The average length was an hour and a half. Each interview was tape recorded with the consent of the interviewee. At the beginning of each interview the normal confidentiality of the information obtained through interviews was stressed.

The climate during the interviews was good. Some of the respondents were somewhat distant, probably because of the generational gap and position between the interviewer and the interviewee. However, some respondents were very open.

The interviews were transcribed word for word. For the purposes of the analysis it was not necessary to transcribe the taped interviews verbatim. As for the quality of the tapes, they were mostly good. A certain amount of editing took place so that all the pauses, laughs, or expletives were left out. The language used by the interviewees was not, however, transformed into written text. Only the selected passages were translated into English using more formal language.

Analysis of the interviews

The interview data were interpreted as describing the realities of the school. They were, in other words, "treated as giving direct access to experience" (Silverman, 2000). The descriptions of the interviewees were taken as presenting facts that are encountered in these schools. When interpreting the data, however, one realizes that by spending time in the school and collecting more diverse data it would be possible to perform a more in-depth analysis of the schools. However, this was not possible in this study.

The analysis of the interview data started with the re-reading of the transcribed texts. The aim was to gain an understanding of the different themes and views the respondents brought up in the interviews. The responses of the principals and teachers in each school were also cross-checked and most of the time teachers and principals expressed similar views. During this process the texts were grouped into themes.[17] In the next phase, each theme was studied and coded separately. During this phase, special attention was paid to the differences between schools within each theme.

Several themes emerged from the texts. Within some of the themes, views and practices of the schools were so similar that it was not possible distinguish different categories and group the schools based on them. These themes were concerned with staff development, evaluation and monitoring practices, and parent-school relations. Four of the themes were such that it was possible to identify different categories within them and group the schools. They were:

- views about the students
- staff relations
- school governance and management
- views about the curriculum and curriculum planning

Each of the categories had their own characterizations. In some cases the differences between the categories were small. Each category within each theme is mutually exclusive. A school is placed into that category based on the views expressed by the interviewees of that school. The placing is based on the majority of the views i.e. in these schools these characterizations came more often up during the interviews.

The results of analysis are presented so that each theme and its categories are discussed in turn, stressing the most distinctive features. In presenting each of the themes and their categories, interview extracts are used to demonstrate the empirical basis for the categories and to verify the views expressed by the respondents. The quotes have been translated from Finnish. Pseudo identifiers referring only to schools are used for the quotes in order to protect the anonymity of the respondents and the schools. These identifiers also show how each school is placed into these categories. After presenting the different categories, their relationship to efficiency is examined with category averages and simple cross-tabulations.

[17] The analysis of the texts was carried out using QSR N6 and VIVO 7 computer software designed for qualitative research analysis.

Finally, it should be noted that no attempt is made to draw any causal conclusions or identify mechanisms that produce inefficiency in this study. For such purposes a more extensive study using longitudinal data would be needed.

4. Views concerning the students

The characteristics of the students are a central part of the school. They mostly determine the work of the teachers and affect the school's reputation. The role of the teachers and schools is usually smaller. Student characteristics also affect teachers' decisions to apply or not to apply to certain schools. It is an established fact in the literature that schools in deprived areas and with students having difficulties have problems in recruiting highly qualified teachers. Staff turnover is also usually higher in such areas.

Student characteristics affect the way teachers perceive their work. However, teachers also differ in how they react to certain characteristics of students. In some sense the views of the teaching staff and personal traits of the students intertwine so that it is difficult to distinguish them from each other. At some point, however, they develop into views that perhaps characterize the whole school rather than being just the views of individuals. The differences in views are more evident when the background of students is kept constant.

Four categories emerged when respondents' descriptions and characterizations of the students were analysed. The teachers' stance or view could be characterized as <u>attentive</u> and sensitive, <u>trustful</u> and confident, <u>educating</u>, or <u>frustrated</u>. Each category has a characterization of its own so that the behaviour of students and the stance of respondents towards it can be seen as interacting.

Attentive and sensitive view

The attentive and sensitive view emphasizes that every student is important and should be taken care of regardless of his or her personal motivation or skills. Some respondents made a comparison with parenthood by considering that part of their role as a teacher included being a mother or a father figure for the students. Attentiveness and sensitiveness also involved an internalization of the school's social task to provide general upper secondary education for everyone, as one respondent comments:

> *"As we are a small school in a rural neighbourhood we have aimed to provide opportunities for those wanting them." [Metsälä]*

The composition of the student body did not differ from the average and there were both high and low performing students with differing motivation to study in an general upper secondary school. Most students had their minds set so that "they had decided that they would pass the matriculation examination or their parents had decided that their child would pass the matriculation examination." [Metsälä] However, there were also students who were not so sure about their goals. These latter students were considered as the ones "with problems" [Metsälä]. Even though the motivation of students varied, the overall situation was seen as quite good because "most of the students have quite good motivation to study when they start school." [Metsälä]

The view towards the least able students was perhaps the most distinctive feature. Students with low motivation were taken as quite a natural part of school life. Instead of seeing them as a burden, an encouring and soliciting stance towards them was taken.

> *"But I guess these things happen in every school. There are the ones that fall behind, especially if you have a low GPA. They realize that they don't necessarily succeed. But then we have tried to encourage them, both teachers and the principal have tried to encourage them." [Metsälä]*

> *"[...][...] I try to keep everyone involved and encourage and try not to put students down, even if they don't know. [...] [...][Ennala]*

Weaker students also required special attention from teachers and the role of 'care takers' was taken quite seriously. Respondents emphasized that "students with a poorer grade point average were also accepted and they have also matriculated". This meant that teachers had then "worked very hard with these students and succeeded". [Metsälä]

There was also some sort of pride when respondents talked about the weaker students and how hard they worked with them and succeeded in preparing them for the matriculation examination. "It is not always a very easy task", though. [Ennala]

> *"In practice, we have to turn the weaker students into matriculating students. And in that, I claim, we are really good. None of our students have failed the matriculation examination." [Metsälä]*

Teachers also felt that weaker students brought a differing aspect to their work. "[…][…] Sometimes I think that, from the teachers' work point of view, the job would be different without the weak ones to be taken care of." [Ennala] The hard work with the weaker students was also considered rewarding.

Although the respondents emphasized the weaker students in the interviews, good students were also mentioned with respect and care.

> *"The good students are really great and wonderful and you also have to take care that they are given demanding tasks and proceed normally." [Ennala]*

The provision of remedial courses was considered important and a part of the process of "also making the weaker students matriculate". Teachers felt that students took these courses quite seriously and almost all students participated in them in most the important subjects at the beginning of their studies.

> *"Yes. Very much. We have this entry-level test and then we recommend that everyone should participate in a remedial course, even the good students. And they do." [Ennala]*

Some were concerned whether these courses were enough. “Sometimes I have the feeling that despite the courses, the weaker students don’t learn. Because they have these gaps in knowledge, it is really difficult to fill them.” […][…] [Ennala]

Trustful and confident view

Trust or confidence perhaps best describes the views that teachers and principals expressed towards the students with high motivation and a high comprehensive school grade point average. Accordingly, students were treated more as grown-ups who were responsible and able to take care of their studies. They were also described as “conscious youngsters whose goals were mostly clear and high set”. In addition, it was not only students that had “ambitious minds”, but this behavior had its roots in their families. Parents “often had quite a high educational level with university degrees” and “the families themselves provided encouragement to study”. This was a fact that was seen to “affect the way students approached studying” [Ilves]. Teachers also appreciated this fact.

Professionalism and formal roles were stressed when teachers characterized their roles. Accordingly, they were “subject teachers”, “student counsellors” or “homeroom teachers”, or even “collegues to other teachers”. While the teachers in attentive and sensitive schools also mentioned their roles as mother and father figures, here these roles were not present. If teachers stressed their ability to understand or be concerned with the lives and sorrows of the students, it was because of their professional background and not because of their personal traits.

Principals characterized themselves and their role by a feeling that they were the “first and last caretakers” in the house, meaning that they were ultimately responsible for the school. They also were quite task oriented in describing their various roles. They mentioned being a pedagogical leader, developer of the school, leader of the personnel, teacher, the one responsible for the finances, a care taker of social relations in the neighbouring community, and the marketing person of the school.

The attitude towards the skill and knowledge differences among students in this setting was not very concerned. The differences were not seen as a problem, but were recognized as being caused by the varying grading practices of comprehensive schools. Teachers also commented that they "learn to know what a nine from a particular school actually is" [Annala]. Apparently, the differences were quite small and did not affect the classroom work of teachers, since none of them mentioned that they somehow adjusted the pace in teaching. Teachers only mentioned that as a consequence of skill differences some of the students start getting lower marks.

> *"Some of those who have been getting tens may continue getting tens, but not so much anymore. And then it is also possible that grades sink to four, five, or six. I mean, in the end there are some differences in what they have learned." [Annala]*

The way teachers motivated students was quite light in nature. As teachers respected the independence of the students, motivation was mainly based on emphasizing the students' own responsibility. Therefore, it was rather a matter of reminding students that they should study or should start studying and it did not involve any discussion about encouragement or concern. This view was partly due to the fact that only a small group of students had to be motivated or reminded every now and then.

> *"[...][...] naturally there are some students who have to be motivated and even I motivate some of the students by saying should you not work now." [Annala]*

The same view emphasizing personal responsibility was applied to final year students with low motivation. Their own initiative and responsibility rather than teachers work was stressed when teachers described that "a few months before the matriculation examination they [students] start studying so that they will pass the examination. And some of the students even graduate with good marks." [Annala]

Students loosing their motivation to study were seen as youngsters searching their identities. In this sense there was little concern about the well-being of the students involved. It could also be argued that the role of the school was externalized and that the students' own views were emphasized and respected.

> *"But there is always someone who has been accepted with a high GPA and who has wanted to study in this school, but then suddenly starts to reconsider during secondary school whether this is what she really wants. It can be considered as searching for the self when being young." [Annala]*

In some cases the stance towards the low performing and low-motivated students could also be characterized as quite harsh. They were considered strictly as "unbelievable loosers" and persons "that cannot be kept under discipline because they are individuals and they work at their own pace and they don't want to hurry into life" [Ilves].

Because most students were high achievers, they did not need any remedial instruction at the beginning of their studies. These courses were offered in the most important subjects (such as mathematics and Swedish) during the first year "but there weren't very many youngsters attending these courses". They thought that "they were good enough, so that it was not necessary for them to attend." [Annala]

Teachers emphasized and took seriously the demanding nature of high achieving students. Such students were also considered as motivating. Therefore, schools carefully thought over how they could serve the students best. "Students know what they are worth and that places certain demands on the instruction" [Annala], meaning that, for instance, the selection of new teachers has to be made carefully. In addition, teachers had to pay extra attention to their preparation of instruction and lessons.

> *"You can't go in front of the class empty-handed. You have to have a certain amount of knowledge and you also have to be able to pass on this knowledge to the students." [Annala]*

Despite the high requirements, the work itself and positive attitude of the students towards studying was so rewarding that “teachers didn’t want to teach in any other school, just because of students...” [Annala]. Teachers also pointed out that they found their job easy in a sense “because they don’t have to make any extra effort to create a positive learning environment. Instead, it is already there when they enter the classroom.” [Annala]

Educating view

The educating view involved a somewhat distant attitude and in some cases also frustrated views about the students and their behaviour. Students were more or less average. There were both high and low performing students and the family backgrounds of students also varied. Teachers perceived their role in a task-related manner and very little else was considered as being part of it. Principals were “running the show” and felt they were “responsible for everything”.

Respondents felt that the motivation of the students was “quite good” [Talvio]. There were students with “clear motivation and with a good study technique” [Talvio], but with some of the students, it “could have also been better” [Talvio]. At the other extreme, there were “students who don’t do anything” [Talvio]. This diversity was considered by some of the interviewees as one of the strengths and facilitated the work. The role of the higher achieving students was to draw the weaker students into achieving better results. Strongly performing and motivated students were, in other words, used as a teaching aid that also motivated weaker students.

> *“What I consider as positive is that there are both students with a high GPA and with a low GPA. […][…] Such a school is an easy school. There the weaker students can also hang on the coat-tails of the better students.” [Talvio]*

The stance towards the students with low motivation was somewhat divided. In some sense, the respondents felt that it was natural and acceptable that “there are some students who start studying (here) because of friends and because they don’t know what else to do” [Talvio] or because they, as a whole, “wonder what to do with the life” [Laurila]. These students were also seen as the “ones that are here for child minding” [Laurila], which was also considered an “acceptable reason” [Laurila]. This view did not, however, include any emphasis on the students’ own responsibility or caring by the teachers.

Besides this, there was also a tougher view according to which some of the students “shouldn’t be in general upper secondary school in the first place” [Laurila]. Some interviewees characterized students proceeding with a slower speed as decelerators. The somewhat negative emphasis of the term can be sensed from the descriptions of one teacher.

> *”[…][…] Then there are weaker students who in the sense opt out to join the group of decelerators. And these decelerators have a high status, she’s a ‘decelerator’. Earlier, they were ‘repeaters’, and that was negative. But a decelerator, who has her own curriculum, has a higher status. And there are students who want to be in that group, even though they have the resources to complete their schooling in three years.” [Laurila]*

A disciplinary approach was also mentioned by some of the respondents. This especially applied to first year students who “come into the classroom unprepared or forget to bring their textbooks”. Teachers felt that “the first year students can be surprisingly carefree.” [Laurila] Motivating in this environment was considered a difficult task.

> *“[…][…] In other words, it is a big question how we can make this whole thing work. And we have been thinking about it a lot, but have not yet found any philosopher’s stone for how it would work better. […][…]” [Laurila]*

Concerning instructional methods, teachers felt that they were unable to use them freely because some methods “would just create chaos” [Talvio]. As for the planning of the instruction and progression in the classroom, the strategies were mixed. Some teachers emphasized that they mostly proceeded in terms of the more motivated students and didn't take that much account of the weaker students. Some teachers even purposefully dropped students who were not interested in their subject.

> *"I don't actually plan it in a different way. We proceed in terms of the best students. During the lessons, of course, to get the attention of the ones who are not interested is a challenge. One should use some tricks. I personally try to use humor so that I don't run anybody down too much." [Laurila]*

> *"I drop the students who are not interested in this subject so that they really don't take this subject. That way I have quite good students in the end. In other words, the group is then smaller but they are also quite interested. […][…]" [Talvio]*

Sometimes, a softer strategy was applied so that only the compulsory topics were covered during the courses. The content of the courses was also facilitated so that everyone could participate and find it interesting.

> *"[…][…] This year I only taught the compulsory topics. […] I have chosen to teach general things to those ones that are not so interested. I give lectures on important persons in this field. In that way it's somewhat easier. I also usually start from the very basics when I start teaching advanced special studies." [Talvio]*

Some measures were taken to improve the skills of the weakest students. Schools offered some remedial courses in some of the subjects. Some of the differences in skills were also due to 'quality' differences between the upper schools at the comprehensive level.

> *"We have noticed that in different upper schools [comprehensive] they give somewhat different grades. The grading practices differ. And naturally that's not fair for the students. […] […]" [Laurila]*

Frustrated view

The greater the proportion of low-performing students, the more there was frustration and concern in the voices of the interviewees. The common denominator was the unhappy and in some cases clearly frustrated views concerning the students and especially the low-performing students. A distintictive feature compared to other categories was the presence of a slightly higher proportion of low-performing students and only a few top-performing students. The heterogeneity of the performance was also high. These factors made teachers' work more demanding and also influenced their views.

Teachers had a very diverse view of their role in this category. Some of them emphasized very strongly their role as a teacher. Some teachers also stressed the administrative roles that they were involved with. For some they provided recreation for the demanding/exhausting teaching. Men also brought up the idea of being a "disciplinarian" or a "father" where male teachers were in a minority in the school. Especially the older teachers felt they were like "grandmothers or grandfathers", taking their responsibilities in a somewhat more flexible manner. Teachers with a longer career felt they represented "authority" and "status". In this group, as with other groups, principals emphasized their role as a person being responsible for the whole school.

Some of the respondents used the term underperformance when describing students' motivation and attitude towards schoolwork. It referred to the fact that many students were happy when they achieved the minimum requirements. The teachers had the opinion that many of the students could do better if they worked harder.

> *"Quite a lot of them are under-performing. They are satisfied if they just graduate, even though they could do much more than that." [Kalervo]*

This low-performing group of students was considered to be in the school "for child-minding" and "they had to be hustled" [Kontio]. The frustration was also clearly spelled out by some teachers who felt that almost every student in their classes was "more or less just passing time". According to them, there were "normally only a few students in a class who could be characterized as normal." [Maijala]

> *"Their skills and knowledge are pretty weak. Take, for example, the answering in an essay form. They do not answer the question. Or it is difficult to figure out what question they are answering. Their skills are mainly weak. It takes basically the first year to teach them how, for example, to answer the questions. It's something one would expect them to learn at comprehensive school. Sometimes I feel like I'm shepherding them. If I look at someone and see that he can hardly write, I start wondering if he's earlier only been playing with the books. And the content is also missing." [Kalervo]*

In addition, for some teachers the low skills of some of the students invoked emotional wonder, disappointment and concern. One teacher asked "what are they basically doing at comprehensive school?" [Maijala], since many students had major deficiencies in their basic skills. Some teachers became emotional when talking about the students.

> *"Sometimes when I have had a maths lesson, I'm very much affected, when I realize that they don't even know this thing. We don't have such easy courses in maths or Swedish that the weakest students could do." [Kalervo]*

Alongside the frustration and compassion, some respondents expressed an attentive view towards the weaker students, emphasizing that they "have to be especially taken care of". Teachers felt they "must be more like mothers here than in some other kind of general upper secondary school". This meant that they "have to take care of practical matters and make sure that they don't get fours or that they complete their courses in time." [Kontio]

Because of the students with low motivation, teachers perceived their work to be quite hard and demanding. They did not expect students to study independently and acquire the knowledge from elsewhere. They emphasized that "here, the teacher has to repeat until the student really grasps it" [Maijala]. The choice of teaching methods was also problematic and teachers had to be prepared to change their style according to the situation. They found this very demanding. The teaching also caused some disappointment because certain teaching methods could not be used.

