The Relative Efficiency of Swedish Secondary Schools
An estimation using Stochastic Frontier Analysis
Johan Holmberg
“Roy [singing]: We don't need no education.
Moss: Yes, you do. You've just used a double negative.”
- The IT-crowd (2006)

Acknowledgements:
I would like to thank my supervisor Tommy Lundgren, and Golnaz Amjadi for their guidance and advice during the process of writing this thesis. Your assistance was indispensable for the realisation of this thesis.
Abstract:
The question of public education is of importance for a society and its citizens as education is contributing to the stability of a democratic society and affects the expected future income levels for the individuals receiving it. A significant share of the Swedish GDP is devoted to the provision of educational services which raises the necessity of monitoring the use of these resources. This thesis endeavoured to estimate a production possibility frontier and the relative efficiency of Swedish upper secondary schools. To accomplish this then Stochastic Frontier Analysis was implemented. Data on student results, teacher ratio and students’ socioeconomic characteristics for individual schools during the scholastic years of 2006/2007 through 2015/2016 gathered from the Swedish National Agency for Education was used in this study. The effects of competition on school performance and the relative efficiency of public and independent schools were two factors of interest in the thesis. The results of the analysis point toward a possible positive relationship between local school competition and student results, that Swedish secondary schools could have high average levels of technical efficiency and that the type of the principal organiser might be of minor importance for school efficiency.
# Table of contents

1. Introduction .................................................................................................................. 1

2. Literature review ........................................................................................................ 3
   2.1 Efficiency analysis of schools and school districts .................................................. 3
   2.2 School competition and student results .................................................................... 4

3. Method .......................................................................................................................... 6
   3.1 Estimating a best practice frontier .......................................................................... 6
   3.2 Model and assumptions: ......................................................................................... 8
   3.3 Two-step estimation ............................................................................................... 12

4. Data and model specification ....................................................................................... 12
   4.1 Output data .............................................................................................................. 13
   4.2 Input data ................................................................................................................. 14
   4.3 Missing data and other issues ................................................................................ 16

5. Results .......................................................................................................................... 18

6. Discussion ..................................................................................................................... 22
   6.1 Concluding remarks .............................................................................................. 24

References ....................................................................................................................... iv

Appendix 1: Tables ........................................................................................................... x
Appendix 2: Stata code ..................................................................................................... xi
1. Introduction

This thesis would argue that the educational system is an institution significant for the welfare of a society and its citizens. Education facilitates literacy, common values and basic knowledge for most citizens, elements essential for a functioning and stable democratic society (Friedman, 1955). Educational acquisitions could also affect civic behaviour and attitudes such as voting participation rates and the probability of supporting free speech (Dee, 2004). Public education increases the individual knowledge and skills of students and have a positive effect on their expected future income (Hirsch and Segelhorst, 1965).

Measuring the relative performance of schools could be viewed as relevant for both the supervisors of the educational system and the recipients of the services provided. Efficiency analysis of an industry is relevant for its economic planning as it shows if the output could be raised by reducing inefficiencies without the need of additional input (Farrell, 1957). According to the Statistics Sweden, the cost of education constituted 6.3 % of the Swedish GDP during the years 2009 to 2013 on average (SCB, 2014). Since the resources devoted to education constitutes a significant share of the Swedish economy then it is of interest to evaluate the performance of those providing the services. If individuals are rational and utility maximising then they would likely prefer attending the schools from which their educational benefits are the greatest.

From a historical perspective, mandatory school in Sweden have primarily been a responsibility of the government while only a limited number of non-government schools acquired public grants (Sandström and Bergström, 2005). In 1992, reforms of the Swedish school system were made giving privately owned schools rights to payments from the municipalities for every student enrolled. The grant per student enrolled was determined to be corresponding to at least 85% of the average self-cost of the municipality providing education in the area (Prop. 1991/92:95). After these reformations, public and independent schools have obtained public grants under approximately equivalent conditions (Lindbom, 2010). Any school, including for-profit corporations, became eligible if they fulfilled the base requirements (Sandström and Bergström, 2005). In contrast to other implementations of school voucher schemes, for example the Milwaukee voucher plan, examined by Greene, Peterson and Du (1999), where a random subset of students were given schools vouchers, the Swedish system included all students.
After the reformations, privately run schools funded by the government became a prominent feature of the educational system in Sweden (Magnússon, Göransson and Nilholm, 2015). These schools will be referred to as independent schools. Between 2006 and 2015, the fraction of students aged 15 enrolled in independent schools in Sweden increased from 8% to 18% (OECD, 2016). Unlike private schools, independent schools are not allowed to charge any fees for enrolling students (Lindbom, 2010). The principal organizer of the independent school must be approved by the Swedish National Agency for Education (NAE) before establishing an independent school (Björklund, Edin, Fredriksson and Krueger, 2004). The local governments are consulted but hold no veto when it comes to the approval of independent schools (Lindbom, 2010).