"I teach history and philosophy. I'll comment first on the history. It is [the student body] very heterogeneous. And just that is the biggest problem in pedagogy. You have to think it over and change the style and the way of teaching all the time. [...][...] It influences the teaching methods in the sense that I haven't been able to get a feeling of success when using interactive teaching methods, even though I always try them once or twice during the course. [...][...]" [Maijala]

Lack of strongly performing students was also considered a problem. Since students are always comparing themselves to their classmates, in a class where hardly anyone is taking schoolwork seriously, there were no good role models to raise the standards. "When they look at the student next to them and see that she can't do better than they do and that she doesn't work any harder than they do, they will not work harder." [Kalervo] For this reason, teachers felt that "it is clearly easier to work with groups having a couple of hard-working and good students." [Kalervo]

"[...][...]And I think that such an example also teaches the weaker ones. And it keeps up with positive learning atmosphere. If all (students) are quite weak ones and in a sense reluctant to learn, it follows that the whole climate in the classroom is unpleasant. And it feels like you would be pulling a sledge full of stones, as is usually said. But even two or three able and positive students will eventually draw the weaker ones along." [Kontio]

Weak skills were taken seriously by devoting some extra resources to remedial courses at the beginning of the general upper secondary school. Teachers also described that they devoted extra work to get students to the required level, especially during the first year. In many subjects, all students are required to participate in these courses. Good study skills also have to be taught to some of the students.

"Quite a lot of time is spent on remedial instruction and therefore we are not able to proceed as far as one would hope. Then the students have another problem, related to their working habits. They think that, for instance, by only attending the course one learns Swedish. Just like the students do in comprehensive school. But naturally that's not how it works." [Kalervo]

5. Staff relations

Collegial and good staff relations are mentioned as one of the key characteristics of effective schools (Sammons, 1999; Teddlie and Reynolds, 2000). This involves good working relations among staff members, sharing good practices, exchanging ideas, supporting each other, observing each other and giving feedback, and working together to improve the teaching programme.

In this study, the focus in staff relations was on staff members' descriptions of the atmosphere, discussion culture, cooperation, and respect for each other. Based on the descriptions of staff relations, three quite traditional categories emerged in this study. In most of the case schools, relations were <u>professional</u> and in some schools they were characterized as <u>friendly</u>. There were also schools that had <u>tense</u> relations among staff. The difference between professional and friendly relations was small. Most of the characteristics were present in both categories, but with a higher degree in friendly relations.

Professional relations

Typical for professional staff relations were respondents' comments on staff relations such as "quite relaxed and unreserved" [Annala] as well as "open" [Laurila]. There was also some formality and mixed feelings involved when they were described as "not so warm, but rather warm than on the minus side" [Annala]. Despite these differences, many respondents stressed that there was "a good climate that could also be sensed in the staff room."

Professionalism manifested itself when it was emphasized that the focus was clearly on school work and "there was this sense of doing and working in the air" [Annala]. People were not so hard and fast about different things. There were certain rules and things were usually performed according to schedule. The relations between staff members were also good in the sense that there were no cliques.

> *"In my view it's very pleasant. There are no cliques. Nothing like I have heard these unbelievable stories that people sit together and some of them haven't even talked for a year. And brawl and quarrel. We don't have anything like that. Our climate is very nice." [Talvio]*

The independence of teachers in their work was emphasized and that there had always been "the freedom of doing". Collegues and their work were also respected so that everyone "was able to do their field of work without someone breathe down their neck". [Ilves] This included the management of the school. Teachers felt that they were "trusted to take care of their own work". [Ilves] Some teachers also mentioned a tolerant climate in the school, which also extended to allowing impetuous opinions and discussion.

> *"I have been happy in this school, I would say very happy. I have felt this is a good and pleasant working place and work community. The students are nice. In principle, the work community is very nice and the climate in the staff room, even though it is large, is nice. There are always some tensions when there are many people, but they have been solved. The general climate is, however, nice. Then I have also appreciated that the principal doesn't intrude in everything but we have very high autonomy." [Laurila]*

Good and open relations among the staff members also created job satisfaction. All the interviewees told that they were happy in their current situation. Some of the teachers were actually very satisfied and many mentioned the good climate in the staff room and among the collegues as one of the reasons for their satisfaction. The professionalism of their collegues was also appreciated by most of the respondents.

The discussion culture was open. Teachers had always engaged in discussion about instructional matters in these schools. The school climate had been supportive and open for them. The discussions were somewhat dependent on the personal chemistry of people. However, for instance, the language teachers very vividly discussed the instructional methods and matters, such as "how you do this thing or I did it like that but it didn't work out" [Annala] or "changing ideas of what to try out and how to do things and how they work out" [Ilves]. In subject group meetings there was also lively discussion [Laurila].

> *"These kinds of discussions have always been unofficially going on in the staff room and after school. And more so that different subject groups discuss objectives and problems. Let's take as an example that some student has many problems in studies and she's not completing courses. Teachers discuss such cases. Equally, if some student is often absent from school. Such problems can be discussed." [Annala]*

Respect for independence also had some drawbacks. Some respondents felt that even though the discussion culture was open, certain topics were somewhat avoided. Discussions concerning instructional methods and the content of subjects and instruction were especially needed. In this area, teachers felt that they were alone and that there was a lack of support.

> *"The discussions are not very general or subject related concerning instructional methods or subject content. Mostly it's concerned about the students, so that teachers think over what to do with this student who hasn't attended lessons and hasn't completed her work. Very practical. And also important. So I think it remains at that stage." [Talvio]*

> *"[…][…]We somehow support each other. Maybe in this social side. But not enough in the content of instruction. Inside the subject I think each teacher is quite alone. It naturally depends on the subject. But I think that everyone is perhaps doing too much in their own way." [Talvio]*

Despite the good overall situation there were also areas that could be improved. Some respondents commented that some of the teachers were more active and some wanted to stand aside. But even there, the teacher's own responsibility was emphasized and it was found more or less useless to try to force teachers to cooperate more.

> *"In some subjects the teachers are active. But I think it's a personal attribute. We also have subjects that don't have any cooperation. They don't discuss with each other, they don't share materials and tips or anything. And that's naturally a very inconvenient situation. And if it's not possible to change, there is no use forcing anyone into it." [Talvio]*

Occasionally, emerging disputes and disagreements between some of the staff members were considered quite normal. Some balancing between different subjects and subject groups also sometimes took place. However, these problems were minor. If disagreements emerged between certain people they were usually dealt within the staff room with understanding. They did not cause long-lasting problems or distress.

> *"Sometimes people get nervous and discharge their feelings very fiercely, but it doesn't cause any catastrophe. There are no cliques so that one teacher belongs to one and one to the other. Instead, if someone is feeling bad, others try to understand that she was maybe tired or something."[Laurila]*

Teachers' independence and responsibility for their own work was also reflected in problem situations. Normally, teachers coped with and settled these problems alone. But it was also emphasized that the climate was open to discussions and teachers openly exchanged information on, for instance, problems with students. It was also quite common to share difficult situations with other colleagues in the staff room.

> *"[...][...] But we also have a sort of openness so that you can hear teachers talking with each other that damn, I had such and such an experience." [Annala]*

Some mostly quite difficult situations were also discussed with the principal. In such cases teachers did not feel that their credibility would be affected in front of the students. In fact, no such theme emerged during the interviews in these schools. Nevertheless, some concerns were expressed about the culture of teachers solving the problems mostly by themselves.

> *"It's also a matter I have been thinking of a lot and in my view there is a lot to improve. In general I think that teachers are quite alone. If a teacher opens up to some of her colleagues, she doesn't necessarily get sympathy and understanding. And especially young teachers may feel alone. [Annala]*

Because of the good interpersonal relations, teachers had many enjoyable after-school activities together such as trips, dinners and parties. The climate was usually nice and relaxed, even though in some schools some teachers also mentioned the slightly formal climate by saying "that the atmosphere wasn't exactly very easy going". Often, groups of teachers also spent time together outside school hours.

Friendly relations

Friendly staff relations occurred in schools where respondents characterized them as "very good" and "relaxed". The situation in the schools in this category was somewhat more open than in schools with professional relations. Everyone praised the climate, "even the substitute teachers coming from outside". The cooperation between staff members was also described as very good. Interviewees were satisfied with the situation and commented that there was "nothing to complain about".

> *"The working climate and relations among staff are very good… Colleagues are easy going. There are no controversies or anything else. We have a nice climate. The visitors have also enjoyed staying with us. I don't have the feeling that I somehow hate to come here and that again I have to see some odd persons. I have no such concerns." [Kalervo]*

> *"It's nice to be here. There is no need to be nervous in the staff room or elsewhere. I think we are enjoying it here." [Kalervo]*

Teachers were clearly more unanimous and the climate for discussion was open. In the interviews, there was a sense that teachers were in the same boat, so that they were doing their work together. Teachers were unanimious about the less motivated students causing problems and therefore problems related to instruction, for instance, were considered natural.

> *"Teachers can freely discuss the problems here. One doesn't have to brood over them by oneself. We all have problems because students don't learn. And we all complain there (in staff room) together. Thank god, it's not so that everyone should do one's stuff and show that everything is under control. And to say that in my class no one behaves like that." [Kalervo]*

A relaxed and good climate was one of the key factors that created job satisfaction in this group. Teachers felt that they were part of the group and not outsiders in any way. So, for them the satisfaction had other sources in addition to the good climate such as the possibility to develop their professional skills, professional freedom and challenging tasks.

The respect for each other was high in these schools. The discussion culture was also open and problems with both students as well as instructional and subject matters were discussed and developed together. Similarly to schools with professional relations, these factors were mostly considered in smaller groups within the subject or subject group.

> *"Teachers of related subjects discuss matters quite a lot. Mathematics teachers have discussions, then language teachers. They have cooperation. They discuss instructional methods and related things." [Ennala]*

Disagreements seem not to cause any disturbing arguments. Differing opinions were considered as natural and they were tolerated as long as they did not harm the individual freedom to work.

> *"There are differences in courses of action. But one can also disagree. Some teachers are much stricter than others and some teachers feel that everyone should follow the same routines. But I don't think that's the case. And that's also accepted, as long as it doesn't disturb the work of someone else." [Kalervo]*

Teachers' collegial relations were also manifested in after-school activities, which seemed to be most plentiful among the case schools. Teachers had parties, trips and other recreational activities together. They were also content with the situation.

> *"They meet a lot. They practice sports together and they travel together." [Ennala]*

> *"We arrange joint gatherings and trips. Tomorrow, for example, we are going on a field trip. So it's not formal by any means. Then we*

arrange Christmas parties at someone's place etc. and also other related activities." [Kalervo,b]

"[…][…] The yearly planning meeting, for example, is always arranged in some nice place so that the environment would be nice. I think they all have been very successful. And when it comes to other things, they have been really nice. We go and have dinner together after some event and things like that." [Ennala]

Tense relations

In schools with tense staff relations the interviewees openly admitted that the situation "could be better" [Kontio] and that there were "obvious conflicts between some of the teachers" [Kontio]. In some cases the descriptions and charges were even more severe, so that the relations and the atmosphere in the staff room were described as "unpleasant" or "unfriendly".

Problems in staff relations were common in this category, but there was variation in their severity. Some respondents had the view that disagreements mainly concerned some of the staff members and most of the teachers got along quite well with each other most of the time. The climate, for example, in the staff room was also mostly relaxed. However, there were also feelings that the atmosphere was oppressive and elusive.

"[…][…]When you come here in the morning and say hello at the door, someone might answer, or then not." [Maijala]

Despite problems between staff members, many interviewees emphasized that the relations between students and teachers were good and the problems did not affect them.

"In the classroom everything is okay. There are no problems with the students." [Maijala]

The tenseness of the school climate also reduced the well-being and job satisfaction of the staff members. Quite many of them commented that they did not feel very happy in their current school. As could be expected, none of the interviewees mentioned that they felt happy because of good relations among staff. Things that were mentioned in connection with job satisfaction were related to nice students or the current position satisfying their expectations. If collegues were mentioned it was because of a long common history, which created some fellowship. Some of the interviewees even emphasized that they purposefully wanted to stay out of disputes and concentrate on their own goals and teaching. Work satisfaction was gained by fulfilling these own professional goals, since the communal or social factor was missing.

> *"I have been happy here just because I can carry out my own goals, as I said earlier. But I would say that there are some conflicts and they certainly disrupt the working climate here. My climate. And it is a serious matter[...][...]" [Maijala]*

Even if it was not admitted by all the interviewees, the problems were also reflected in the broader discussion culture of the schools. Some respondents had the opinion that a real discussion culture was lacking and people did not express their views openly. Or sometimes they were expressed somewhat too openly, causing open arguments. All of this created disappointments and withdrawals. To some extent the frictions also prevented the development of the school as a whole. Problems in the discussion culture and considerable differences in views certainly complicated cooperation.

> *"I have this impression, and I have been also told, that there isn't any culture for discussion, I mean real culture for discussion. [...][...]" [Kontio]*

> *"The tradition in this school does not encourage to such discussions. Therefore, it relies on occasional relations. It only depends on some random factors and in that sense there are things that have not been put into practice as well as they could." [Maijala]*

In daily matters the culture was partly similar to that in other categories. Teachers discussed school matters mainly with other teachers of the same subject group. Some of the staff members cooperated, but in some cases conflicts between individuals prevented or disturbed this. Some teachers had the opinion that cooperation and the exchange of ideas was unsystematic and it was limited to school matters, and problems related to individual teachers were not discussed.

> *"There is quite a lot of discussion going on during the breaks and lunch hours. And it's mainly between the teachers of similar subjects, like language teachers. To a smaller extent with teachers teaching very different subjects. This would also be quite difficult, since the content is so different." [Kontio]*

> *"The discussions are not very organized. Individual teachers discuss with each other. I have discussions a couple of times during the school year with the science and humanities teachers. I also cooperate quite a lot with my closest colleague." [Maijala]*

Awareness of the problems was clear and there had been attempts to solve them with different approaches. Discussions had been started after open disputes and they had somewhat improved the situation. Other measures included the use of outside speakers or consultants. Teachers, however, felt that the latter approach was useless and therefore the results had been limited.

In terms of problems encountered by the teachers, for instance, in the classroom, the strategy of coping with them somewhat varied and the problems in staff relations clearly affected this. Where the problems were less severe, teachers were able to discuss them and unload their classroom experiences openly in the staff room.

> *"[…][…]It's quite usual to explode so that one comes indignant during the break and everyone then gets to know how terrible it has been. This is a way to relieve the pressure and often it's resolved this way. Naturally, we also discuss if someone has concrete problems and try to solve them. [Kontio]*

When relations were tenser the attitude of collegues was less emphatic. In such cases, teachers basically tried to solve their problems independently and alone. None of them mentioned an open and inclusive climate, e.g. the possibility to share unpleasant experiences in the staff room. Teachers' independent role was emphasized and even the help of the principal was considered as creating a credibility problem. For example, consulting the principal was considered questionable because teachers felt they could loose their face in front of the students.

> *"As an old hand, the starting point is, I think, to take care of the problems under one's own steam. Because once you go and ask for your superior's assistance, students know your rating. Okay, in this kind of situation she needed the principal's help. You kind of narrow your margins. During my stay here I have asked for the principal's assistance once, because one has to be very careful with this. And in that case there was also some humour in my mind." [Maijala]*

A poor climate and frictions among the staff members were also reflected in after-school activities. Teachers had them quite rarely and the cooperation between teachers was mainly related to instruction during the school hours. Teachers either did not feel that it was necessary or they did not have the time and interest to spend time together after school. Some of the teachers commented on the joint events with a highly negative tone. They felt that it was pretence to try spending leisure time together with people that you hardly say hello to during the day.

6. School governance and management

Educational leadership has been one of the key characteristics of effective schools, as was mentioned earlier. It refers to a wide range of factors such as how the school leader provides information, articulates the school goals, coordinates, and orchestrates participative decision making. In a narrower sense it refers to the leadership directed at the school's primary process and its immediate facilitative conditions, such as time devoted to educational versus administrative tasks, controlling of classroom processes, quality controlling of teachers, facilitating work-oriented teams, and initiating and facilitating staff professionalization (Scheerens and Bosker, 1997).

This study also examined some of the above-mentioned factors, but the main focus has been on governance and management rather than on school leadership. Therefore, governance and management are discussed here. This refers to the formal organizational structure, decision making and staff participation in decision making, and the development of school practices as a whole. The collaboration between the staff members is also discussed in this connection.

General upper secondary schooling has been quite tightly regulated in Finland and this has also been reflected in school governance, because it was quite similar in most of the case schools. The most common case was characterized as participative. There was also a case school with hierarchical governance and management. At the other end of the continuum it was characterized as equal.

Participative

In the majority of the case schools, management was open and participative. The roles of the school management and teachers were clear and accepted by the staff members. The principal had the leading role and was assisted by the vice-principal. Communication between the principal and vice-principal was characterized as open. The tasks and roles of the vice-principal varied to a small extent depending on the schools. In some of them, principals pointed out that the role and tasks of the vice-principal were considerable. As one of them playfully stated, “our division of tasks is such that vice-principal is doing everything and I’m doing nothing” [Talvio]. At the other extreme was the vice-principal’s view that she was more in a role of “lessening the workload of the principal” [Kontio].

Typical responsibilities of the vice-principal included preparing work schedules with the list of courses, a weekly info magazine, chairing the quality committee, space/facility arrangements, maintenance of the student database, chairing some special working groups, some responsibilities related to school resources and maintenance of computer facilities.

Only one of the schools had a management group, but in all of them there were some assisting groups such as a development group, planning group, quality group, quality committee, or subject groups. Their tasks differed little from those of a management group. Common properties for the groups were that they had representatives from different subject groups and their task was to assist the principal and vice-principal in some broader issues and prepare matters that could then be discussed in the staff meetings. The tasks of the subject groups were mainly related to curriculum development.

In most of the schools some additional working groups were also established. They ensured that teachers' views were not dismissed in important matters. Many of these groups had meetings, for example, once a month or once a period, but this varied to some extent depending on the group and its purpose.

> *"Always when we have some project to take care of we have this practice of appointing a working group. The members then sit down and work on their task… The results are then discussed in the staff meeting. Everyone has a chance to participate." [Ilves]*

The staff meetings were usually arranged once a week or every two weeks, but there were also exceptions that were based on teachers' requests. If meetings were less frequent, some discontent arose because principals consequently felt that professional discussions between teachers were reduced. In order to plan the upcoming semester and school work, additional planning meetings were held two or three times a year. Sometimes these meetings also had some special theme related to school development.

The division of the teaching load did not cause any major problems in these schools. In small schools it was considered more or less 'automatic' because in most of the subjects there was only one teacher. In that sense, "it followed the titles" [Ennala]. In larger schools, the division of the teaching load and preparation of the working plan was usually completed either by the vice-principal or some other person. After the proposal had been prepared, teachers discussed it by themselves and divided up the courses. It was usually settled in a good spirit. The principal naturally settled the matter in cases where there were differences in opinions.

> *"We always decide by ourselves in subject groups. In my subject it has gone well. There have never been any disagreements and I haven't heard of anyone else having them either." [Annala]*

In three of the schools, the principal was in her first year in the post. All the principals had brought changes in school leadership and these were greeted with a positive attitude. Greater openness and less authority were noteworthy features of the new principals. There had been "a change in generation" in that sense [Annala]. Moreover, teachers' opportunities to influence the development of the school had clearly improved. According to the teachers, this change was absolutely positive.

Decision making in these schools could be characterized as fairly open and participatory. Teachers' opinions were listened to and teachers were mainly satisfied with how their views were taken into account. Some principals stressed the importance of teachers' views by pointing out that they were not able to run the school by themselves. Therefore, it was also important that teachers felt free to express their opinions about school matters.

> *"Even if some teachers think that in some matters I stipulate, I would say in substantial matters, everyone is able to have an influence. It wouldn't make any sense if I would do it all by myself. I would burn myself out if I would stubbornly do it in my way." [Kontio]*

According to teachers it was usually easy to discuss different matters with their principal and vice-principal. In this sense the communication between staff members and the management of the school was working in most cases.

> *"I can say that I find it personally easy to discuss with them (principal and vice principal) about every matter, even some troublesome matter. It's easy to talk to them and we have always reached some kind of agreement." [Laurila]*

Teachers also felt that their views were taken seriously. All important matters were discussed together in the staff meetings and “all the decisions that were essential for the operation of the school have usually been quite freely made at the staff meetings” [Ennala]. According to the teachers “there were no such notion that someone would dictate from above” [Ennala]. Teachers also felt that everyone “takes part in the discussions and brings their own ideas and thoughts” [Metsälä]. However, the main point that “surely we all have a possibility to influence” [Metsälä], was clearly expressed.

Financial matters were also usually discussed in staff meetings, but the principal was the final decision maker. Teachers, however, stressed that there was open discussion about the principles and even difficult discussions were not avoided.

> *“[…][…]After the resources have been distributed there is always information about it and we get the numbers in paper. Everyone can then go through them and sometimes there have been very strict and difficult discussions about the principles. But they have always been open…” [Laurila]*

Most teachers stressed the significance that important matters were discussed with them. Questions related to their own subject were naturally emphasized most, but general matters were also considered important. In matters concerning their own subject, teachers were usually very independent.

> *“The curriculum work is important and also the development of one’s own teaching subject. But I also find it important that teachers can take a stand on the general goals. It’s not working if everything comes from the top. I don’t think it would work out if the principal would just say that that’s how we are going to do. I think that it is important that teachers experience/feel that this is our shared view. Teachers have to feel that they are pulling together, that this is our common interest. [Kontio]*

Teachers felt it was usually fairly easy to get new ideas through and the success was mostly dependent on the nature of the ideas. Decision making was a joint process in which not only the principal’s opinions affected the final result, but also the opinions of other staff members.

> *"I would say we would take it up in staff meeting and discuss about it. And if other staff members would also agree it was a good idea, I think it would be realized. We have had different kinds of projects that have required extra work from teachers and the involvement of students. They all have been realized. I don't remember that anybody has been turned off because of new ideas. In other words, I have the impression that if someone has good ideas and there are enough financial resources, they will be carried out." [Kontio]*

Since decision-making power always invloves responsibility, some teachers emphasized that they did not want to make decisions in matters they were not able to take responsibility for. As one teacher expressed it:

> *"There are matters that I don't want to be involved in when making decisions, since I'm not the one making decisions in matters that I'm not responsible for. I only make decisions on matters that I also take responsibility for." [Ilves]*

Equal

In one school the governance could be characterized as being equal. The character of the principal had mainly created this climate. The principal could be described as being "one of the staff members". Teachers expressed their satisfaction with the principal's style of managing the school. The small size of the school also made it easy to take such an open and equal approach. One teacher commented on the decision making and teachers' participation by saying that

> *"Our school is of the right size in a sense... In our staff room we have a large table that everyone fits around. Everyone comes together there and drinks coffee. It's a daily routine that we sit and discuss things there together. There is no tribalism in the sense that the teachers of mathematical subjects get together in one place and the language teachers or teachers who smoke or teachers who don't smoke in another place..." [Kalervo]*

The division of work between the principal and vice-principal followed the same principles as in the schools characterized as open and communicative. Because of the size of the school there were mostly no permanent groups assisting in school management. Working groups were formed to take care of certain specific tasks that were related to preparing, for example, the curriculum or other comparable tasks or reports. In addition to weekly staff meetings there were the usual planning meetings two or three times a year.

The division of the teaching load did not cause any problems. Where there was more than one teacher teaching the same subject they were able to negotiate the division of teaching lessons.

> *"I think where there are several teachers teaching the same subject, it is possible to bargain with them. Since the principal makes the schedule in the end she also is responsible for it. They can't be arranged however one likes, there are many things to be taken into account. But teachers can have an influence. If some teacher wants to lecture advanced or special courses, I'm sure it's okey." [Kalervo]*

Teachers' opinions were carefully listened to. In this school the principal also pointed out that she was not able to run the school by herself and therefore it was important that teachers also participated. The power of teachers to influence school matters was considered important.

> *"I always try to get everyone with me when I present some ideas. Because I can't do anything alone...It has to come from the teachers and I hope that teachers would have more ideas of all kinds." [Kalervo]*

The atmosphere was open and teachers seemed to be very satisfied with it. They were especially happy with the principal's open style of management. There was an open dialog between the teachers and the principal. Teachers felt that they were able to influence the decision making and they also were quite active in school matters.

> *[So teachers participate quite a lot in decision making?] "Yes, very much. The principal also pays attention to the teachers' views and listens them very much in matters related teaching and their own work." [Kalervo]*

Hierarchical

The governance in one school was characterized as hierarchical and individualistic. The principal had centralized quite a lot of power and decision making and as a consequence teachers felt somewhat constrained. Teachers also had the opinion that the principal was quite authoritarian.

As for the role and tasks of the vice-principal, they were quite limited. They mainly included “preparing monitoring lists, monitoring at the matriculation examinations, arranging a proper space for meetings and instruction, and teacher training. There were no management or other permanent assisting groups. The decision of not having a management group was a conscious one. According to principal, such a group would not work and could actually easily create friction among teachers because of jealousy.

> *“In some schools that have management groups they cause this tendency of many teachers feeling that the members form an odd clique that gains extra benefits. And often they also do. And this does not necessarily benefit the school but can actually harm it.” [Maijala]*

Even though there was no management group, some working groups were sometimes formed. Their tasks were related to preparing, for example, the curriculum or other comparable tasks or reports.

The role of the staff meetings was emphasized by the principal. They were open to all teachers and were considered as the main instrument to inform teachers of the current issues. She also underlined that transparency and the equal distribution of information is best ensured through collective decision making and thus that all teachers attend the staff meetings.

In addition to staff meetings there were some planning meetings two or three times a year. In addition to planning the upcoming semester, another goal was to increase “the fellowship among the staff” [Maijala]. In that sense, the practices were very similar to other case schools.

Regarding the decision making, teachers' discretion concerning their own classroom work was considerable and there were no constraints so that they "were able to do everything they wanted" [Maijala]. They also had "all the necessary support and training". Teachers naturally appreciated this very much.

> *"I would thank the principal that she doesn't interfere in my teaching. In some sense it's nice that she gives responsibility… So she doesn't interfere and there aren't any hindrances or anything on how to proceed. I mean, you can teach here very independently." [Maijala]*

The problem was the developent of the school as a whole, which created discontent among teachers. Even if teachers had initative they had the opinion that it was difficult to get approval for their ideas where they concerned the whole school. Therefore, teachers felt more or less constrained and frustrated when it came to school development outside their own classroom work.

> *"I would say that teachers participate very much. Teachers have many propositions and ideas. But they are perhaps difficult to realize because the principal makes the final decision. And there is real rigmarole when we try to get something done […][…]" [Maijala]*

The principal tried to emphasize open discussion in all matters, but the teachers did not have the same experiences in actual situations. Rather, they commented that there was no discussion culture and there were no attempts to develop any practices for the whole work community. Problems in staff relations also further complicated the situation.

> *"… But let's say that there are no opportunities for an open discussion. The pedagogical development takes place at the subject level and in your own brains. But if we take matters concerning the development of the whole school, they pass unnoticed and stay in the spare time and at the pub discussion level. In other words, the ideas don't find their way to the development of the day school and the improvement of the process of developing." [Maijala]*

"[...][...] But, on the other hand, there is no platform where new ideas could come up and be kicked. It falls down to single individuals. And if you work like that for many years you get into a rut. So, I think it would be a good idea to think over these matters together in the work community." [Maijala]

7. School curriculum

At the beginning of the 1990s, school-level autonomy and decision making was considerably increased in general upper secondary schools in Finland. As a consequence of this reform, the new national curriculum for general upper secondary schools published in 1994 was less detailed and schools were obliged to devise their own school curriculum. By the end of 1996 all schools had completed this work. At the time of the interviews, all case schools were in the process of renewing and making some required amedments to the curriculum.

The results of school effectiveness research showed that in successful schools teachers were involved in curriculum planning and played a major role in developing their own curriculum guidelines. As for the case schools, there were some clear differences in the process of making and renewing the curriculum. Partly they reflected the overall situation of the school and partly they were a result of teachers' own attitudes towards the development of the school processes and school work. In some cases the school curriculum was considered as an <u>administrative measure</u> that has to be written. For some, the process and curriculum was a <u>way to develop the work</u> and the school. There were also <u>mixed views</u> and practices containing features from both of these categories.

Administrative tool

In this category the school curriculum was mainly considered as an administrative measure. Many of the interviewees in these schools questioned it's relevance for practical work. It was seen as one of those papers that have to be written with no other significance.

"It's much of an administrative measure. And what is written in a certain section in the curriculum does not have any significance in daily work."[...][...] [Kalervo]

The national curriculum and text books were considered as one of the reasons for seeing the curriculum an administrative paper. Some of the interviewees had the opinion that it is quite difficult for teachers to have any other goals than that of students absorbing the required knowledge. In practice, instruction is dependent on the content of text books and thus, the content of the curriculum is restricted. Therefore, teachers felt that they didn't have much of a choice.

> *"In the end, you have to follow the text books. You can't do everything by yourself. But you can emphasize certain things. And therefore, I sometimes when I read the original curriculum find it somewhat too grand and important if you compare it to what happens in practice. But I guess they are some kind of backbone but the meat is somewhere else." [Talvio]*

The process of writing the curriculum was usually experienced as somewhat problematic. Neither the teachers nor the principals were totally satisfied with it. In some cases, the principal was somewhat disappointed about the passiveness of the teachers. In other cases, things were described as even worse.

> *"Unfortunately it's also here so that I have made the draft. Then I have asked teachers to read and comment on it during the development days. I think it's not a curriculum if it's prepared by the principal. But there were very few proposed alterations. So I really don't know how to do it." [Talvio]*

Mostly, the work itself was organized so that the general part was written by the principal and a group of teachers. Every teacher was responsible for their own subject. If there was more than one teacher teaching the same subject the work was conducted together. All the subject-specific parts were pulled together by the working group that was responsible for the general part.

> *"All teachers are responsible for their own subject and in case there are three teachers, together they are responsible. And then there is this general part which was also worked out together in a sense. The text I or the working group had first written was handed out and we went through it together." [Maijala]*

Problems in staff relations also manifested themselves through this process. The discussions were quietened by some of the staff members and teachers felt the situation was very unequal and depressing. They also felt that they were unable to influence the content of the curriculum as much as they wanted. Partly, that was clearly due to some tensions between certain teachers and partly it was due to inter-generational differences in views.

> *"The curriculum work was, I would say, a quite a negative experience of the cooperation that existed in this school. There were a few older teachers who practically dictated what was written down. Persons that love their own voice and complicated sentences and empty mumble. We have one teacher who made a short and succinct and clear version, but it was completely torpedoed by those people. It was something that indiated right away: be quiet, you don't understand anything." [Maijala]*

There were also exceptions to the negative views where case staff relations were working. Even though the staff members did not seem to value the curriculum very highly, the process of renewing it was considered good and productive since it opened up a possibility for thorough discussions. The work itself gave a chance for people to think over the school work and discuss the content of it with colleagues.

> *"The ideas were discussed, which probably had some effect. We could have done it in other ways, too. But once in a while it is good to sit down and discuss and think over of what is most important." [Kalervo]*

Tool for school development

The views towards the curriculum were quite positive and a realistic approach to preparing the curriculum was characteristic of this category. The process of updating and writing the curriculum had been rich and people were usually satisfied with it. It gave the staff members a possibility for open discussion about the school goals. Teachers also gradually took the curriculum as tool to develop their own work.

When the work started interviewees took a somewhat reserved stand towards it. In some schools this was because teachers were somewhat afraid of expressing their personal opinions and felt a lack of confidence when faced with new tasks and responsibilities.

> *[...][...] "It was a new and strange situation when you had to start thinking of some principles by yourself. Naturally, quite many were longing for clear guidelines on how to proceed. It was the way it had always been so why couldn't it still be that way." [...][...][Laurila]*

However, as soon as teachers realized all the possibilities involved they started seeing the curriculum in a more positive light. They even felt the work to be challenging and interesting, since in subject groups they were able to plan their courses and make new courses. They were also able to plan courses that had no text books that would guide the progress of students. Teachers "were able to plan the content of the whole course by themselves." [Laurila] Curriculum work was also seen as a way to develop one's own work, especially for those with a longer career.

> *"Since I have been here a long time I felt it more as a reminder that things can also be done differently. What can be stressed and what can be left out and how things can be carried forward together and what is our shared interest. And from the personal point of view, to show that I can also renew and not get into a rut." [Metsälä]*

It took quite some time before teachers realized that they have the freedom to plan their own courses. At the beginning they more or less followed old procedures. Eventually, teachers realized that curriculum development is an ongoing process. And now the curriculum is living all the time so that teachers "are constantly bringing new ideas and making new courses". [Laurila] Doing it in the old way was a frustrating process. In the new setting, the curriculum started to have a real meaning for teachers.

Naturally, there were also some differences among teachers in how they viewed the curriculum. Some of the teachers "never bother to read it through or think what it means". [Laurila] Some of the teachers, after writing and reading the curriculum, started seeing students "as more comprehensive" by emphasizing things that are set as goals in the curriculum. And "somehow these things are also thought over in instruction" [Laurila].

To some extent the interest in the curriculum and its development was associated with the age of the teachers, so that younger teachers expressed somewhat more positive views.

> *"In the first time, when we revised the curriculum, there was quite a lot of this dismissive attitude. It was said that it's only a paper. Well, I know that some had a lack of time and lot of work, so in the end the attitude was maybe not so negative. But then there are also some who have taken it as a challenge. But I must say that the attitude varies according to the age of the teacher." [Ilves]*

The argumentation over the content of curriculum and instruction was reversed in many of these schools. It also reinforced and increased the meaningfulness of the curriculum work. Most of the time the argumentation about the curriculum work had the direction that what was written in the curriculum must become materialized in the daily teaching work. Many of the interviewees turned things the other way around, so that teachers documented in the curriculum their teaching work and practices.

> *"Since I have long experience, I would say that they come out of my teaching profession. [...] I am, at least, very aware of the excesses. I mean, writing things that I know are nonsense. The things written down in the curriculum must be realistic. One must bear in mind the environment one is working in. One must take into account the capacity of the children. If the student body radically changes after two years we must adjust our curriculum to these changes. So I would see it like this. And not so that we first write something and then try to apply it. It doesn't work. Then it's only a written word on the bookshelf and not included in the actual work. [Ilves]*

However, some teachers felt constrained by the text books and the thought that books mainly steer the process of writing the curriculum at the school, since the instructions follow the books. The content in each subject and course is therefore very much dependent on the content of the books. (These, in turn, follow the content of the national core curriculum). Therefore, the curriculum "was some kind of nonsense all together" [Ennala]. According to this view, its credibility suffered because when books were published their content did not exactly follow the core curriculum.

Used to be an administrative tool

In this category, views were divided. In some cases the curriculum was considered an administrative document with little relevance for everyday work. However, changing views were also present largely because of the successful process of updating the curriculum. People were usually content with the process of preparing and updating the curriculum because they felt their views were listened to and they had the possibility to influence the content.

With the former school curriculum the respondents had the view that it was most likely written "because it had to be done and it was done with the minimum effort" [Annala]. Teachers also expressed views that the curriculum work was "let's say, the nonsense of the school department" [Annala]. The curriculum itself was considered "rather as a paper". The main benefit of it was that maybe "you check out at the beginning of each period what has been written down." In that way it affects as an ulterior motive the goals in each course and serves as "a whip against teachers' backs". Teachers also felt that "most likely not all teachers have internalized or even read the curriculum." [Kontio]

During the past years, views had started to change so that "teachers were developing the curriculum together" and they had realized "that curriculum work is actually school and job development. And with the curriculum work they have the keys to that." [Annala] Practical views were also expressed according to which the significance of the curriculum depended on the content being written realistically and so that would be possible to implement.