There is currently a debate regarding private enterprises operating in the welfare sector in Sweden. In 2016, a public inquiry of the private undertakings of publicly financed services was published. The inquiry, SOU 2016:78, suggested an upper limit on the profits made in the welfare sector. This report was criticised by the Swedish National Audit Office for prematurely discarding other possible options like regulations on the quality of the services provided and not sufficiently considering the effects of firms leaving the market (Riksrevisionen, 2017). They also criticised the inquiry for not considering the potential economic effects the significant departure from the fundamental principles of Swedish corporate law an implementation of the suggestion would constitute (Riksrevisionen, 2017). The business community also criticised the report for using an unfit method for calculating the upper limit of profits and for setting such a low limit on profits that it would increase the sensitivity to economic fluctuations of the sectors (PWC LLP, 2016). The inquiry arguably raised the attention paid to the issue of how tax money is used in these sectors and what services are received from the public means spent. An important question not as frequently discussed in the Swedish debate is the question of the quantity and quality of education which could be achieved with the resources used.

The aim of this thesis is to estimate the productivity in generating educational value for graduating students and the relative efficiency of public- and independent secondary schools in Sweden.

The thesis was organised as follows, section 2 and 3 describes previous literature on the efficiency analysis of schools and the effects of school competition respectively. Subsequently a chapter on the methods implemented followed by a description of the data used. Then the
main results of the analysis will be presented. Finally, then the results of the analysis will be discussed.

2. Literature review

2.1 Efficiency analysis of schools and school districts

Technical efficiency could be described as the ability of an agent to operate on the border of the production possibility set. Farrell (1957) provided a seminal article on the estimation of technical efficiency demonstrated using data from the agricultural sector in the United States. There exists a body of literature on the estimation of the technical efficiency of schools. Levin (1974) provides an early example of measuring school efficiency using frontier estimation in which data from a 1965 school survey was analysed using the linear programming method proposed by Aigner and Chu (1968). Bessent and Bessent (1980) introduced the Data Envelope Analysis (DEA) approach to the measurement of relative school efficiency and applied the method to the schools in an unspecified urban school district. This approach for measuring the relative efficiency of schools by estimating a best practice frontier has later been used in several other articles. Färe, Grosskopf and Weber (1989) used the DEA approach to estimate the relative efficiency of Missouri school districts, Grosskopf et al. (1999) used it to measure the relative efficiency of Texas school districts and the NAE used it when evaluating the relative efficiency of Swedish secondary schools (Skolverket, 2005). Scippacerola and D’ambra (2014) conducted a Stochastic Frontier Analysis (SFA) of schools in Campania to measure their relative efficiency and due to the limited range of the technical efficiency term a Tobit regression model was implemented to assess the factors affecting efficiency. SFA is a parametric approach for frontier estimation and this will be further discussed in the method chapter. Gronberg et al. (2017) did a cost function model on the relative efficiency of Texas charter schools focused on students at risk of dropping out. Scale advantages was found for campus size influencing average cost per student, holding quality constant, with a U-shaped relationship (Gronberg et al., 2017).

The contribution of this thesis is to measure the relative efficiency of Swedish secondary schools using a production frontier approach. Skolverket (2005) provides the only other known example of a study doing this. However, Skolverkets report differs from this thesis in various aspects. Firstly, a deterministic method was used in Skolverket (2005), this will be further

---

1 Survey on Equal Educational Opportunity
discussed in the method chapter. Secondly, the scope of the thesis at hand was wider as it analysed the results of Swedish schools over ten years while Skolverket (2005) only analysed three.

2.2 School competition and student results

School choice and school voucher systems have been debated for some time and some recurring theoretical arguments favour these approaches for developing educational systems instead of strictly government operated ones. Friedman (1955) argued that a school voucher system would increase the variety of institutions involved in education and that this would incite competition among providers of education. Theoretically, the competition for students from privatisation could create incentives for schools to improve their quality and stimulates the creation of new solutions and methods for teaching (Böhlmark and Lindahl, 2015). Competitions among providers of public services could also provide better matches between receivers’ preferences and services provided, if the recipients can choose from a set of possible options (Tiebout, 1956). Some like Böhlmark and Lindahl (2015) have claimed that the school reforms made in Sweden in the 1990s virtually emulated the system proposed by Friedman (1955).

Bergström and Sandström (2001) provided the first analysis of the impact of independent school competition on the performance of local schools in Sweden (Björklund et al., 2004). Bergström and Sandström (2001) claimed to have found evidence of independent school competition leading to higher educational standards for public schools within the same municipality. The report received severe criticism from teachers amongst others (Lidbom, 2010). Wibe (2002) presented some issues with the report for example that the data used was not representative for the entirety of Sweden and that a significant amount of data was missing. A reworked version of the paper was later published in 2005 (see Sandström and Bergström, 2005) which claimed to provide evidence of competition being beneficial to the quality of education. Ahlin (2003) did a study on the effects of local competition on school performance in Sweden using value-added models and found a significant positive effect on the results of the Swedish national standardised tests in math, but not Swedish and English. Some scholars, like Lindbom (2010) exclude to mention these insignificant results when discussing this report. Björklund et al. (2004) found evidence for a positive effect from independent school competition on the results in English and Swedish for students in public schools, but the effects on the results in math were mixed. Böhlmark and Lindahl (2015) did a study on both the short- and long-run effects of school competition on the students’ outcomes for Swedish students graduating from
compulsory school between 1988 and 2009 using a fixed effects panel regression. The long run effects studies were upper secondary school grades, university attendance and years of schooling. The study found that an increase in the share of students attending independent schools had a positive effect on both short- and long-run performance of students (Böhlmark and Lindahl, 2015).

School choice and competition for students between schools are not exclusively Swedish phenomena. Greene et al. (1999) did a report on the effects from private school enrolment using “the Milwaukee choice experiment” in which students were randomly assigned vouchers as data. The method used was a fixed effects OLS. The results found positive and significant effects from the years enrolled in private school and results in math and reading (Greene et al., 1999). Tiebout (1956) suggested that if individuals may choose their municipality of residence based on the bundle of public services and taxes they offer, they choose the one with the bundles conforming the most with their preferences. Hoxby (2000) found that an increase in the level of Tiebout choice, based on indexes of enrolment levels and geographical variables, had a positive effect on student results and a negative effect on school spending for school districts in the US. This would increase the efficiency of the local public goods markets as the preferences and provision of services get harmonised. Grosskopf et al. (2001) examined the determinants of technical- and allocative efficiency for different school districts in Texas using a Shephard input distance function and found that competition raised allocative efficiency but had no significant effect on the levels of technical efficiency. Bettinger (2005) estimated the effect charter schools had on the results of students attending adjacent public and charter schools in Michigan. The US charter schools share more similarities with the Swedish independent schools than the US private schools (Björklund et al., 2004). Using both difference-in-difference- and lagged dependent variable specification approaches and analysing both student- and school level data, no significant effect on student results were found (Bettinger, 2005). Hsieh and Urquiola (2006) found no evidence that the increase in school choice after the reforms in the 1980s in Chile had a positive effect on the average results for Chilean students. Grosskopf, Hayes and Taylor (2009) studies the efficiency of charter schools in Texas using a Shephard (1970) input distance function to inquire whether these schools are more, or less, efficient than public schools. The study found evidence of charter schools being more technical- and scale efficient than their public counterparts. A third of this difference was attributed to the regulatory conditions under which these schools operated (Grosskopf et al., 2009).
It seems as the question of efficiency gains in student performance due to privatization or competition between schools is an unresolved question. There is evidence suggesting that competition could have positive effects on student results and that private schools could be more efficient than their public counterparts. At the same time, these results are not consistent across the literature.

3. Method

Each school was viewed as an independent decision-making unit determining their own set of educational input factors. The definition of efficiency in this thesis was the maximum knowledge accumulation per student attainable from a given set of inputs. It was assumed that schools are maximising the results of their students given the resources available.

3.1 Estimating a best practice frontier

Different methods for estimating a best practice production frontier have been used in the literature. Two previously mentioned methods are SFA and DEA which both have their foundations in the economic production theory. The aim is to find a standard or benchmark to use as a comparison for all production units. Technical efficiency or inefficiency of a producer is given by the distance between the estimated best practice frontier and the observed production. Technical (in)efficiency in these perspectives is the variation in productivity among homogenous output maximising or cost minimising producers. SFA is a parametric frontier model with an error term specified in two components. This method was developed simultaneously by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). They were first to model stochastic production frontiers, or efficient frontiers where the output of a producer is affected by both individual inefficiency and exogenous shocks. Using SFA requires a priori specification of the relationship between input factors and output and a set of assumptions that vary in strength depending on the model at hand (Horrace and Schmidt, 1996). DEA is a non-parametric deterministic approach to estimate the efficient frontier using linear programming (Cooper, Seiford and Tone, 2007). No a priori specification of the production function is required as it is determined by enveloping the data using programming methods (Färe et al., 1989). The objective is to determine the most efficient producer, whether it is a producer observed in the data or a linear combination of several efficient producers (Mizala et al., 2002). DEA is a method based on extremal observations and is therefore sensitive to extreme observations (Scippacercola and D’Ambra, 2013).
The SFA approach was used in this thesis. Although DEA requires fewer assumptions, it presumes an absence of random errors and thus considers all deviations from the efficient frontier as inefficiency. This would have led to some issues for this thesis. First, inefficiency due to factors which the schools themselves have no possibility to control should not be held against them. The data available was limited and therefore potentially relevant variables like average per student spending or any measure of the capital stock per student could not be included in the models. Due to DEA not allowing for random errors, a few extreme observations could give exaggerated estimations of the production possibility for a given set of inputs. SFA is better equipped to separate random shocks from technical inefficiency. SFA is a statistical approach which recognises the existence of uncertainty and has the capacity to quantify it (Horrace and Schmidt, 1996). The basic SFA model can be written as:

$$y_{it} = f(x_{it}; \beta) + \epsilon_{it}, \quad i = 1, \ldots, N, \ t = 1, \ldots, T$$ (1)

where

$$\epsilon_{it} = v_{it} + u_{it}$$ (2)

A common definition of $y_{it}$ is that it is the logarithm of output and $x_{it}$ is a vector containing the logarithms of input variables (Horrace and Schmidt, 1996). $y_{it}$ was the logarithm of output for the $i$th school during period $t$ in this thesis. $\beta$ is a vector of coefficients and the error term is defined as $\epsilon_i$. Its first component, $v_i$, is a stochastic term describing statistical noise or exogenous shocks that affect the production frontier for individual producers. This allows for variations in the efficient frontier for different production units (Aigner et al., 1977). $v_i$ is assumed to be independent and identically distributed, with a normal distribution, and independent of all input factors for all observations and time-periods. The other component of the error term, $u_i$, have a one-sided distribution representing the inefficiency of the individual subject. This term is assumed to be distributed independently and identically among the producers. The individual producing unit is assumed to be in control of this error and its distribution follows the fact that the production unit must lie on or below its efficient frontier (Aigner et al., 1977).