> *"It makes sense if it's made this way. In other words, if it is simple enough. And so that only those things are included that can be implemented." [Kontio]*

Despite some negative views, respondents also commented that the curriculum had begun to 'live', so that its content was annually checked and revised if necessary. For example, the course offerings were checked and courses that did not attract students were removed. This change was new compared to earlier practices.

In this group, teachers seemed to be happy with the process of revising the curriculum. They were involved throughout, "each, more or less, naturally according to their own interests, but all were involved" [Kontio]. They also had a possibility to influence the content of the curriculum. The basic rule was that the general part was thoroughly discussed and changes were made if necessary. In at least one school there was also a student member taking part in the discussions.

> *"That's the principle. The content of the general part is agreed together. It is talked through together and changes are made if necessary, sometimes in a very detailed manner by changing the words. [So content is discussed.] Yes, especially the educational goals and general trends and emphases." [Kontio]*

"Everyone was taking part really well and with the attitude that we will really think over this thing. And we generated goals and pondered. Some matters were discussed much and some less. But it was completed together." [Kontio]

8. Interrelationship between the categories

It is apparent from the preceding descriptions that the different categories are at least to some extent interrelated or correlated. If there are problems in staff relations, processes involving interaction between staff members such as participation in decision making and curriculum work may be difficult to carry out. Table 2 decribes the interrelationships between the different categories by showing the category of each school within the themes.

As expected, schools with professional staff relations usually have school governance and management practices that enable teachers to become involved in school matters and decision making so that they are satisfied with the situation. If there were problems in staff relations, the governance and management of the school was also more centralized in nature and there was dissatisfaction among teachers concerning their role in school decision making.

A less systematic connection existed between how the students were viewed and the rest of the themes. In schools with caring and respecting views of the students, staff relations were characterized as either professional or friendly. However, the three schools with hesitant views more often did not share any similar patterns. Some had tense staff relations but others very friendly relations. The school governance and management practices and views concerning the curriculum work were also different.

Based on the categories, case schools can be further grouped into four groups. The first group is formed by Ennala and Metsälä Schools. These were characterised by attentive views about the students, friendly or professional staff relations, communicative governance and management, and curriculum development that was considered productive and as a tool for work and school development.

Annala and Ilves Schools had a very similar categorization based on the themes. Staff members had respecting views about the students, staff relations were professional and there was a mutual exchange of ideas in the decision making of the school. The views about the curriculum somewhat diverged so that they were either developing or the curriculum was already considered as a development tool. Laurila and Talvio Schools were slightly different from Annala and Ilves Schools. The views about the students were educating, i.e. more formal. In Talvio School the school curriculum was considered more as an administrative task.

Three of the schools with hesitant views about the students had mixed results concerning their staff relations, governance and management practices, and views about the curriculum. Two of them (Kontio and Maijala Schools) also had tense staff relations, but the views and practicies differed in school governance and curriculum work. Staff members had a hesitant view about the students but the relations between staff members were very good and management practices supported teachers' initiative and participation in decision making. The possibilities offered by the school curriculum were seen as limited, but the process of preparing and revising it was considered very fruitful.

*Table 2. Placement of schools into categories and formation of groups (Gr 1-4)**

	Annala	Ennala	**Ilves**	***Kontio***	***Kalervo***	*Laurila*	Metsälä	***Maijala***	*Talvio*
Views concerning the students									
Attentive		Gr 1					Gr 1		
Respecting	**Gr 2**		**Gr 2**						
Educating						*Gr 3*			*Gr 3*
Frustrated				***Gr 4***	***Gr 4***			***Gr 4***	
Staff relations									
Tense				***Gr 4***				***Gr 4***	
Professional	**Gr 2**		**Gr 2**			*Gr 3*	Gr 1		*Gr 3*
Friendly		Gr 1			***Gr 4***				
Governance and management									
Centralized								***Gr 4***	
Participative	**Gr 2**	Gr 1	**Gr 2**	***Gr 4***		*Gr 3*	Gr 1		*Gr 3*
Equal					***Gr 4***				
School curriculum									
Administrative tool					***Gr 4***			***Gr 4***	*Gr 3*
Purpose unclear	**Gr 2**			***Gr 4***					
Tool for work and school development		Gr 1	**Gr 2**			*Gr 3*	Gr 1		

*The groupings are formed in the text.

9. Categories, groups and inefficiency

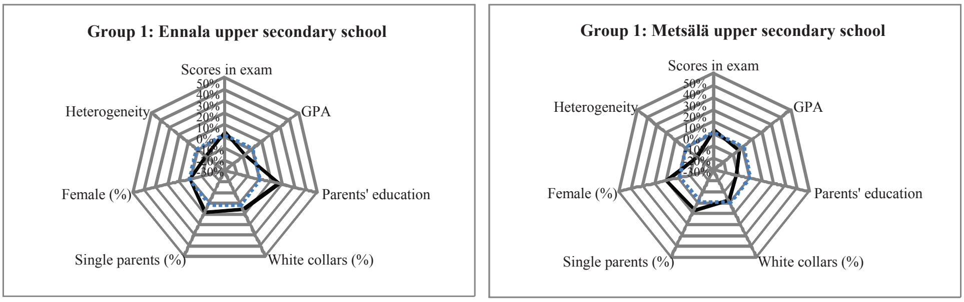
Finally, the question of whether the findings concerning case schools are related to their inefficiency scores is discussed. The relationship between the previously-described groups and inefficiency is depicted by calculating the average inefficiency of the schools in each group (see Table 3). It must, however, be noted that the following results are only tentative because of the low number of cases. For the same reason, no statistical tests for significance are performed.

Table 3. Average inefficiency in different groups

	Random effects	Fixed effects
Group 1 (Ennala, Metsälä)	0.04	0.14
Group 2 (Ilves, Annala)	0.04	0.11
Group 3 (Laurila, Talvio)	0.10	0.17
Group 4 (Kontio, Kalervo, Maijala)	0.16	0.24

As shown in Table 3 groups 1 and 2 had a lower average inefficiency than groups 3 and 4. In the first group, the intake of students in terms of GPA and heterogeneity of the student body was infavorable (see Figure 3). Despite these facts, students scored little above the national average in the matriculation examination. The views of the staff members concerning the students were attentive, so that interviewees emphasized the commitment to give every student a chance to matriculate. Extra attention was especially devoted to the low-achieving students and teachers were also proud when these students succeeded.

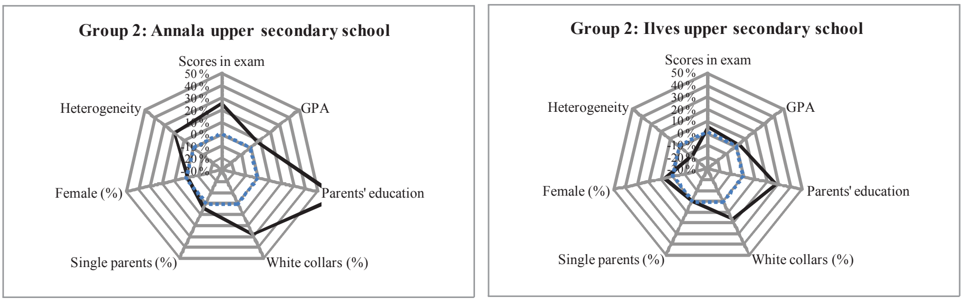
Figure 3. *Background of students in group 1 in 2000. Percentage deviation from the national average (dotted circle in the figure). Values outside the circle indicate favorable situation for the school*



Note: see Figure 2.

In group 2 the background of students was clearly favorable (see Figure 4). Students were motivated and in some cases highly motivated, and and they performed above the national average in matriculation examination. These facts were reflected in interviewees' views of the students which were characterized as trusting and confident. This view stressed the students' own responsibility for the school work. Low motivation and performance was uncommon and it was treated more as identity searching which did not require any major measures by the teachers. Respondents were generally confident that students would cope with their problems by themselves. The attention of the teachers was more on promoting the high performance of students.

Figure 4. *Background of students in group 2 in 2000. Percentage deviation from the national average (dotted circle in the figure). Values outside the circle indicate favorable situation for the school*



Note: see Figure 2.

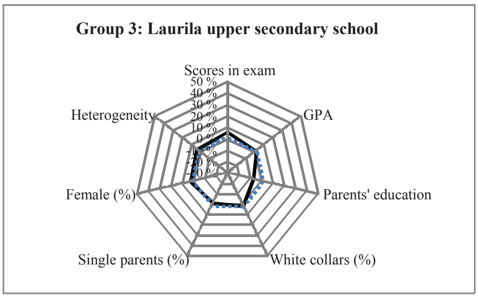
Staff relations in both groups were open and smooth and they were characterized as professional. There was a sense of collegiality involved. The discussion culture was open and it focused on school work. Staff members were also able to settle disputes without long-lasting consequences. There were no cliques and teachers discussed and exchanged ideas and information about, for instance, instruction and the problems of students. Teacher independence was also emphasized so that everyone was comfortable doing their tasks without the feeling that someone was breathing down their neck.

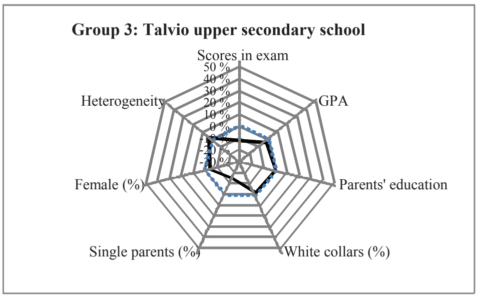
Governance and management of the schools was participative. Cooperation between the principal and vice-principal was good. The roles in the school were clear and accepted by the staff members. The principal had the leading role and was assisted by the vice-principal. The tasks of the vice-principals varied to some extent from school to school. There were one or more groups (management or subject groups) assisting school management in matters requiring more extensive preparation. Teachers participated into decision making and felt that their views and suggestions were taken into account in all important matters. In their instruction and classroom work, teachers were very independent.

The school curriculum and curriculum work was perhaps earlier seen more as an administrative task with only limited relevance to everyday school work. This view, however, had been changing as interviewees realized that the curriculum was actually living. The curriculum was updated regularly and during this work it was to a growing extent seen as a tool that could also be used in developing one's own work. Teachers had realized that they were able to influence the content of courses and school matters through the curriculum. The curriculum work in these schools had mostly been participative and in some schools also created lively discussions about the ultimate goals of the instruction and school work. The process of writing and renewing the curriculum was also described as one way to exchange ideas and increase the mutual understanding and internal cohesion among staff members.

The average efficiency was lower in groups 3 and 4. Concerning student background, schools in group 3 were close to the national average and similar to schools in group 1 (see Figure 5). Views about the students in this group were characterized as educating, emhasizing the students' own responsibilities. Low-achieving students were seen in a conflicting light in the sense they were understood, but there were also views that students were mostly looking for child minding and that their place was not in general upper secondary school. Some attention was paid to these students, but the approach was such that the students' own initiative was stressed. Staff relations, school governance and management and school curriculum work in group 3 was very similar to those schools in group 1 and group 2.

Figure 5. *Background of students in group 3 in 2000. Percentage deviation from the national average (dotted circle in the figure). Values outside the circle indicate favorable situation for the school*

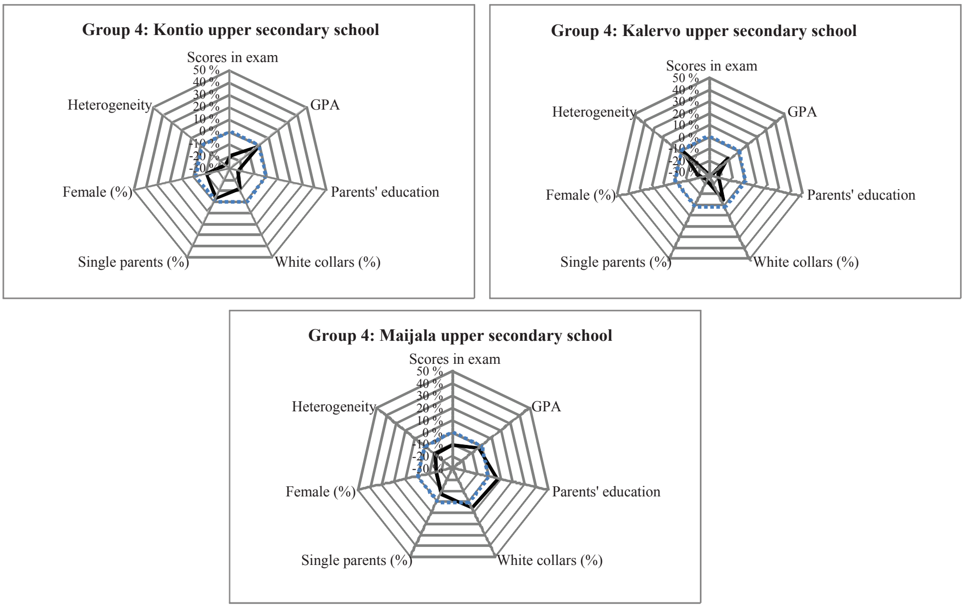




Note: see Figure 2.

The least efficient schools (group 4) in this study had in most cases a larger share of low-performing students than the other case schools (see Figure 6). This fact was also reflected in respondents' views about the students. They were to some extent frustrated and hesitant. Students were characterized with terms such as underperformance. Some teachers also had the view that the students were mainly passing time in school. Disappointment and in some cases also some caring reactions were noted. In some of the schools with many low performers, teachers felt that the absence of role models further lowered the standards of the students. Large skill differences were also found problematic for instruction.

Figure 6. *Background of students in group 4 in 2000. Percentage deviation from the national average (dotted circle in the figure). Values outside the circle indicate favorable situation for the school*



Note: see Figure 2.

In addition to problems with the students, in two of the schools there were problems in staff relations that affected the job satisfaction of staff members. Staff members openly admitted that the situation could be better and that there were problems among some of the staff members. The discussion culture of the schools suffered because of these problems and people were not able to express their views openly. Because of personal problems and disputes the cooperation was limited to school matters. Even though teachers emphasized that the relations between teachers and students were good, the unsatisfactory situation between the staff members created friction and discontent.

Problems in staff relations also had consequences for the governance and management of the schools, especially if they were severe.

Governance and management was hierarchical so that most of the tasks were centralized to the principal and teachers felt rather constrained. The vice-principal's role was limited and the development of school matters was placed in the hands of a few. Teachers lacked the possibilities to have an influence when it came to the development of matters concerning the whole school. The lack of an open culture of discussion between the teachers and problems in staff relations also complicated the matter. As for the teachers' own classroom work, the situation was considered good and they had quite free hands and the support of the principal.

In this group, the relevance of the school curriculum and the curriculum work to the practical work done in the school was more often questioned. Probably because of these views and problems in staff relations, the process of preparing the curriculum was mostly experienced as problematic and there was dissatisfaction concerning the final outcome.

In Group 4, one school was an exception to the above descriptions in terms of staff relations and governance and management. Despite the clearly unfavorable intake of students, teachers felt comfortable in the school, staff relations were characterized as friendly. Teachers openly discussed the students and difficult classroom situations. Staff members were very satisfied with the situation and they praised the climate and the discussion culture.

The friendly relations were accompanied by open and equal governance and management. The difference in having open and participative governance and management was small, but the principal's attitude of being one of the others made the difference from the other groups. In this setting, teachers' opinions and ideas were carefully listened to and their contribution was considered very important. Teachers were very content with the open style and there was a constant dialog between the staff members and the principal.

Finally, the relationship between inefficiency and the themes is depicted in Table 4. It shows that schools with attentive and trusting views were more efficient than schools with educating and frustrated views. Problematic staff relations were reflected in inefficiency since schools with tense relations had the highest inefficiency compared to schools with professional and friendly relations. Schools with participative school governance and management turned out to be more efficient than rest of the schools. Concerning the school curriculum and curriculum work, schools taking the curriculum as a way to develop the school and work were most efficient.

Table 4. Inefficiency and views concerning the students, staff relations, school governance and management and curriculum work

	Random effects	Fixed effects
Views concerning the students		
Attentive (Ennala, Metsälä)	0.035	0.141
Trusting (Annala, Ilves)	0.042	0.109
Educating (Laurila, Talvio)	0.100	0.175
Frustrated (Kontio, Kalervo, Maijala)	0.163	0.239
Staff relations		
Tense (Kontio, Maijala)	0.144	0.223
Professional (Annala, Ilves, Laurila, Metsälä, Talvio)	0.066	0.147
Friendly (Ennala, Kalervo)	0.113	0.193
School governance and management		
Hierarchical (Maijala)	0.134	0.214
Participative (Annala, Ennala, Ilves, Kontio, Laurila, Metsälä, Talvio)	0.073	0.154
Equal (Kalervo)	0.202	0.272
School curriculum		
Administrative tool (Kalervo, Maijala, Talvio)	0.167	0.239
Used to be an administrative tool (Annala, Kontio)	0.092	0.167
Tool for school development (Ennala, Ilves, Laurila, Metsälä)	0.040	0.129

10. Conclusions

Studies on the efficiency of schools have traditionally concentrated on determining the efficiency differences with various statistical or non-parametric methods. Seldom has a further step been taken to investigate the organizational or pedagogical characteristics and practices of schools that might produce the efficiency differences. In educational research this kind of approach has been more common in school effectiveness studies. This study analyzed the views of staff members concerning the students, staff relations, school governance and curriculum work in nine case schools that differed in efficiency based on the results of an earlier study. The analysis was based on interviews with principals and teachers.

The findings of this study show that teachers in efficient schools were attentive, implying that they considered every student as important. There was a sense of pride when respondents described how they succeeded with low-performing students. In schools with a high proportion of strongly-performing students, teachers' views were trusting and students were described as being conscientious youngsters. In inefficient schools the views were more frustrated and disappointed. Weaker students were more often seen as those who should not be in a general upper secondary school and there was less talk about taking care of all students.

Partly these views were related to the background of students. In schools with clearly lower GPA and higher heterogeneity of skills teachers felt frustrated and disappointed. These views are understandable in such a case. It is also possible that this situation is hard to overcome just by trying to change the views and attitudes. It may also require other measures and additional resources. However, in some schools same views were expressed even though student characteristics did not differ from the average one. Therefore, it is not just the poor performing students that cause frustration but some common view or attitude that was present in those schools. These views influence teachers' behavior and gradually also students' behavior as they begin to act according to expectiations. Hence, these views are also reflected in efficiency.

Staff relations in all efficient schools were good and they were characterized as professional in this study. Very good staff relations were not, however, only a hallmark of efficient schools, since in one inefficient school the situation was also very good. In schools with professional staff relations, staff members got along well with each other. There were no major problems and there was a sense of collegiality so that teachers were sharing matters related to instruction and students. Collaboration between the teachers and principal was usually frictionless and staff members were happy with the situation. In inefficient schools, there were more often problems in staff relations and these were openly admitted. These problems also complicated the collaboration between the teachers.

As for school governance and management, in efficient schools the roles of the principal and teachers were clear. The management and decision making was participative so that teachers were satisfied with how their views were taken into account. In inefficient schools, principals were more often isolated and had centralized most of the responsibility and tasks for themselves. Decision making was also centralized so that teachers were discontent with their possibilies to influence school matters.

The school curriculum and curriculum development were more often seen as way to develop the school and one’s own work in efficient schools. In inefficient schools they were considered more as an administrative measure with a little relevance to everyday work. The process of writing the curriculum was also considered problematic, either because of the passiveness of teachers, the lack of a culture of discussion and problems in staff relations. Teachers were also more often discontent with the final outcome.

In addition to the above-mentioned topics, the interview data contained information on evaluation and monitoring practices of schools, staff development, and parent-school relations. These practices were very similar in every school and they were therefore not analysed any further in this paper. Monitoring of student performance was usually conducted after every period in staff meetings. Students with problems were identified and discussed. It was admitted in almost every school that they had some problems in keeping a record of every student. The system with no fixed classes provided an opportunity for some students to fall behind without teachers noticing it early enough. Staff development was usually based on individual interests and there were very few school-wide development programs. Teachers mainly attended courses a few days a year that were related to their own subject. Parent-school relations were considered mostly as of minor importance, since students were on their way to independence. Teachers and principals rather emphasized the importance of discussing all important matters directly with the students.

Many of the findings of this study related to staff relations, school governance and management and curriculum work are similar to earlier school effectiveness studies (Teddlie and Reynolds, 2000; Sammons, 1999; Sammons *et al*., 1998). Views concerning students have not usually been addressed and the analysis of this study showed that they differed between the efficient and inefficient schools and they are important to take into consideration. The study also showed that some of the findings of school effectiveness studies related to evaluation and monitoring practices, staff development and parent-school relations didn't apply to Finnish general upper secondary schools or at least to case schools of this study.

As for the method used for efficiency measurement, the inefficiency scores based on stochastic frontier analysis most likely captured some more permanent characteristics of schools related to their organizational practices. Hence, this method is a promising instrument to use in efficiency evaluation of schools. To gain more certainty on this matter, a large survey data would be needed that would enable the statistical testing of the case study results.

References

Bradley, S. – Johnes, G. – Millington, J. (2001): The Effect of Competition on the Efficiency of Secondary Schools in England. *European Journal of Operational Research* 135, 545–568.

Brookover, W. B. – Beady, C. – Flood, P. – Schweitzer, J. – Wisenbaker, J. (1979): *Schools, Social Systems and Student Achievement: Schools Can Make a Difference*. New York: Praeger.

Coe, R. – Fitz-Gibbon, C.T. (1998): School Effectiveness Research: Criticism and Recommendations. *Oxford Review of Education* 24, 421–438.

Coelli, T. – Rao, P.D.S. – Battese, G. E. (1999): *An Introduction to Efficiency and Productivity Analyses*. Boston: Kluwer Academic Publishers.

Duncombe, W. – Miner, J. – Ruggiero, J. (1997): Empirical Evaluation of Bureaucratic Models of Inefficiency. *Public Choice* 93, 1–18.

Dodd, A. (2006): Investigating the Effective Use of Resources in Secondary Schools. Research Report to DfES. *Reserch Report* No 779. Hedra plc, London.

Goldstein, H. (2003): *Multilevel Statistical Models*. 3rd ed. London: Edward Arnold.

Goldstein, H. – Woodhouse, G. (2000): School Effectiveness Research and Educational Policy. *Oxford Review of Education* 26, 353–363.

Gray, J. – Wilcox, B. (1995): *Good School, Bad School: Evaluating Performance and Encouraging Improvement*. Buckingham: Open University Press.

Greene, W. (2005a): Fixed and Random Effects in Stochastic Frontier Models. *Journal of Productivity Analysis* 23, 7–32.

Greene, W. (2005b): Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model. *Journal of Econometrics* 126, 269–303.

Grosskopf, S. – Hayes, K. – Taylor, L. – Weber, W. (2001): On the Determinants of School District Efficiency: Competition and Monitoring. *Journal of Urban Economics* 49, 453–478.

Johnes, J. (2004): Efficiency Measurement. In Johnes, G. and Johnes, J. (eds.) *International Handbook on the Economics of Education*. Cheltenham/ Northampton: Edwar Elgar Publishing.

Kirjavainen, T. (2008): Understanding Efficiency Differences of Schools: Practitioners' Views on Students, Staff Relations, School Management and the

Curriculum. *VATT Discussion Papers 450*. Helsinki: Government Institute for Economic Research.

Kirjavainen, T. (2007a): The Efficiency of Finnish Upper Secondary Schools in 2000–2004. (In Finnish). *VATT Research Reports 131*. Helsinki: Government Institute for Economic Research.

Kirjavainen, T. (2007b): Efficiency of Finnish Upper Secondary Schools: An Application of Stochastic Frontier Analysis with Panel Data. *VATT Discussion Papers 428*. Helsinki: Government Institute for Economic Research.

Kirjavainen, T. (1999): Efficiency of Finnish Senior Secondary Schools: Explaining High and Low Performance with Resources, Environmental, and Organizational Factors. *Helsinki School of Economics and Business Administration Working Papers W–221*. Helsinki: Helsinki School of Economics and Business Administration.

Kirjavainen, T. (1998): *Efficiency of Schools: The Impact of Environmental and Organizational Factors on Efficiency Measured by DEA*. An unpublished licenciate thesis. Helsinki School of Economics and Business Administration.

Kirjavainen, T. – Loikkanen, H. A. (1998): Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit-Analysis. *Economics of Education Review* 17, 377–394.

Kumbhakar, S. C. – Lovell, K. C. A. (2000): *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

Levine, D. U. – Lezotte, L. W. (1990): *Unusually Effective Schools: A Review and Analysis of Research and Practice*. Madison, WI: National Center for Effective Schools Research and Development.

Lupton, R. (2005): Social Justice and School Improvement: Improving the Quality of Schooling in the Poorest Neighbourhoods. *British Educational Research Journal* 31, 589-604.

Luyten, H., – Visscher, A. – Witziers B. (2005): School Effectiveness Research: From a Review of the Criticism to Recommendations for Further Development. *School Effectiveness and School Improvement* 16, 249-279.

Mortimore, P. – Sammons, P. – Stoll, L. – Lewis, D. - Ecob, R. (1988): *School Matters: The Junior Years*. Wells, Somerset: Open Books.

Nicolaidou, M. – Ainscow, M. (2005): Understanding Failing Schools: Perspectives from the Inside. *School Effectiveness and School Improvement* 16, 229-248.

Portela M. C. – Camanho, A. S. (2007): Performance Assessment of Portuguese Secondary Schools. Working paper, Economics 07/2007. Universidade Catolica Portuguesa (Porto). http://dspace.feg.porto.ucp.pt:8080/dspace/

bitstream/2386/100/1/072007+-+Portela+e+Camanho+-+Performance+assessment+of+portuguese+secondary+schools.pdf

Purkey, S. C. – Smith, M. S. (1983): Effective Schools: A Review. *The Elementary School Journal* 83, 427–452.

Raudenbush, S. – Bryk, A. (2002): Hierarchical Linear Models: Application and Data Analysis Methods. 2nd ed. Newbury Park: Sage Publications.

Reynolds, D. – Packer, A. (1992): School Effectiveness and School Improvement in the 1990s. In Reynolds, D. – Cuttance, P. (eds.) *School Effectiveness: Research, Policy and Practice*. London: Cassell.

Reynolds, D. – Teddlie, C. (2001): Reflections on the Critics and Beyond Them. *School Effectiveness and School Improvement* 12, 24–31.

Reynolds, D. – Teddlie, C. (2000): The Processes of School Effectiveness. In Teddlie, C. and Reynolds, D. (eds.) *The International Handbook of School Effectiveness Research*. London/New York: Routledge.

Rutter, M. – Maughan, B. – Mortimore, P. – Ouston, J. (1979): *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children*. Boston: Harvard University Press.

Sammons, P. (1999): *School Effectiveness. Coming of Age in the Twenty-First Century*. Lisse: Swets & Zeitlinger.

Sammons, P. – Thomas, S. – Mortimore, P. – Walker, A. – Cairns, R. – Bausor, J. (1998): Understanding Differences in Academic Effectiveness: Practioners' Views. *School Effectiveness and School Improvement* 9, 286–309.

Scheerens, J. – Bosker, R. (1997): *The Foundations of Educational Effectiveness*. Oxford/New York/Tokyo: Pergamon.

Silverman, D. (2000): *Doing Qualitative Research. A Practical Handbook.* London: Sage Publications Ltd.

Teddlie, C. – Reynolds, D. (2000): *The International Handbook of School Effectiveness Research*. London/New York: Routledge.

Teddlie, C. – Stringfield, S. (1993): *Schools Make a Difference: Lessons Learned From a 10-Year Study of School Effects*. New York: Teachers College Press.

Worthington, A. C. (2001): An Empirical Survey of Frontier Efficiency Measurement Techniques in Education. *Education Economics* 9, 245–268.

Article V

Reprinted with the permission of Elsevier Limited, Copyright © 2003 Elsevier B.V.

Pergamon

Economics of Education Review 22 (2003) 329–335

Economics of Education Review

www.elsevier.com/locate/econedurev

School resources and student achievement revisited: new evidence from panel data

Iida Häkkinen [a], Tanja Kirjavainen [b], Roope Uusitalo [a,*]

[a] *Government Institute for Economic Research, PO Box 269, 00531 Helsinki, Finland*
[b] *National Board of Education, Hakaniemenkatu 2, 00530 Helsinki, Finland*

Received 29 January 2001; accepted 17 April 2002

Abstract

In this study we analyze the effects of changes in school spending on changes in student performance. We use a large sample of matriculation examination scores of Finnish senior secondary school students from the years 1990–1998. We estimate fixed-effect panel data models that use the dramatic changes in the school spending caused by the 1990s' recession as identifying variation. According to the results, changes in teaching expenditure did not have a significant effect on the test scores. The grade point average in comprehensive school and the parents' education are the strongest explanatory variables for student achievement.
© 2002 Elsevier Science Ltd. All rights reserved.

JEL classification: I2

Keywords: Senior secondary schools; School spending; Student achievement

1. Introduction

The effect of school resources on student achievement has been examined in hundreds of studies during the past thirty years. Still, the question is far from settled. Existing studies have failed to show conclusive evidence that additional resources to schools would improve learning. In his latest survey, Hanushek (1997) concluded that there is no systematic evidence that more resources, such as higher teacher–student ratios or per-student expenditures would improve student learning.[1]

A basic problem in analyzing the effect of resources on student achievement is that resources are likely to be correlated with unobserved characteristics that affect achievement. Typically poor school districts can afford to spend less on education, but it is not clear whether low spending levels or other characteristics of poor neighborhoods are responsible for the low achievement levels in these schools. On the other hand, rural schools with small numbers of students often have smaller classes and higher per-student spending. The students in these rural schools may perform differently from urban schools for reasons unrelated to school resources.

Studies that use experimental data with random assignments to small classes provide the most convincing evidence on the effect of smaller classes and higher per-student spending. Random assignment removes any systematic correlation between the class size and unobserved variables, and, hence, provides an unbiased estimate of the effect of class size on student achievement. One of the best-known examples is Krueger (1999), who used data from the Tennessee STAR experiment, and found that students in smaller classes performed better.

Some recent studies rely on "natural experiments" to

* Corresponding author. Tel.: +46-18-471 1115; fax: +46-18-471 1478.

E-mail address: roope.uusitalo@nek.uu.se (R. Uusitalo).

[1] Other recent surveys on school resources and student achievement include Hanushek (1986); Betts (1996); Hedges and Greenwald (1996), and Card and Krueger (1996).

0272-7757/03/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0272-7757(02)00060-2

create exogenous variation in class size. Angrist and Lavy (1999) identify the effect of class size from discontinuous changes in class size imposed by Maimonides rule that determines the relationship between school enrolment and class size in Israel. Hoxby (1998) identifies the effect of class size from natural population variation in Connecticut school districts. Angrist and Lavy find that class size reductions induce a substantial increase in math and reading achievement. Hoxby finds no significant class size effects.

However, most studies on the effects of school spending use cross-section data. Even if the studies use a "value-added" specification as an attempt to control for the initial achievement level of the students, there is usually no attempt to control for the non-random assignment of students to different schools or to different classes. Since randomized experiments are rare and clever natural experiments hard to come by, it is useful to see if more traditional estimation methods could be used to control for the unobserved school-specific factors.

In this paper we study the effects of the changes in school spending on matriculation examination results using a large sample of Finnish senior secondary school students. We have a representative sample of students in all Finnish senior secondary schools for the years 1990–1998. Nine cohorts of students allow us to control for the time-invariant differences across the schools, and our estimates are based on the differences in the changes in the school spending. Therefore, we can eliminate any permanent differences across the schools and control for the neighborhood effects in a manner that is not possible in the cross-section studies. We can also control for family background differences and the initial level of achievement when students apply to the senior secondary schools.

Our data cover a period when school spending varied considerably across schools. Some local governments, who are responsible for school funding, suffered more than others from the recession in the early 1990s. Spending per student, adjusted for inflation, decreased on average by 25 percent, but the decrease was by no means uniform. In addition, the local government financing system was reformed in 1993, giving the local governments more discretion on how to spend the state subsidies they receive for running the school system.

Most studies on the effects of school resources use standardized test scores as a performance indicator. Typically, these tests suffer from a selection bias, since only a non-random sample of students take part in testing. In this study, we use results from the Finnish senior secondary school matriculation examination, which is compulsory for all students.

The rest of this paper is organized as follows. Section 2 summarizes the main features of the Finnish senior secondary school system. Section 3 discusses the data. Section 4 describes the recent developments in school resources. Section 5 presents the estimation results and Section 6 concludes.

2. Features of the Finnish senior secondary schools

The Finnish senior secondary schools provide three years of general education for students who are between 16 and 19 years of age. About 37,000 new students start senior secondary school each year. This is roughly 50 percent of the age cohort. The other half of the age group enter vocational education. Only approximately 5 percent of the age group quit school after compulsory comprehensive school.

Admission to the senior secondary schools is selective, and it is based on the grade point average (GPA) in the comprehensive school. Competition to the best senior secondary schools is harder, and competition is generally harder in the larger cities. Therefore, the admission criteria and the average GPA of the accepted students vary across schools.

The senior secondary school concludes with a national matriculation examination that gives students a general qualification to apply for universities or for tertiary-level vocational studies. The examination is compulsory for all senior secondary school students. It is drawn up nationally, and there is a centralized body to grade the exam according to uniform criteria. The results are also standardized to be comparable across the years.

There are four compulsory exams in the matriculation examination: mother tongue, the second official language, one foreign language[2], and either mathematics or science and humanities exam.[3] In addition, candidates may voluntarily take additional exams in other foreign languages, or take both the math, and the science and humanities exams. The exams are held each spring and autumn during a two-week examination period. From 1996 onwards the candidates have been able to take the exam over the maximum of three examination periods. Before, the full exam had to be taken within the same period, usually in the spring term of the senior year.

Responsibility for funding senior secondary schools is divided between the state and the municipal governments. The system was reformed in 1993. Before, the municipal governments received a cost-based subsidy. In the new system, the Ministry of Education confirms each year unit prices (expenditure per student) that are used

[2] For the Finnish-speaking majority, the other national language is Swedish and the compulsory foreign language is usually English.

[3] The science and humanities exam includes questions from physics, chemistry, biology, psychology, geography, religion, and history. Students can choose to answer questions from any subject area.

to calculate the state subsidy. These unit prices are higher in small schools and in the municipalities that offer schooling in both Finnish and Swedish. The state subsidy covers 57 percent of the costs calculated by multiplying unit prices by the number of students. The actual level of school spending is decided by the local governments, the state grant does not depend on actual spending.

3. Data and variables

Our basic sample is drawn from the Employment Statistics (ES). This is the main labor market database of Statistics Finland with information on individual's income, employment status, education, household composition, etc. Data are based on about 30 different official registers. Currently the ES cover the years 1987–1997. For our sample we merge additional information from the National Joint Application Register and the Matriculation Examination Board Register. We also add the information on the parents of the sampled individuals from the ES database.

The Matriculation Examination Board Register contains the grades in all subject exams and the time of examination for all students attending the exam. We use data for the years 1990–1998. We defined the year of matriculation examination as the year when the student has taken all four compulsory exams (not necessarily passed). If the student has later supplemented his/her examination, we have added the new scores in the results, but if a student has attended again an exam that she has already passed or failed, we have used the score from the first attempt at this exam. Our sample includes only students who have attended the matriculation examination; we were not able to identify drop-outs from those who have been accepted to senior secondary school, but have never started it.