Both Aigner et al. (1977) and Meeusen and van den Broeck (1977) used models based on cross-sectional data. The usage of cross-sectional data for estimating technical efficiency relies on
inefficiency being random and strict distributional assumptions (Kumbhakar, 1990). Since 1977, SFA models for panel data have been developed and refined. Pitt and Lee (1981) provided the first extension of SFA using panel data (Belotti et al., 2012). There are several benefits when using panel data instead of cross-sectional data. Pitt and Lee (1981) mentions four reasons why panel data is preferable over cross-sectional when estimating their efficient production frontier, two of which were relevant for this thesis. With panel data, there are possibilities to test for structural changes in the production function and for testing the over-time changes in technical efficiency (Pitt and Lee, 1981).

3.2 Model and assumptions:
The use of panel data makes it possible to check for unobserved heterogeneity among the different producers. One could argue that estimating this heterogeneity could obscure the inefficiency of individual producers. The main reason why this was done in this thesis was a lack of data on possibly relevant input variables.

It is possible to introduce heterogeneity among producers in the model using fixed- or random effects. Hausman specification tests for fixed versus random effects were conducted on the models used. This method was developed by Hausman (1978) and test whether random effects estimators yield inconsistent estimates by comparing fixed- and random effect estimations. If the assumption that all explanatory variables used in the model are uncorrelated with the producer specific heterogeneity does not hold then only fixed effects estimators yields consistent estimates. These tests found that fixed effects were better fits for the models used, the results of the tests are presented in Appendix 1 Table A1.1.

Computing fixed effects are problematic when using maximum likelihood and large panels (Greene, 2005). The models used data from a maximum of 2036 different schools, depending on the number of missing observations for the explanatory variables used. This caused computational problems which hindered the direct use of fixed effects estimators. To circumvent this problem the non-binary variables used in the models where de-meaned using within transformations, a method suggested by Wang and Ho (2010). This procedure meant that the sample averages from each school group were subtracted from the variables in the data. This is the same as when using fixed effects estimators, the only difference is that it is done before regressing. A Greene (2005) true random effects model with time varying inefficiencies using the de-meaned data was then implemented for estimating production frontier and the
efficiencies. The use of a random effects model for the de-meaned data was due to the issues with using a fixed effects model directly and limitations of the software used during the process of writing this thesis. Using random effects was deemed plausible as the size of the random effects estimators were expected to be close to zero since individual time invariant heterogeneity had already been removed from the data. As no individual time invariant heterogeneity was left in the de-meaned data then there was nothing for the random effects estimator to capture. This was tested by calculating the inefficiency with the random effect and with the random effects set to zero using the method developed in Jondrow et al. (1982). These results showed a negligible effect from the random effects as mean efficiencies and standard deviations were unaffected, summary statistics of the estimated inefficiencies can be found in Appendix 1 Table A1.2.

There exist different methods for estimating inefficiency as well as unobserved heterogeneity using SFA. Pitt and Lee (1981) were the first to use SFA with random effects and time invariant inefficiency. Schmidt and Sickles (1984) were the first to my knowledge to use a SFA model with fixed effects and time-invariant inefficiency. These models were not used in the thesis for two reasons. Firstly, as Greene (2005) argued the treatment of the effect for the models in Pitt and Lee (1981) and Schmidt and Sickles (1984) neglects the possibility of other non-efficiency related time-invariant heterogeneity which could affect the estimations of efficiency yielding biased results. Secondly, the assumptions of time-invariant inefficiencies could be considered inappropriate for long time series as this implies that inefficiency cannot change over time. Imposing time-invariant inefficiencies without testing its legitimacy could also lead to inconsistent estimations of the inefficiency and parameters (Kumbhakar, 1990). Therefore, a possible linear- and quadratic time trend in efficiency will be included in the model used in this thesis.

The incidental parameter problem was a possible argument against using fixed effects estimators in this thesis. This problem was originally pointed out by Neyman and Scott (1948). Even if the data was sampled over the scholastic years starting 2006-2015, many groups suffer from few observations due to missing data or since they were not operating for some years included in the study. Using fixed effects with short time series combined with many groups could lead to inconsistency in the MLE estimators (Hausman, Hall and Griliches, 1984). If many subjects are observed over a small or fixed number of periods this lead to asymptotically inconsistent estimations of the variance as shown by Lancaster (2000).
transformation method removes the potential issues caused by the incidental parameter problem (Wang and Ho, 2010). The true random effect model suggested by Greene (2005) is invariant to unbalanced panel data which was an argument behind using this method. Drawing inspiration from the specifications of the random effects model in Greene (2005) the model used in this thesis can be written:

\[ y_{it} = (\alpha + w_i) + \beta' x_{it} + v_{it} + u_{it} \]  

(3)

Where \( y_{it} \) represents production unit i’s output during year t, \( \alpha \) is the estimated intercept, \( w_i \) is time-invariant normally distributed unobserved heterogeneity for production unit i. \( \beta' \) is a vector of coefficients and \( x_{it} \) is a vector containing the set of inputs used by production unit i for year t. The last two terms represent the composed error term where \( v_{it} \) represents the stochastic element for production and \( u_{it} \) represents the individual firm inefficiency and can be described as:

\[ u_{it} \sim \mathbb{E}(\sigma_{it}^2) \]  

(4)

In the models used then the conditional variance of \( u_{it} \) was regressed as:

\[ \sigma_{it}^2 = e^{\gamma' z_{it}} \]  

(5)

\( z_{it} \) is a vector containing a constant and variables possibly affecting the variance of the distribution and \( \gamma \) is a vector of coefficients. The conditional variance measures production uncertainty and the method used in this thesis was the one suggested by Bera and Sharma (1999). The main reason why was that the results and interpretations of the conditional inefficiencies does not rely on the distributional assumptions when using this method (Bera and Sharma, 1999). An alternative method for estimating a conditional inefficiency is the conditional mean procedure developed by Battese and Coelli (1995).

A Cobb-Douglas production function was assumed when estimating this model. This production technology was assumed as it in the literature a commonly used benchmark function (used in Aigner et al., 1977; Battese and Coelli, 1988; Greene, 2005) and it is well behaved.
The logarithm of all non-binary variables was used in the regression. Hypothesis tests were conducted using heteroskedasticity robust standard errors. Wooldridge tests were used to control for autocorrelation. All variables containing fractions used in the analysis was in percent.

The method of predicting technical efficiency presented by Batteese and Coelli (1988) was used in this thesis. For this thesis, the technical efficiency could be written as:

\[
TE_{it} = \frac{\mathbb{E}[y_{it} | u_{it}, x_{it}]}{\mathbb{E}[y_{it} | u_{it} = 0, x_{it}]} \tag{5}
\]

Technical efficiency is the ratio between the mean production realised by the producing unit and the best practice production possibility given the input bundles used (Battese and Coelli, 1988). The measure of technical efficiency can range from zero to one and gives a measure of how large fraction of possible production that was achieved by subject \(i\) during period \(t\). The measure the Batteese and Coelli (1988) definition of technical efficiency is output oriented and gives a measure the maximum possible percentage increase of output without requiring additional inputs.

The distributional assumption of the individual inefficiencies is of importance due to the use of maximum likelihood estimation. In this thesis, an exponential distribution of \(u\) and a normal distribution of \(\nu\) and the unobserved heterogeneity was assumed. Two occurring distributional assumptions for \(u\) in the literature is the half-normal distribution and the exponential distribution. Both distributions were suggested and used in Aigner et al. (1977) and the exponential distribution was suggested and used by Meeusen and van den Broeck (1977). These distributions are one-sided and assume a mode fixed at zero. The truncated normal distribution was suggested by Stevenson (1980) and allows for the mode of the distribution to be non-zero. Battese and Coelli (1988) argued that the truncated normal distribution yields a more adequate representation of individual technical efficiencies than the half-normal distribution when there is a high probability of them not being gathered around the maximum. Some like Greene (2005) argue that the distributional assumption is of minor importance as it infrequently, if all, affects the results and states that the pivotal elements of the regression is the model specifications and its theoretical motivations. Skolverket (2005) estimated the technical efficiency of Swedish upper secondary schools using an output-oriented measure of technical efficiency and found
that the average level of technical efficiency was 0.92 with a relatively low dispersion. The exact method used for calculating the estimates for relative efficiencies was unfortunately not explicitly stated in Skolverket (2005). According to Skolverket (2005) then the schools studied on average achieved 92% of the maximum possible output estimated by the best practice frontier. An estimate of relative technical efficiency of 92% mean that the school could raise output with 8 percentage points by adopting the current best practice methods. This suggests that the mode of the distribution for the technical efficiency could likely be close or equal to zero, therefore was an exponential distribution of the inefficiency term assumed in this thesis.

3.3 Two-step estimation
The production frontier and the relative technical efficiency were estimated at the simultaneously in this thesis. Some scholars have implemented a two-stage approach to SFA, where the (in)efficiency is estimated first and then a regression is made on the estimated technical (in)efficiency. Pitt and Lee (1981) and Scippacerola and D’ambra (2014) are examples of articles using a two-step approach. This approach implies that input variables would be considered exogenous which contradict the assumptions necessary for estimating the technical efficiency (Kumbhakar, 1987). In a two-stage model, due to omitted variable bias in the first step regression the error term will be dependent on the (in)efficiency determining variables, this leads to the estimated (in)efficiencies becoming underdispersed which makes the estimated effects of these efficiency determining variables biased toward zero (Wang and Schmidt, 2002).

4. Data and model specification
The data used in this study was gathered from the Swedish National Agency for education’s (NAE) database SIRIS2 gathered for the scholastic years 2006/2007 through 2015/2016. A panel of data from all individual Swedish schools with graduating 9th graders over this period was created containing 16 728 observations. The panel was unbalanced since schools were both established and closed during this period. Data was also gathered from the Swedish municipality and county database Kolada3. The data from Kolada was only used for filtering the data, this was further discussed in the subchapter on missing data.