In this study, we use the sum of test scores in six exams as a dependent variable. The scores in each exam range from improbatur (failed) to laudatur (excellent). To calculate the overall score the sub-scores in each exam are converted to a scale from 0 to 6.[4] About 10 per cent of the candidates reach excellent (5.5–6.0) average test scores each year. Less than 5 per cent of the candidates fail the examination.

Achievement level of the students at the point when they enter senior secondary school can be found from the National Joint Application Register. We had data from the years 1987–1998 with information on students' grade point average (GPA) in the comprehensive school. Grading scale in the comprehensive school is from 4 (fail) to 10. Some senior secondary schools accept all applicants regardless of their GPA, while for some schools the minimum GPA requirement is over 9.[5] The average GPA in our sample is 8.5.

We measure the students' family background by the parents' education using the ES data. The parents' education is recorded at the time when the students enter secondary school. We converted the level of education into years of schooling using the mean years of schooling usually required to complete different levels of education. We have also municipality-level information on the education level of the population and information on unemployment rates at the local labor market.

Other student-related information include the student's gender and whether the student worked during the school year. Work during the school year is defined based on the months worked in the ES data. We assume that if a student has worked for more than two months, she must have been working not only during the summer holidays but also during the school year.

We have cost information on 444 senior secondary schools from 276 municipalities. Data cover all day schools in Finland. Roughly three quarters of the total school expenditure is teaching expenditure, which comprises teachers' salaries, teaching materials, teacher training, and other teaching-related costs. Data on the teaching expenditure for the years 1987–1992 are collected from the educational expenditure registers of the National Board of Education. Data for years 1993–1997 are based on the municipality databases of Statistics Finland. In general, our cost data is very accurate. The main problem in data is that if several senior secondary schools are located in the same municipality, the costs cannot be attributed to the individual schools. Since most municipalities in Finland have only one senior secondary school, this is an issue that concerns mainly larger cities. To ensure that our results are not affected by the aggregation level, we perform the empirical analyses also with a smaller sample of municipalities with only one senior secondary school.

The expenditure data include information on teaching expenditure and other costs both as absolute figures and as average costs per student. We deflated all cost variables to the 1997 prices.[6] Our principal measure of school resources is teaching expenditure per student, cal-

[4] In 1996, laudatur was split into eximia cum laude approbatur and laudatur. To maintain comparability across years, we coded both eximia and laudatur as 6.

[5] As the fraction of the age cohort admitted to senior secondary schools has increased over the years, admission has become less selective. This causes a negative trend in the comprehensive school GPA of the accepted students.

[6] Wages, pension expenditures, and social security payments are deflated using the wage and salary earnings index of the municipal workers. Other costs are deflated using the cost-of-living index.

culated as an average over the three years that the student was enrolled in the school. We consider teaching expenditure a summary measure of class size, teaching hours, and teachers' experience. It should be noted that teacher salaries are based on a nationwide union contract, and there is no across-school variation in salaries. Therefore, the variation in teaching expenditure per student depends almost entirely on the number of teaching hours per student.

The two data sets—the one including the students and the other including the schools—were merged by Statistics Finland. In order to prevent identifying individual students from the data, the identity of the schools was concealed in the merging process. From the 23,199 students whose examination results we collected from the Matriculation Examination Board Register, we were able to match school and background information for 20,505 students.

4. Developments in the school resources

The average teaching expenditure per student calculated over the period from 1987 to 1997 was FIM 16,000.[7] As shown in Fig. 1, there has been substantial variation both between schools and over time. Average expenditure per student peaked in 1989. After that it decreased by 25 percent in five years. The decline was due to a severe recession that reduced tax revenue, increased social expenditure for the unemployed, and forced local governments to cut spending in all other activities. Fig. 1 also shows that there is large variation in expenditure across schools in any given year. The average difference between the 1st and the 5th quintile is FIM 8600. The most important factor explaining the between-school variation in spending is the school size. The schools with less than 100 students have considerably higher teaching expenditure per student than the larger schools.

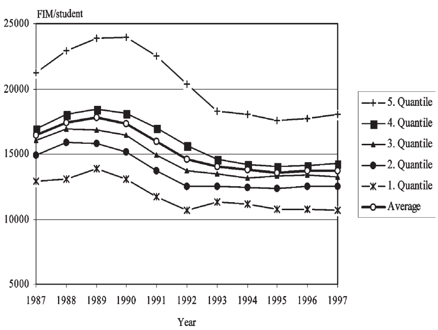


Fig. 1. Changes in teaching expenditure per student.

We also demonstrate that there are large differences across schools in the changes of expenditure per student. Fig. 2 shows changes between 1989 and 1994, i.e., during the largest drop in the average spending. There were only a few schools that did not experience any decline in expenditure. In thirty schools, teaching expenditure decreased by more than FIM 10,000 per student. On average, small schools cut their expenditures more than large schools, but there is considerable variation in the spending changes, even controlling for the school size.

Detailed information on how these spending cuts were achieved is not available in our data. The National Board of Education reports that savings were made by increasing the class size, and by decreasing the supply of voluntary courses and remedial instruction. Teacher salaries have been decreased by cutting holiday benefits, and reducing compensation for other than teaching duties. Temporary substitute teachers were no longer hired. Also, spending on teacher training and teaching materials was reduced.

Eventually, we will use the teaching expenditure per student to explain student achievement. However, simply regressing achievement on resources leads to biased estimates on the effect of resources if the resources are correlated with other factors that affect student achievement. To examine the differences in school spending, we regress the teaching expenditure per student on the local government tax revenue, the local unemployment rate, the average education level in the municipality, and the municipality's gross margin (tax revenues + state subsidies – spending). The equation also includes yearly dummies interacted with the inverse of the number of students in the municipality, which allows for a different spending rule in each year. Since school subsidies are specified as marks per student, and the amount per student varies across years, it is important not to restrict the effect of the number of students to be equal across the years.

Another way to interpret our regression estimates is that if expenditures depend on the economic situation in the municipality and the year-specific rules on expenditure per student, the total school expenditure can be written as

$$Exp_{it} = \alpha_t + \beta_t N_{it} + \gamma X_{it} + \varepsilon_{it} \quad (1)$$

where Exp_{it} is expenditure in municipality i in year t, N_{it} is the number of students and X_{it} is a measure of economic situation, e.g., the tax revenue of the municipality. α_t and β_t are year-specific coefficients. Dividing all terms by the number of students yields

$$Exp_{it}/N_{it} = \alpha_t \frac{1}{N_{it}} + \beta_t + \gamma X_{it}/N_{it} + \nu_{it}. \quad (2)$$

We first estimate the model with random

[7] 1 FIM (Finnish markka) ≈ 0.17 euro.

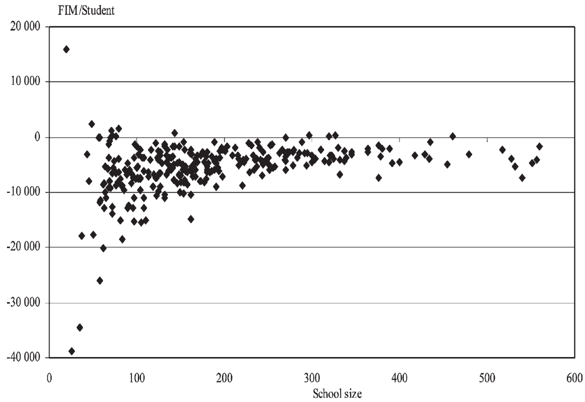


Fig. 2. Change in teaching expenditure by school size in 1989-94. Notes: There were 6 schools with more than 600 students that do not fit into the graph. In these schools changes in expenditures were less than FIM 2000 per student.

(municipality) effects panel regression. The results in the first column of Table 1 show that the economic situation of the local governments matters for school spending. The municipalities with higher tax revenue spend more per student, and the municipalities in high unemployment areas considerably less. Municipality's gross margin and local education level had no significant effects on school spending. However, these variables are highly correlated with the tax revenue. When the tax revenue was not included in the model, both variables had a strong positive effect on the school spending. In column 2, we add an indicator variable for urban area to the model. Apparently urban municipalities have lower per-student teaching expenditure than rural municipalities even when the school size is controlled for.

In column 3, we estimate the model with fixed municipality effects. The fixed effect specification removes all time-invariant differences between municipalities. According to the results, only the unemployment rate has a significant effect on school spending. The municipalities that experienced the highest increase in unemployment cut their school spending by the largest amount.

Above, we showed that in a cross-section, school resources are correlated with the size of the school, local education level, local unemployment rate, municipality tax revenue and population density. However, changes in spending per student appear to be correlated only with changes in unemployment. Therefore, we are reasonably confident that changes in spending are related to the financial situation of local government, and can be treated as exogenous in explaining changes in student achievement.

5. Results

The main question of this paper is whether the reduction in the school spending that occurred during the

Table 1
Determinants of teaching expenditure[a]

	RE (1)	RE (2)	FE (3)
Tax revenue / inhabitant	**0.23** (0.09)	**0.27** (0.10)	–0.04 (0.11)
Local unemployment rate	**–65.60** (24.12)	**–64.05** (24.31)	**–208.91 (26.71)**
Municipality's gross margin	–0.07 (0.07)	–0.08 (0.07)	0.01 (0.07)
Municipality's educational level	–3.28 (8.66)	4.45 (9.95)	
Urban area		–854.62 (563.02)	
Densely populated area		**–1,224.04** (438.17)	
Number of observations	2,464	2,464	2,464
Number of municipalities	274	274	274
R^2	0.74	0.74	0.66

[a] The dependent variable is teaching expenditure per student. All models include year dummies and year dummies interacted with the inverse of the number of students. Tax revenue per inhabitant, unemployment rate and municipality's gross margin are lagged by one year. Standard errors are in parentheses.

1990s has had a negative impact on the matriculation examination results. However, there are two issues that make this analysis more complicated.

First, since the matriculation exam results are standardized across the years, there is no annual variation in the test scores. If the reduction in school spending had been equally large in all schools, the annual variation in school spending could not be used to evaluate the effects of resources on performance. Second, although the effect of spending on performance can be, and often has been, calculated from a single cross-section, it is not clear that cross-section estimates reveal a causal relationship. A positive correlation between school spending and student performance can be due to unobserved differences in schools. For example, parents that are more concerned about their children may get their children into the better schools. Parental involvement may also have a positive effect on learning, creating a correlation between school quality and achievement, even if school quality had no causal effect on learning. Even with data on the initial achievement level and the family background of the students, there is no way to be certain that the partial correlation between school resources and student achievement is a causal effect and not a spurious association.

Since we have repeated observations from each school, we can use panel data methods to control for the school-specific time-invariant factors. Our approach involves explaining the matriculation exam score with the grade point average in the comprehensive school, measures of family background, and school spending:

$$A_{ist} = \alpha + \beta R_{st} + \chi F_{ist} + \delta GPA_{ist} + u_s + v_t + \varepsilon_{ist} \tag{3}$$

where A is the matriculation exam score, R is a measure of school resources averaged over the three years that the student was in school, F is a vector of family background characteristics, and GPA is the grade point average in the comprehensive school. Index i refers to individual, s to school and t to time. We use a two-way fixed-effect panel estimator. Therefore, u_s is a fixed school effect, v_t a fixed time effect, and ϵ_{ist} a random error term.

Table 2 presents the estimation results of the fixed-effect model. The data cover years 1990–1998. The dependent variable is a sum of test scores in six exams. First, in column 1 we use data from all the 444 senior secondary schools.

As expected, the grade point average in comprehensive schools has a very large effect on the matriculation examination results. A unit increase in the GPA increases the matriculation exam score by almost eight points, which is equivalent to an increase of score in each exam by more than one grade. Also, parents' education has a strong effect on the matriculation examination score, even when the comprehensive school GPA is controlled for. All else equal, boys do slightly better than girls.

The results show no effects of teaching expenditure on student performance. The estimates are positive but not significantly different from zero. The effect is rather precisely estimated. A 95 percent confidence interval for the effect of FIM 1000 (5%) increase in expenditure per student is from –0.04 to 0.08 points in the matriculation exam.

In Column 2 of Table 2, we add to the equation the local unemployment rate, the school size, and a dummy variable that indicates whether the student has been working during the school year. The school size and the local unemployment rate have no effect. Working during the school year appears to decrease the exam scores. However, adding these variables has no impact on the coefficients of teaching expenditure, family background or GPA. Altogether, the variables in the model explain 53 percent of the variance in the matriculation examination scores. About 10 percent of the remaining variance is attributed to the between-school variation.

As noted before, it is not possible to divide teaching expenditure between different schools in the same municipality. Therefore, in Columns 3 and 4 of Table 2, we repeated the analysis using only private schools and the municipalities with only one senior secondary school. The results on this smaller sample are similar to the estimates including all schools. The only qualitative differences are that the coefficient of the school size decreases and becomes statistically significant, and that boys are not performing better any more.

We experimented with a number of different specifications to check the robustness of our results. First, we replaced the sum of test scores with the average score in the compulsory exams in order to focus on the tests that all students take. Second, we followed a suggestion by a referee and dropped the observations where per-student expenditure was below the 5th or above the 95th percentile to lessen the influence of extreme observations. Third, we estimated the effect of school spending on the 10th, 25th, 50th, 75th, and 90th quantiles of the achievement distribution. As argued by Eide and Showalter (1998) it is possible that cuts in spending hurt the weakest students even if average achievement is unaltered. Finally, we estimated the effect of resources on the various sub-scores of the exam. We were particularly interested whether the cuts in resources would affect scores in voluntary exams as voluntary courses are more likely targets for the spending cuts than the "core" courses. None of these specifications produced significant estimates for the effect of teaching expenditure.[8]

[8] Results on the robustness checks are not reported here. Full results are available from the authors. More details can also be found in our discussion paper (Häkkinen, Kirjavainen and Uusitalo (2000)).

Table 2
Determinants of student achievement[a]

	(1)	(2)	(3)	(4)
Comprehensive school GPA	**7.89 (0.06)**	**7.90 (0.06)**	**7.90 (0.09)**	**7.91 (0.09)**
Mother's education	**0.32 (0.02)**	**0.32 (0.02)**	**0.30 (0.03)**	**0.30 (0.03)**
Father's education	**0.19 (0.01)**	**0.19 (0.01)**	**0.18 (0.02)**	**0.18 (0.02)**
Teaching exp./student(in thousands)	0.02 (0.03)	0.02 (0.03)	0.02 (0.04)	0.006 (0.04)
Male	**0.15 (0.07)**	**0.16 (0.07)**	–0.08 (0.11)	–0.07 (0.11)
Work during senior secondary school	-	**–0.62 (0.10)**	-	**–0.52 (0.14)**
Unemployment rate	-	–0.03 (0.02)	-	0.003 (0.03)
Number of students in school	-	–0.0002 (0.0002)	-	**–0.003 (0.001)**
Number of observations	20,505	20,504	9,442	9,442
Number of schools	444	444	245	245
R^2	0.530	0.531	0.535	0.536

[a] Dependent variable is sum of test scores in six exams (mother tongue, the other national language, mathematics, compulsory foreign language, additional foreign language, science and humanities). Standard errors corrected for the school-year clustering are in parentheses. In columns (3) and (4) only private schools and municipalities with one school are included. Work during senior secondary school is a dummy variable indicating that the student worked during the school year (1 if work months >2, 0 otherwise). All columns include the year dummies and the school fixed effects.

6. Conclusion

Our results are in accord with much of the earlier research on the effects of school resources. Students' family background and earlier achievement have a large effect on the matriculation examination results. As for the school resources, we found no significant effects on any of the exam results. This conclusion was not sensitive to changes in the model specification or on the choice of the dependent variable. At least in the short-term, the schools have been able to reduce their spending without significant decreases in the test scores. This does not imply that resources would be completely irrelevant; schools may save on other activities, and focus on teaching in the subjects that are included in matriculation examination. As some cuts in average spending were related to increases in the school size, it may be that schools have become more effective due to increasing returns to scale. Other possible explanations include that teachers have exerted extra effort in the short-run, or that the test standards have been lowered. Nevertheless, our results show that at least the worst fears on the consequences of cutting resources from schools are not supported by the empirical facts.

References

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, *114*, 533–575.

Betts, J. (1996). Is there a link between school inputs and earnings? Fresh scrutiny of an old literature. In G. Burtless (Ed.), *Does money matter? The effect of school resources on student achievement and adult success*. Washington DC: Brookings Institution Press.

Card, D., & Krueger, A. (1996). School resources and student outcomes: an overview of the literature and new evidence from North and South Carolina. *Journal of Economic Perspectives*, *10*, 31–40.

Eide, E., & Showalter, M. H. (1998). The effect of school quality on student performance: a quantile regression approach. *Economics Letters*, *58*, 345–350.

Häkkinen, I., Kirjavainen T. & Uusitalo, R. (2000). School resources and student achievement revisited: new evidence using panel data. *Government Institute for Economic Research*, *Discussion Papers* 227. Helsinki: VATT.

Hanushek, E. (1986). The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, *24*, 1141–1177.

Hanushek, E. (1997). Assessing the effects of school resources on student performance: an update. *Educational Evaluation and Policy Analysis*, *19*, 141–164.

Hedges, L., & Greenwald, R. (1996). Have times changed? The relation between school resources and student performance. In G. Burtless (Ed.), *Does money matter? The effect of school resources on student achievement and adult success*. Washington DC: Brookings Institution Press.

Hoxby, C. (1998). The effects of class size and composition on student achievement: new evidence from natural population variation. *NBER Working Paper* 6869.

Krueger, A. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, *114*, 497–532.

HELSINGIN KAUPPAKORKEAKOULUN JULKAISUJA
Publications of the Helsinki School of Economics

A-SARJA: VÄITÖSKIRJOJA - DOCTORAL DISSERTATIONS. ISSN 1237-556X.

A:287. TIMO JÄRVENSIVU: Values-driven management in strategic networks: A case study of the influence of organizational values on cooperation. 2007.
ISBN-10: 952-488-081-4, ISBN-13: 978-952-488-081-7.

A:288. PETRI HILLI: Riskinhallinta yksityisen sektorin työeläkkeiden rahoituksessa. 2007.
ISBN-10: 952-488-085-7, ISBN-13: 978-952-488-085-5.
E-version: ISBN 978-952-488-110-4.