2 Skolverkets Internetbaserade Resultat- och kvalitetsInformationsSystem
3 Kommun- och landstingdatabasen
4.1 Output data

The data used for output were the individual schools’ average grade per subject for 9th grade students graduating during each scholastic year. Only students who received the grade pass in at least one subject were included in the calculation of the average grade score. The numeric scale for grades ranges from 0 to 20, where 0 represents a fail and 20 the highest grade available. Many scholars working with the effect of schools agree that the marginal effect of the school on educational outcomes is the most well-suited measure of school output (Grosskopf et al., 2001). Data on the individual level was not accessible during the process of writing this thesis therefore a value-added model was not an option. Others like Levin (1974) have also pointed out that focusing solely on the measures of cognitive achievements is a misspecification of the schools’ output as it fails to capture other important functions such as helping to develop the students’ values and attitudes. No data for these kinds of outputs were available and hence not included in this thesis. Even though the final grades are an imperfect summation of the educational effort it was the best measure of output available.

The choice of using the results of 9th graders was due to two reasons. Firstly, there exists no data for grades for 6th graders from years before the scholastic year of 2012/2013 in Sweden, since 6th graders did not receive grades before that. Secondly, results from upper secondary school4 students would represent the cumulative results of 9 years of schooling in addition to the value added from the schools evaluated. No data for estimating difference in results from the point of enrolment and graduation for upper secondary school students was accessible and taking the final grades directly would not capture the isolated educational value created by the upper secondary schools.

There were several reforms in the Swedish school system implemented in 2011 which included new curricula, a new school law and a new grading system (Skolverket, 2015). The numeric valuation of the maximum and minimum grade was not changed, but two additional steps were added in between the existing ones. For the details of the grading system the reader is referred to prop. 2008/09:66. The scholastic year 2013/2014 was the first cohort of 9th grade students educated and graded using the post reform standards. This posed a problem for this study as the results from the years after the reform were not equivalent to the ones from before. Since the method used measures relative efficiency and since this reform affected all schools in the same

---

4 Gymnasium in Swedish
way at the same time, it is viewed as possible to compare the results between these periods. The grading score was not changed to any larger extent and the grades are viewed as sufficiently comparable to use in this thesis. To account for the possible effects these reforms had on productivity an indicator variable named “Reform” was generated and used in the regressions.

During the scholastic years of 2006/2007 to 2013/2014 the grade score was reported as the average sum of the highest 16 grades for students. If the student had taken any additional courses, such as a 3rd language, then the highest 16 of the students’ grades were used in the calculation of the grade score. From the scholastic year of 2014/2015 and onwards then this was changed and schools reported average grade score from a summation of 17 grades, regardless whether the student had participated in any additional courses or not. To cope with the changes in the reporting of the statistics for average grade for each year was used instead. For the scholastic years 2006/2007 – 2013/2014 an average of the grade score divided over 16 subjects were used and for the scholastic years 2014/2015 – 2015/2016 the grade score divided by 17 subjects was used instead. Since reading 17 subjects is not mandatory then this is a possible source of negative bias in the average grades. An indicator variable for the time before and after this change in the reporting called “Report” was generated and used in the models.

4.2 Input data

Data over the students per teacher ratio was used as an input factor in the model. Teachers were measured as the number of teachers, including supply teachers employed for a period longer than a month, employed in October recalculated to full-time employments. Students was measured as students enrolled in the school during October. A problem with this variable is that it includes all students and teachers per school. This means that this is not necessarily representative for the resources used for educating the 9th graders. It is a reasonable assumption that the general student teacher ratio for the school an approximatively accurate measure of students per teacher for 9th graders attending that school. The students per teacher measure was inverted and the logarithm was taken of the inversion, this variable was referred to as “Teachers per student”. This was done to get a variable with a coefficient expected to be positive in the production frontier. The fraction of teachers employed with a pedagogical degree was used as a measure of the quality of the teachers employed. Teachers with a non-Swedish degree that have not been approved by the Swedish authorities are not included in this measurement. This variable was named “% Pedagogical degree”. Other variables, like years in the profession, have been used in other studies, for example Skolverket (2004), but these types of data were not
available for this thesis. There was also data on the fraction of teachers with a special pedagogical degree, this was not used since the demand of this kind of competence was expected to be highly correlated with the number of students with special needs attending the school.

Whether public- or independent schools differ in terms of efficiency was a factor of interest. One reason was the ongoing debate regarding private providers of public services in Sweden. The relative efficiency of public and non-public schools has been evaluated in several articles discussed in previous chapters. An indicator variable for public or independent schools was therefore included in the efficiency term in the models.

It was of interest for this thesis to see whether local school competition could affect the productivity of Swedish secondary schools. The fracture of independent schools within a municipality was used as an approximation of local school competition in the regression. This is not a perfect measure of competition, since for example the size of the schools are not included in the measure. For this thesis then the variable is considered a sufficiently accurate approximation. Since all municipalities did not have any independent schools every year $10^{10}$ was added to each observation before taking logarithms. This variable was named “% Independent schools”. Battese (1997) suggested an alternative solution for this problem using interaction terms and dummy variables for zero-observations.