A:289. ULLA KRUHSE-LEHTONEN: Empirical Studies on the Returns to Education in Finland. 2007. ISBN 978-952-488-089-3, E-version ISBN 978-952-488-091-6.

A:290. IRJA HYVÄRI: Project Management Effectiveness in Different Organizational Conditions. 2007. ISBN 978-952-488-092-3, E-version: 978-952-488-093-0.

A:291. MIKKO MÄKINEN: Essays on Stock Option Schemes and CEO Compensation. 2007.
ISBN 978-952-488-095-4.

A:292. JAAKKO ASPARA: Emergence and Translations of Management Interests in Corporate Branding in the Finnish Pulp and Paper Corporations. A Study with an Actor-Network Theory Approach. 2007. ISBN 978-952-488-096-1, E-version: 978-952-488-107-4.

A:293. SAMI J. SARPOLA: Information Systems in Buyer-supplier Collaboration. 2007.
ISBN 978-952-488-098-5.

A:294. SANNA K. LAUKKANEN: On the Integrative Role of Information Systems in Organizations: Observations and a Proposal for Assessment in the Broader Context of Integrative Devices. 2006. ISBN 978-952-488-099-2.

A:295. CHUNYANG HUANG: Essays on Corporate Governance Issues in China. 2007.
ISBN 978-952-488-106-7, E-version: 978-952-488-125-8.

A:296. ALEKSI HORSTI: Essays on Electronic Business Models and Their Evaluation. 2007.
ISBN 978-952-488-117-3, E-version: 978-952-488-118-0.

A:297. SARI STENFORS: Strategy tools and strategy toys: Management tools in strategy work. 2007. ISBN 978-952-488-120-3, E-version: 978-952-488-130-2.

A:298. PÄIVI KARHUNEN: Field-Level Change in Institutional Transformation: Strategic Responses to Post-Socialism in St. Petersburg Hotel Enterprises. 2007.
ISBN 978-952-488-122-7, E-version: 978-952-488-123-4.

A:299. EEVA-KATRI AHOLA: Producing Experience in Marketplace Encounters: A Study of Consumption Experiences in Art Exhibitions and Trade Fairs. 2007.
ISBN 978-952-488-126-5.

A:300. HANNU HÄNNINEN: Negotiated Risks: The Estonia Accident and the Stream of Bow Visor Failures in the Baltic Ferry Traffic. 2007. ISBN 978-952-499-127-2.

A-301. MARIANNE KIVELÄ: Dynamic Capabilities in Small Software Firms. 2007. ISBN 978-952-488-128-9.

A:302. OSMO T.A. SORONEN: A Transaction Cost Based Comparison of Consumers' Choice between Conventional and Electronic Markets. 2007. ISBN 978-952-488-131-9.

A:303. MATTI NOJONEN: Guanxi – The Chinese Third Arm. 2007. ISBN 978-952-488-132-6.

A:304. HANNU OJALA: Essays on the Value Relevance of Goodwill Accounting. 2007. ISBN 978-952-488-133-3, E-version: 978-952-488-135-7.

A:305. ANTTI KAUHANEN: Essays on Empirical Personnel Economics. 2007. ISBN 978-952-488-139-5.

A:306. HANS MÄNTYLÄ: On "Good" Academic Work – Practicing Respect at Close Range. 2007. ISBN 978,952-488-1421-8, E-version: 978-952-488-142-5.

A:307. MILLA HUURROS: The Emergence and Scope of Complex System/Service Innovation. The Case of the Mobile Payment Services Market in Finland. 2007. ISBN 978-952-488-143-2

A:308. PEKKA MALO: Higher Order Moments in Distribution Modelling with Applications to Risk Management. 2007. ISBN 978-952-488-155-5, E-version: 978-952-488-156-2.

A:309. TANJA TANAYAMA: Allocation and Effects of R&D Subsidies: Selection, Screening, and Strategic Behavior. 2007. ISBN 978-952-488-157-9, E-version: 978-952-488-158-6.

A:310. JARI PAULAMÄKI: Kauppiasyrittäjän toimintavapaus ketjuyrityksessä. Haastattelututkimus K-kauppiaan kokemasta toimintavapaudesta agenttiteorian näkökulmasta. 2008. Korjattu painos. ISBN 978-952-488-246-0, E-version: 978-952-488-247-7.

A:311. JANNE VIHINEN: Supply and Demand Perspectives on Mobile Products and Content Services. ISBN 978-952-488-168-5.

A:312. SAMULI KNÜPFER: Essays on Household Finance. 2007. ISBN 978-952-488-178-4.

A:313. MARI NYRHINEN: The Success of Firm-wide IT Infrastructure Outsourcing: an Integrated Approach. 2007. ISBN 978-952-488-179-1.

A:314. ESKO PENTTINEN: Transition from Products to Services within the Manufacturing Business. 2007. ISBN 978-952-488-181-4, E-version: 978-952-488-182-1.

A:315. JARKKO VESA: A Comparison of the Finnish and the Japanese Mobile Services Markets: Observations and Possible Implications. 2007. ISBN 978-952-488-184-5.

A:316. ANTTI RUOTOISTENMÄKI: Condition Data in Road Maintenance Management. 2007. ISBN 978-952-488-185-2, E-version: 978-952-488-186-9.

A:317. NINA GRANQVIST: Nanotechnology and Nanolabeling. Essays on the Emergence of New Technological Fields. 2007. ISBN 978-952-488-187-6, E-version: 978-952-488-188-3.

A:318. GERARD L. DANFORD: INTERNATIONALIZATION: An Information-Processing Perspective. A Study of the Level of ICT Use During Internationalization. 2007. ISBN 978-952-488-190-6.

A:319. TIINA RITVALA: Actors and Institutions in the Emergence of a New Field: A Study of the Cholesterol-Lowering Functional Foods Market. 2007. ISBN 978-952-488-195-1.

A:320. JUHA LAAKSONEN: Managing Radical Business Innovations. A Study of Internal Corporate Venturing at Sonera Corporation. 2007.
ISBN 978-952-488-201-9, E-version: 978-952-488-202-6.

A:321. BRETT FIFIELD: A Project Network: An Approach to Creating Emergent Business. 2008.
ISBN 978-952-488-206-4, E-version: 978-952-488-207-1.

A:322. ANTTI NURMI: Essays on Management of Complex Information Systems Development Projects. 2008. ISBN 978-952-488-226-2.

A:323. SAMI RELANDER: Towards Approximate Reasoning on New Software Product Company Success Potential Estimation. A Design Science Based Fuzzy Logic Expert System.
2008. ISBN 978-952-488-227-9.

A:324. SEPPO KINKKI: Essays on Minority Protection and Dividend Policy. 2008.
ISBN 978-952-488-229-3.

A:325. TEEMU MOILANEN: Network Brand Management: Study of Competencies of Place Branding Ski Destinations. 2008. ISBN 978-952-488-236-1.

A:326. JYRKI ALI-YRKKÖ: Essays on the Impacts of Technology Development and R&D Subsidies. 2008. ISBN 978-952-488-237-8.

A:327. MARKUS M. MÄKELÄ: Essays on software product development. A Strategic management viewpoint. 2008. ISBN 978-952-488-238-5.

A:328. SAMI NAPARI: Essays on the gender wage gap in Finland. 2008.
ISBN 978-952-488-243-9.

A:329. PAULA KIVIMAA: The innovation effects of environmental policies. Linking policies, companies and innovations in the Nordic pulp and paper industry.
2008. ISBN 978-952-488-244-6.

A:330. HELI VIRTA: Essays on Institutions and the Other Deep Determinants of Economic Development. 2008. ISBN 978-952-488-267-5.

A:331. JUKKA RUOTINEN: Essays in trade in services difficulties and possibilities.
2008. ISBN 978-952-488-271-2, E-version: ISBN 978-952-488-272-9.

A:332. IIKKA KORHONEN: Essays on commitment and government debt structure.
2008. ISBN 978-952-488-273-6, E-version: ISBN 978-952-488-274-3.

A:333. MARKO MERISAVO: The interaction between digital marketing communication and customer loyalty. 2008. ISBN 978-952-488-277-4, E-version 978-952-488-278-1.

A:334. PETRI ESKELINEN: Reference point based decision support tools for interactive multiobjective optimization. 2008. ISBN 978-952-488-282-8.

A:335. SARI YLI-KAUHALUOMA: Working on technology: a study on collaborative R&D work in industrial chemistry. 2008. ISBN 978-952-488-284-2

A:336. JANI KILPI: Sourcing of availability services - case aircraft component support. 2008. ISBN 978-952-488-284-2, 978-952-488-286-6 (e-version).

A:337. HEIDI SILVENNOINEN: Essays on household time allocation decisions in a collective household model. 2008. ISBN 978-952-488-290-3, ISBN 978-952-488-291-0 (e-version).

A:338. JUKKA PARTANEN: Pk-yrityksen verkostokyvykkyydet ja nopea kasvu - case: Tiede- ja teknologiavetoiset yritykset. 2008. ISBN 978-952-488-295-8.

A:339. PETRUS KAUTTO: Who holds the reins in Integrated Product Policy? An individual company as a target of regulation and as a policy maker. 2008. ISBN 978-952-488-300-9, 978-952-488-301-6 (e-version).

A:340. KATJA AHONIEMI: Modeling and Forecasting Implied Volatility. 2009. ISBN 978-952-488-303-0, E-version: 978-952-488-304-7.

A:341. MATTI SARVIMÄKI: Essays on Migration. 2009. ISBN 978-952-488-305-4, 978-952-488-306-1 (e-version).

A:342. LEENA KERKELÄ: Essays on Globalization – Policies in Trade, Development, Resources and Climate Change. 2009. ISBN 978-952-488-307-8, E-version: 978-952-488-308-5.

A:343. ANNELI NORDBERG: Pienyrityksen dynaaminen kyvykkyys - Empiirinen tutkimus graafisen alan pienpainoyrityksistä. 2009. ISBN 978-952-488-318-4.

A:344. KATRI KARJALAINEN: Challenges of Purchasing Centralization – Empirical Evidence from Public Procurement. 2009. ISBN 978-952-488-322-1, E-version: 978-952-488-323-8.

A:345. Jouni H. Leinonen: Organizational Learning in High-Velocity Markets. Case Study in The Mobile Communications Industry. 2009. ISBN 978-952-488-325-2.

A:346. Johanna Vesterinen: Equity Markets and Firm Innovation in Interaction.
- A Study of a Telecommunications Firm in Radical Industry Transformation. 2009. ISBN 978-952-488-327-6.

A:347. Jari Huikku: Post-Completion Auditing of Capital Investments and Organizational Learning. 2009. ISBN 978-952-488-334-4, E-version: 978-952-488-335-1.

A:348. TANJA KIRJAVAINEN: Essays on the Efficiency of Schools and Student Achievement. 2009. ISBN 978-952-488-336-8, E-version: 978-952-488-337-5.

B-SARJA: TUTKIMUKSIA - RESEARCH REPORTS. ISSN 0356-889X.

B:77. MATTI KAUTTO – ARTO LINDBLOM – LASSE MITRONEN: Kaupan liiketoimintaosaaminen. 2007. ISBN 978-952-488-109-8.

B:78. NIILO HOME: Kauppiasyrittäjyys. Empiirinen tutkimus K-ruokakauppiaiden yrittäjyysasenteista. Entrepreneurial Orientation of Grocery Retailers – A Summary. ISBN 978-952-488-113-5, E-versio: ISBN 978-952-488-114-2.

B:79. PÄIVI KARHUNEN – OLENA LESYK – KRISTO OVASKA: Ukraina suomalaisyritysten toimintaympäristönä. 2007. ISBN 978-952-488-150-0, E-versio: 978-952-488-151-7.

B:80. MARIA NOKKONEN: Näkemyksiä pörssiyhtiöiden hallitusten sukupuolikiintiöistä. Retorinen diskurssianalyysi Helsingin Sanomien verkkokeskusteluista. Nasta-projekti. 2007. ISBN 978-952-488-166-1, E-versio: 978-952-488-167-8.

B:81. PIIA HELISTE – RIITTA KOSONEN – MARJA MATTILA: Suomalaisyritykset Baltiassa tänään ja huomenna: Liiketoimintanormien ja -käytäntöjen kehityksestä. 2007. ISBN 978-952-488-177-7, E-versio: 978-952-488-183-8.

B:82. OLGA MASHKINA – PIIA HELISTE – RIITTA KOSONEN: The Emerging Mortgage Market in Russia: An Overview with Local and Foreign Perspectives. 2007. ISBN 978-952-488-193-7, E-version: 978-952-488-194-4.

B:83. PIIA HELISTE – MARJA MATTILA – KRZYSZTOF STACHOWIAK: Puola suomalaisyritysten toimintaympäristönä. 2007. ISBN 978-952-488-198-2, E-versio: 978-952-488-199-9.

B:84. PÄIVI KARHUNEN – RIITTA KOSONEN – JOHANNA LOGRÉN – KRISTO OVASKA: Suomalaisyritysten strategiat Venäjän muuttuvassa liiketoimintaympäristössä. 2008. ISBN 978-953-488-212-5, E-versio: 978-952-488-241-5.

B:85. MARJA MATTILA – EEVA KEROLA – RIITTA KOSONEN: Unkari suomalaisyritysten toimintaympäristönä. 2008. ISBN 978-952-488-213-2, E-versio: 978-952-488-222-4.

B:86. KRISTIINA KORHONEN – ANU PENTTILÄ – MAYUMI SHIMIZU – EEVA KEROLA – RIITTA KOSONEN: Intia suomalaisyritysten toimintaympäristönä.2008. ISBN 978-952-488-214-9, E-versio: 978-952-488-283-5

B:87. SINIKKA VANHALA – SINIKKA PESONEN: Työstä nauttien. SEFE:en kuuluvien nais- ja miesjohtajien näkemyksiä työstään ja urastaan. 2008. ISBN 978-952-488-224-8, E-versio: 978-952-488-225-5.

B:88. POLINA HEININEN – OLGA MASHKINA – PÄIVI KARHUNEN – RIITTA KOSONEN: Leningradin lääni yritysten toimintaympäristönä: pk-sektorin näkökulma. 2008. ISBN 978-952-488-231-6, E-versio: 978-952-488-235-4.

B:89. Ольга Машкина – Полина Хейнинен: Влияние государственного сектора на развитие малого и среднего предпринимательства в Ленинградской области: взгляд предприятий.2008. ISBN 978-952-488-233-0, E-version: 978-952-488-240-8.

B:90. MAI ANTTILA – ARTO RAJALA (Editors): Fishing with business nets – keeping thoughts on the horizon Professor Kristian Möller. 2008. ISBN 978-952-488-249-1, E-version: 978-952-488-250-7.

B:91. RENÉ DE KOSTER – WERNER DELFMANN (Editors): Recent developments in supply chain management. 2008. ISBN 978-952-488-251-4, E-version: 978-952-488-252-1.

B:92. KATARIINA RASILAINEN: Valta orkesterissa. Narratiivinen tutkimus soittajien kokemuksista ja näkemyksistä. 2008. ISBN 978-952-488-254-5, E-versio: 978-952-488-256-9.

B:93. SUSANNA KANTELINEN: Opiskelen, siis koen. Kohti kokevan subjektin tunnistavaa korkeakoulututkimusta. 2008. ISBN 978-952-488-257-6, E-versio: 978-952-488-258.

B:94. KATRI KARJALAINEN – TUOMO KIVIOJA – SANNA PELLAVA: Yhteishankintojen kustannusvaikutus. Valtion hankintatoimen kustannussäästöjen selvittäminen. 2008. ISBN 978-952-488-263-7, E-versio: ISBN 978-952-488-264-4.

B:95. ESKO PENTTINEN: Electronic Invoicing Initiatives in Finland and in the European Union – Taking the Steps towards the Real-Time Economy. 2008.
ISBN 978-952-488-268-2, E-versio: ISBN 978-952-488-270-5.

B:96. LIISA UUSITALO (Editor): Museum and visual art markets. 2008.
ISBN 978-952-488-287-3, E-version: ISBN 978-952-488-288-0.

B:97. EEVA-LIISA LEHTONEN: Pohjoismaiden ensimmäinen kauppatieteiden tohtori Vilho Paavo Nurmilahti 1899-1943. 2008. ISBN 978-952-488-292-7,
E-versio: ISBN 978-952-488-293-4.

B:98. ERJA KETTUNEN – JYRI LINTUNEN – WEI LU – RIITTA KOSONEN: Suomalaisyritysten strategiat Kiinan muuttuvassa toimintaympäristössä. 2008 ISBN 978-952-488-234-7,
E-versio: ISBN 978-952-488-297-2.

B:99. SUSANNA VIRKKULA – EEVA-KATRI AHOLA – JOHANNA MOISANDER – JAAKKO ASPARA – HENRIKKI TIKKANEN: Messut kuluttajia osallistavan markkinakulttuurin fasilitaattorina: messukokemuksen rakentuminen Venemessuilla. 2008.
ISBN 978-952-488-298-9, E-versio: ISBN 978-952-488-299-6.

B:100. PEER HULL KRISTENSEN – KARI LILJA (Eds): New Modes of Globalization: Experimentalist Forms of Economics Organization and Enabling Welfare Institutions – Lessons from The Nordic Countries and Slovenia. 2009. ISBN 978-952-488-309-2,
E-version: 978-952-488-310-8.

B:101. VIRPI SERITA – ERIK PÖNTISKOSKI (eds.)
SEPPO MALLENIUS – VESA LEIKOS – KATARIINA VILLBERG – TUUA RINNE – NINA YPPÄRILÄ – SUSANNA HURME: Marketing Finnish Design in Japan. 2009.
ISBN 978-952-488-320-7. E-version: ISBN 978-952-488-321-4.

B:102. POLINA HEININEN – OLLI-MATTI MIKKOLA – PÄIVI KARHUNEN – RIITTA KOSONEN: Yritysrahoitusmarkkinoiden kehitys Venäjällä. Pk-yritysten tilanne Pietarissa. 2009.
ISBN 978-952-488-329-0. E-version: ISBN 978-952-488-331-3.

B:103. ARTO LAHTI: Liiketoimintaosaamisen ja yrittäjyyden pioneeri Suomessa. 2009.
ISBN 978-952-488-330-6.

B:104. KEIJO RÄSÄNEN: Tutkija kirjoittaa - esseitä kirjoittamisesta ja kirjoittajista akateemisessa työssä. 2009. ISBN 978-952-488-332-0. E-versio: ISBN 978-952-488-333-7.

N-SARJA: HELSINKI SCHOOL OF ECONOMICS. MIKKELI BUSINESS CAMPUS PUBLICATIONS.
ISSN 1458-5383

N:63. SOILE MUSTONEN – ANNE GUSTAFSSON-PESONEN: Oppilaitosten yrittäjyyskoulutuksen kehittämishanke 2004–2006 Etelä-Savon alueella. Tavoitteiden, toimenpiteiden ja vaikuttavuuden arviointi. 2007. ISBN: 978-952-488-086-2.

N:64. JOHANNA LOGRÉN – VESA KOKKONEN: Pietarissa toteutettujen yrittäjäkoulutusohjelmien vaikuttavuus. 2007. ISBN 978-952-488-111-1.

N:65. VESA KOKKONEN: Kehity esimiehenä – koulutusohjelman vaikuttavuus. 2007.
ISBN 978-952-488-116-6.

N:66. VESA KOKKONEN – JOHANNA LOGRÉN: Kaupallisten avustajien – koulutusohjelman vaikuttavuus. 2007. ISBN 978-952-488-116-6.

N:67. MARKKU VIRTANEN: Summary and Declaration. Of the Conference on Public Support Systems of SME's in Russia and Other North European Countries. May 18 – 19, 2006, Mikkeli, Finland. 2007. ISBN 978-952-488-140-1.

N:68. ALEKSANDER PANFILO – PÄIVI KARHUNEN: Pietarin ja Leningradin läänin potentiaali kaakkoissuomalaisille metallialan yrityksille. 2007. ISBN 978-952-488-163-0.

N:69. ALEKSANDER PANFILO – PÄIVI KARHUNEN – VISA MIETTINEN: Pietarin innovaatiojärjestelmä jayhteistyöpotentiaali suomalaisille innovaatiotoimijoille. 2007. ISBN 978-952-488-164-7.

N:70. VESA KOKKONEN: Perusta Oma Yritys – koulutusohjelman vaikuttavuus. 2007. ISBN 978-952-488-165-4.

N:71. JARI HANDELBERG – MIKKO SAARIKIVI: Tutkimus Miktech Yrityshautomon yritysten näkemyksistä ja kokemuksista hautomon toiminnasta ja sen edelleen kehittämisestä. 2007. ISBN 978-952-488-175-3.

N:72. SINIKKA MYNTTINEN – MIKKO SAARIKIVI – ERKKI HÄMÄLÄINEN: Mikkelin Seudun yrityspalvelujen henkilökunnan sekä alueen yrittäjien näkemykset ja suhtautuminen mentorointiin. 2007. ISBN 978-952-488-176-0.

N:73. SINIKKA MYNTTINEN: Katsaus K-päivittäistavarakauppaan ja sen merkitykseen Itä-Suomessa. 2007. ISBN 978-952-488-196-8.

N:74. MIKKO SAARIKIVI: Pk-yritysten kansainvälistymisen sopimukset. 2008. ISBN 978-952-488-210-1.

N:75. LAURA TUUTTI: Uutta naisjohtajuutta Delfoi Akatemiasta – hankkeen vaikuttavuus. 2008. ISBN 978-952-488-211-8.

N:76. LAURA KEHUSMAA – JUSSI KÄMÄ – ANNE GUSTAFSSON-PESONEN (ohjaaja): StuNet -Business Possibilities and Education - hankkeen arviointi. 2008. ISBN 978-952-488-215-6.

N:77. PÄIVI KARHUNEN – ERJA KETTUNEN – VISA MIETTINEN – TIINAMARI SIVONEN: Determinants of knowledge-intensive entrepreneurship in Southeast Finland and Northwest Russia. 2008. ISBN 978-952-488-223-1.

N:78. ALEKSANDER PANFILO – PÄIVI KARHUNEN – VISA MIETTINEN: Suomalais-venäläisen innovaatioyhteistyön haasteet toimijanäkökulmasta. 2008. ISBN 978-952-488-232-3.

N:79. VESA KOKKONEN: Kasva Yrittäjäksi – koulutusohjelman vaikuttavuus. 2008. ISBN 978-952-488-248-4.

N:80. VESA KOKKONEN: Johtamisen taidot - hankkeessa järjestettyjen koulutusohjelmien vaikuttavuus. 2008. ISBN 978-952-488-259-0.

N:81. MIKKO SAARIKIVI: Raportti suomalaisten ja brittiläisten pk-yritysten yhteistyön kehittämisestä uusiutuvan energian sektorilla. 2008. ISBN 978-952-488-260-6.

N:82. MIKKO SAARIKIVI – JARI HANDELBERG – TIMO HOLMBERG – ARI MATILAINEN: Selvitys lujitemuovikomposiittituotteiden mahdollisuuksista rakennusteollisuudessa. 2008. ISBN 978-952-488-262-0.

N:83. PÄIVI KARHUNEN – SVETLANA LEDYAEVA – ANNE GUSTAFSSON-PESONEN – ELENA MOCHNIKOVA – DMITRY VASILENKO: Russian students' perceptions of entrepreneurship. Results of a survey in three St. Petersburg universities. Entrepreneurship development –project 2. 2008. ISBN 978-952-488-280-4.

W-SARJA: TYÖPAPEREITA - WORKING PAPERS . ISSN 1235-5674.
ELECTRONIC WORKING PAPERS, ISSN 1795-1828.

W:412. LOTHAR THIELE – KAISA MIETTINEN – PEKKA J. KORHONEN – JULIAN MOLINA: A Preference-Based Interactive Evolutionary Algorithm for Multiobjective Optimization. 2007. ISBN 978-952-488-094-7.

W:413. JAN-ERIK ANTIPIN – JANI LUOTO: Are There Asymmetric Price Responses in the Euro Area? 2007. ISBN 978-952-488-097-8.

W:414. SAMI SARPOLA: Evaluation Framework for VML Systems. 2007. ISBN 978-952-488-097-8.

W:415. SAMI SARPOLA: Focus of Information Systems in Collaborative Supply Chain Relationships. 2007. ISBN 978-952-488-101-2.

W:416. SANNA LAUKKANEN: Information Systems as Integrative Infrastructures. Information Integration and the Broader Context of Integrative and Coordinative Devices. 2007. ISBN 978-952-488-102-9.

W:417. SAMULI SKURNIK – DANIEL PASTERNACK: Uusi näkökulma 1900-luvun alun murroskauteen ja talouden murrosvaiheiden dynamiikkaan. Liikemies Moses Skurnik osakesijoittajana ja -välittäjänä. 2007. ISBN 978-952-488-104-3.

W:418. JOHANNA LOGRÉN – PIIA HELISTE: Kymenlaakson pienten ja keskisuurten yritysten Venäjä-yhteistyöpotentiaali. 2001. ISBN 978-952-488-112-8.

W:419. SARI STENFORS – LEENA TANNER: Evaluating Strategy Tools through Activity Lens. 2007. ISBN 978-952-488-120-3.

W:420. RAIMO LOVIO: Suomalaisten monikansallisten yritysten kotimaisen sidoksen heikkeneminen 2000-luvulla. 2007. ISBN 978-952-488-121-0.

W:421. PEKKA J. KORHONEN – PYRY-ANTTI SIITARI: A Dimensional Decomposition Approach to Identifying Efficient Units in Large-Scale DEA Models. 2007. ISBN 978-952-488-124-1.

W:422. IRYNA YEVSEYEVA – KAISA MIETTINEN – PEKKA SALMINEN – RISTO LAHDELMA: SMAA-Classification - A New Method for Nominal Classification. 2007. ISBN 978-952-488-129-6.

W:423. ELINA HILTUNEN: The Futures Window – A Medium for Presenting Visual Weak Signals to Trigger Employees' Futures Thinking in Organizations. 2007. ISBN 978-952-488-134-0.

W:424. TOMI SEPPÄLÄ – ANTTI RUOTOISTENMÄKI – FRIDTJOF THOMAS: Optimal Selection and Routing of Road Surface Measurements. 2007. ISBN 978-952-488-137-1.

W:425. ANTTI RUOTOISTENMÄKI: Road Maintenance Management System. A Simplified Approach. 2007. ISBN 978-952-488-1389-8.

W:426. ANTTI PIRJETÄ – VESA PUTTONEN: Style Migration in the European Markets 2007. ISBN 978-952-488-145-6.

W:427. MARKKU KALLIO – ANTTI PIRJETÄ: Incentive Option Valuation under Imperfect Market and Risky Private Endowment. 2007. ISBN 978-952-488-146-3.

W:428. ANTTI PIRJETÄ – SEPPO IKÄHEIMO – VESA PUTTONEN: Semiparametric Risk Preferences Implied by Executive Stock Options. 2007. ISBN 978-952-488-147-0.

W:429. OLLI-PEKKA KAUPPILA: Towards a Network Model of Ambidexterity. 2007. ISBN 978-952-488-148-7.

W:430. TIINA RITVALA – BIRGIT KLEYMANN: Scientists as Midwives to Cluster Emergence. An Interpretative Case Study of Functional Foods. 2007. ISBN 978-952-488-149-4.

W:431. JUKKA ALA-MUTKA: Johtamiskyvykkyyden mittaaminen kasvuyrityksissä. 2007. ISBN 978-952-488-153-1.

W:432. MARIANO LUQUE – FRANCISCO RUIZ – KAISA MIETTINEN: GLIDE – General Formulation for Interactive Multiobjective Optimization. 2007. ISBN 978-952-488-154-8.

W:433. SEPPO KINKKI: Minority Protection and Information Content of Dividends in Finland. 2007. ISBN 978-952-488-170-8.

W:434. TAPIO LAAKSO: Characteristics of the Process Supersede Characteristics of the Debtor Explaining Failure to Recover by Legal Reorganization Proceedings. 2007. ISBN 978-952-488-171-5.

W:435. MINNA HALME: Something Good for Everyone? Investigation of Three Corporate Responsibility Approaches. 2007. ISBN 978-952-488-189.

W:436. ARTO LAHTI: Globalization, International Trade, Entrepreneurship and Dynamic Theory of Economics.The Nordic Resource Based View. Part One. 2007. ISBN 978-952-488-191-3.

W:437. ARTO LAHTI: Globalization, International Trade, Entrepreneurship and Dynamic Theory of Economics.The Nordic Resource Based View. Part Two. 2007 ISBN 978-952-488-192-0.

W:438. JANI KILPI: Valuation of Rotable Spare Parts. 2007. ISBN 978-952-488-197-5.

W:439. PETRI ESKELINEN – KAISA MIETTINEN – KATHRIN KLAMROTH – JUSSI HAKANEN: Interactive Learning-oriented Decision Support Tool for Nonlinear Multiobjective Optimization: Pareto Navigator. 2007. ISBN 978-952-488-200-2.

W:440. KALYANMOY DEB – KAISA MIETTINEN – SHAMIK CHAUDHURI: Estimating Nadir Objective Vector: Hybrid of Evolutionary and Local Search. 2008. ISBN 978-952-488-209-5.

W:441. ARTO LAHTI: Globalisaatio haastaa pohjoismaisen palkkatalousmallin. Onko löydettävissä uusia aktiivisia toimintamalleja, joissa Suomi olisi edelleen globalisaation voittaja? 2008. ISBN 978-952-488-216-3.

W:442. ARTO LAHTI: Semanttinen Web – tulevaisuuden internet. Yrittäjien uudet liiketoimintamahdollisuudet. 2008. ISBN 978-952-488-217-0.

W:443. ARTO LAHTI: Ohjelmistoteollisuuden globaali kasvustrategia ja immateriaalioikeudet. 2008. ISBN 978-952-488-218-7.

W:444. ARTO LAHTI: Yrittäjän oikeusvarmuus globaalisaation ja byrokratisoitumisen pyörteissä. Onko löydettävissä uusia ja aktiivisia toimintamalleja yrittäjien syrjäytymisen estämiseksi? 2008. ISBN 978-952-488-219-4.

W:445. PETRI ESKELINEN: Objective trade-off rate information in interactive multiobjective optimization methods – A survey of theory and applications. 2008. ISBN 978-952-488-220-0.

W:446. DEREK C. JONES – PANU KALMI: Trust, inequality and the size of co-operative sector – Cross-country evidence. 2008. ISBN 978-951-488-221-7.

W:447. KRISTIINA KORHONEN – RIITTA KOSONEN – TIINAMARI SIVONEN – PASI SAUKKONEN: Pohjoiskarjalaisten pienten ja keskisuurten yritysten Venäjä-yhteistyöpotentiaali ja tukitarpeet. 2008. ISBN 978-952-488-228-6.

W:448. TIMO JÄRVENSIVU – KRISTIAN MÖLLER: Metatheory of Network Management: A Contingency Perspective. 2008. ISBN 978-952-488-231-6.

W:449. PEKKA KORHONEN: Setting “condition of order preservation” requirements for the priority vector estimate in AHP is not justified. 2008. ISBN 978-952-488-242-2.

W:450. LASSE NIEMI – HANNU OJALA – TOMI SEPPÄLÄ: Misvaluation of takeover targets and auditor quality. 2008. ISBN 978-952-488-255-2.

W:451. JAN-ERIK ANTIPIN – JANI LUOTO: Forecasting performance of the small-scale hybrid New Keynesian model. 2008. ISBN 978-952-488-261-3.

W:452. MARKO MERISAVO: The Interaction between Digital Marketing Communication and Customer Loyalty. 2008. ISBN 978-952-488-266-8.

W:453. PETRI ESKELINEN – KAISA MIETTINEN: Trade-off Analysis Tool with Applicability Study for Interactive Nonlinear Multiobjective Optimization. 2008. ISBN 978-952-488-269-9.

W:454. SEPPO IKÄHEIMO – VESA PUTTONEN – TUOMAS RATILAINEN: Antitakeover provisions and performance – Evidence from the Nordic countries. 2008. ISBN 978-952-488-275-0.

W:455. JAN-ERIK ANTIPIN: Dynamics of inflation responses to monetary policy in the EMU area. 2008. ISBN 978-952-488-276-7.

W:456. KIRSI KOMMONEN: Narratives on Chinese colour culture in business contexts. The Yin Yang Wu Xing of Chinese values. 2008. ISBN 978-952-488-279-8.

W:457. MARKKU ANTTONEN – MIKA KUISMA – MINNA HALME – PETRUS KAUTTO: Materiaalitehokkuuden palveluista ympäristömyötäistä liiketoimintaa (MASCO2). 2008. ISBN 978-952-488-279-8.

W:458. PANU KALMI – DEREK C. JONES – ANTTI KAUHANEN: Econometric case studies: overview and evidence from recent finnish studies. 2008. ISBN 978-952-488-289-7.

W:459. PETRI JYLHÄ – MATTI SUOMINEN – JUSSI-PEKKA LYYTINEN: Arbitrage Capital and Currency Carry Trade Returns. 2008. ISBN 978-952-488-294-1.

W:460. OLLI-MATTI MIKKOLA – KATIA BLOIGU – PÄIVI KARHUNEN: Venäjä-osaamisen luonne ja merkitys kansainvälisissä suomalaisyrityksissä. 2009. ISBN 978-952-488-302-3.

W:461. ANTTI KAUHANEN – SATU ROPONEN: Productivity Dispersion: A Case in the Finnish Retail Trade. 2009. ISBN 978-952-488-311-5.

W:462. JARI HUIKKU: Design of a Post-Completion Auditing System for Organizational Learning. 2009. ISBN 978-952-488-312-2.

W:463. PYRY-ANTTI SIITARI: Identifying Efficient Units in Large-Scale Dea Models Using Efficient Frontier Approximation. 2009. ISBN 978-952-488-313-9.

W:464. MARKKU KALLIO – MERJA HALME: Conditions for Loss Averse and Gain Seeking Consumer Price Behavior. 2009. ISBN 978-952-488-314-6.

W:465. MERJA HALME – OUTI SOMERVUORI: Study of Internet Material Use in Education in Finland. 2009. ISBN 978-952-488-315-3.

W:466. RAIMO LOVIO: Näkökulmia innovaatiotoiminnan ja –politiikan muutoksiin 2000-luvulla. 2009. ISBN 978-952-488-316-0.

Z-SARJA: HELSINKI SCHOOL OF ECONOMICS.
CENTRE FOR INTERNATIONAL BUSINESS RESEARCH. CIBR WORKING PAPERS. ISSN 1235-3931.

Z:16. PETER GABRIELSSON – MIKA GABRIELSSON: Marketing Strategies for Global Expansion in the ICT Field. 2007. ISBN 978-952-488-105-0.

Z:17. MIKA GABRIELSSON – JARMO ERONEN – JORMA PIETALA: Internationalization and Globalization as a Spatial Process. 2007. ISBN 978-952-488-136-4.

Kaikkia Helsingin kauppakorkeakoulun julkaisusarjassa ilmestyneitä julkaisuja voi tilata osoitteella:

KY-Palvelu Oy
Kirjakauppa
Runeberginkatu 14-16
00100 Helsinki
Puh. (09) 4313 8310, fax (09) 495 617
Sähköposti: kykirja@ky.hse.fi

Helsingin kauppakorkeakoulu
Julkaisutoimittaja
PL 1210
00101 Helsinki
Puh. (09) 4313 8579, fax (09) 4313 8305
Sähköposti: julkaisu@hse.fi

All the publications can be ordered from

Helsinki School of Economics
Publications officer
P.O.Box 1210
FIN-00101 Helsinki
Phone +358-9-4313 8579, fax +358-9-4313 8305
E-mail: julkaisu@hse.fi