The grades in Sweden are set by individual teachers without any outside audit (Skolverket, 2005). Examination instruments like standardised tests or external referees are not used in the Swedish school system, and while there are some standardised tests they are not mandatory to report (Wikström & Wikström, 2005). This poses a problem since the appraisal of student performances are likely heterogenous among teachers, meaning that the underlying knowledge which the grade represent might vary. To cope with this problem then the fraction of grades set higher than the students’ result on the standardised tests in math, Swedish and English was used in the production frontier for two of the models. Firstly, this means that schools not reporting the results of the tests to the NAE were not included in the analysis. Secondly, this provides an indication of whether the school set higher grades than the requirements of the curriculum. The aim was to reduce the potential efficiency advantage of schools inflating grades. These variables were referred to as “% Higher grade subject” for math, English and Swedish. These control variables are although not optimal due to how the tests are corrected. Only a fraction of these
tests gets corrected centrally while the bulk of the tests are corrected locally by the schools themselves (Skolinspektionen, 2016). The tests that are corrected centrally are not done so to calibrate grades (Wikström & Wikström, 2005). If a school systematically inflates both final grades and results on the standardised tests then these control variables would no longer be valid.

There are studies on the effects the socioeconomic status of a students’ parents has on student achievement. Ho and Willms (1996) found that students from households with relatively higher socioeconomic status had an advantage in academic achievement. The parental educational level is a factor which could influence the expected student results. In this thesis fraction of students with at least one parent with a post upper secondary school education was used. The variable was called “% Parent with higher education”. This variable was not entirely unproblematic due to the NEA removing observations based on fewer than ten students, this will be further discussed in the missing data section. Out of the possible variables available this was the one deeming most suitable and it was also the one with the least missing values. Other studies like Greene et al. (1999), Sandström and Bergström (2005) and Wikström and Wikström (2005) have used similar variables controlling for parental educational level.

The fraction of students with a Swedish background was used as a variable for the student characteristics. Swedish background was defined as students born in Sweden with at least one parent born in Sweden. Other measures of student nationality were available, but had a significantly higher rate of missing observations since the NEA removed all observations based on fewer than 10 subjects. Various studies have used similar variables as domestic students could have a relative advantage over foreign students as they for example could be more proficient in the language of instruction. Bergström and Sandström (2005), Scippacercola and D’Ambra (2014) and Gronberg et al. (2017) are examples of studies using similar variables.

4.3 Missing data and other issues
The data used in this thesis was not complete. Starting with missing data for the output variable then 359 observations out of 16729 were missing due to a lack of data. Statistics based on fewer than 10 students had also been removed by the Swedish National Agency of Education and it was not possible to get access to their raw data. This led to an additional loss of 1703 observations for the data on average grade score. These 2062 observations were dropped from the panel.
The praxis of removing observations also significantly affected data on parental education and student background. For parental educational level 7838 of the original 16729 observations were removed by the NAE, including missing data then 7996 observations were missing in total. To test the robustness of the estimations models, two models were estimated using this variable for parental education and two without.

In the case of student background then the fraction of students with a Swedish background was the alternative least affected by the NAE praxis. 2608 observations were missing due to being based on too few observations adding up to a total of 2686 missing observations. These missing values were reduced to 785 after the removal of observations without output data or with implausible student per teacher ratios.

There was also missing data for the difference between result on the standardised tests in math, Swedish and English. For English, then 3895 observations were missing in the raw data and 1688 in the data used. In the case of Swedish, there were 3457 missing observations in the raw data and 1286 in the cleaned data. Finally, for math there were 2950 missing observations in the raw data and 813 in the final panel. To test for the robustness of the estimates, two models without these variables were regressed.

Some observations of the number of students per teacher were dropped due to implausible values, in the raw data the number of students per teacher ranged from 0.1 to 720. The selection of variables was based on data of the municipality averages from the Kolada database. For each year the minimum and maximum student per teacher ratio was selected and two standard deviations was subtracted and added respectively. This gave a plausible range of student per teacher ratio. Out of the 14667 remaining observations after removing the ones with missing data on grades an additional 773 observations were dropped. Winsorization was considered but discarded as a selection criteria based on extremity of the individual values was deemed more adequate than dropping the most extreme percentiles.

---

5 Except from 2013 when the municipality Dorotea reported 0.1 students per teacher as a municipality average. This was a sharp decline from both the previous and significantly lower than the following year. This value was implausible and was not used as the 2013 minimum value.
Data on average yearly costs per student, expenditures on teaching equipment, libraries, student health, student meals, classrooms, etc. was only gathered at a municipality level and hence not usable for this thesis. There was also no data on capital goods available. The within-transformations was done to at least partly cope with this heterogeneity problem, as the method is only equipped to identify time invariant heterogeneity. This means that the model could likely suffer from omitted variable bias. This could lead to an underestimation of relative technical efficiency for schools with high over time variations in conditions and performance compared to schools with consistent performance over the period observed, regardless of whether the consistently performing schools have high- or low levels of productivity.

A potential problem was a lack of data for identifying schools not giving grades in accordance to the general curriculum or which operates under significantly different conditions. Schools like Waldorf-schools, hospital-schools and school that only educates children seeking asylum are examples of groups that would preferably have been removed as in Skolverket (2005). Since up to 2036 unique schools were included in the frontier estimations then it was not realistic to identify and remove these schools, due to time constraints. The schools for the newly arrived were most likely excluded due to the use of the fraction of Swedish students graduating from the school in the production frontier estimation. This since all observations based on fewer than 10 students were removed by the NAE.

5. Results
Wooldrige tests for autocorrelation were carried out using the four different model specifications. The results of these tests are presented in Table A1.1 in appendix 1. Evidence of autocorrelation was found in the models excluding parental educational level but not in the models including this variable. The main results of models 1 and 2 are given in Table 1. Table 2 presents the main results from models 3 and 4.

The estimated effect of local competition measured as portion of independent schools in a municipality on student performance was positive and significant in models 1 through 4. The coefficients can be interpreted as elasticities due to the use of a log-log model. For example, in model 1 if the fraction of schools within the municipality that are independent rises with 1 % then it is predicted that the average grade of students within that municipality rises with 0.000422 %.
Table 1: Main results from model 1 and 2.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Frontier 1</th>
<th>( \sigma^2_{it} ) 1</th>
<th>Frontier 2†</th>
<th>( \sigma^2_{it} ) 2†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers per students</td>
<td>-0.0235***</td>
<td>-(0.00800)</td>
<td>-0.0178***</td>
<td>(0.00620)</td>
</tr>
<tr>
<td>% Pedagogical degree</td>
<td>0.0157*</td>
<td>(0.00936)</td>
<td>0.0107</td>
<td></td>
</tr>
<tr>
<td>% Independent schools</td>
<td>0.000422**</td>
<td>(0.000192)</td>
<td>0.000340**</td>
<td>(0.000149)</td>
</tr>
<tr>
<td>% Higher grade math</td>
<td>-0.00169</td>
<td>(0.00124)</td>
<td>-0.00140</td>
<td>(0.00107)</td>
</tr>
<tr>
<td>% Higher grade English</td>
<td>0.00298***</td>
<td>(0.00108)</td>
<td>0.00314***</td>
<td>(0.000904)</td>
</tr>
<tr>
<td>% Higher grade Swedish</td>
<td>0.00831***</td>
<td>(0.00140)</td>
<td>0.00796***</td>
<td>(0.00110)</td>
</tr>
<tr>
<td>% Swedish students</td>
<td>0.0788***</td>
<td>(0.0102)</td>
<td>0.0760***</td>
<td>(0.00914)</td>
</tr>
<tr>
<td>% Parent with higher education</td>
<td>0.0565***</td>
<td>(0.00469)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reform</td>
<td>-0.00715***</td>
<td>(0.00270)</td>
<td>0.0164***</td>
<td>(0.00177)</td>
</tr>
<tr>
<td>Report</td>
<td>0.0446***</td>
<td>(0.00325)</td>
<td>0.0179***</td>
<td>(0.00221)</td>
</tr>
<tr>
<td>Linear time trend</td>
<td>-0.141</td>
<td>-(0.0945)</td>
<td>-0.260***</td>
<td>(0.0598)</td>
</tr>
<tr>
<td>Quadric time trend</td>
<td>0.0232***</td>
<td>(0.00766)</td>
<td>0.0297***</td>
<td>(0.00551)</td>
</tr>
<tr>
<td>Public</td>
<td>0.0902</td>
<td>(0.120)</td>
<td>-0.0930</td>
<td>(0.0882)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0167***</td>
<td>-(0.00115)</td>
<td>0.0157***</td>
<td>-(0.000986)</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.47562</td>
<td>0.47562</td>
<td>0.52321</td>
<td>0.52321</td>
</tr>
<tr>
<td>Log simulated likelihood</td>
<td>11230.5970</td>
<td>11230.5970</td>
<td>17759.4250</td>
<td>17759.4250</td>
</tr>
<tr>
<td>Observations</td>
<td>7043</td>
<td>7043</td>
<td>11205</td>
<td>11205</td>
</tr>
<tr>
<td>Number of schools</td>
<td>1682</td>
<td>1682</td>
<td>1926</td>
<td>1926</td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All non-binary variables were de-meaned and are in logarithm form. \( \sigma^2_{it} \) is the inefficiency. † autocorrelation in the model.
The fraction of students receiving higher grades than their results on the standardised tests in English and Swedish had positive and significant coefficients in model 1 and 2. For “% Higher math” then the estimated effect was negative but insignificant in both models.

Table 2: Main results from model 3 and 4.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Frontier 3</th>
<th>( \sigma_i^2 ) Frontier 3</th>
<th>Frontier 4†</th>
<th>( \sigma_i^2 ) Frontier 4†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers per students</td>
<td>-0.0219***</td>
<td>(0.00771)</td>
<td>-0.0183***</td>
<td>(0.00595)</td>
</tr>
<tr>
<td>% Pedagogical degree</td>
<td>0.0145</td>
<td>(0.00900)</td>
<td>0.00992</td>
<td>(0.00733)</td>
</tr>
<tr>
<td>% Independent schools</td>
<td>0.000445**</td>
<td>(0.000191)</td>
<td>0.000331**</td>
<td>(0.000150)</td>
</tr>
<tr>
<td>% Swedish students</td>
<td>0.0777***</td>
<td>(0.00957)</td>
<td>0.0725***</td>
<td>(0.00858)</td>
</tr>
<tr>
<td>% Parent with higher education</td>
<td>0.0564***</td>
<td>(0.00448)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reform</td>
<td>-0.00610**</td>
<td>(0.00252)</td>
<td>0.0169***</td>
<td>(0.00165)</td>
</tr>
<tr>
<td>Report</td>
<td>0.0451***</td>
<td>(0.00303)</td>
<td>0.0190***</td>
<td>(0.00202)</td>
</tr>
<tr>
<td>Linear time trend</td>
<td>-0.172**</td>
<td>(0.0843)</td>
<td>-0.191***</td>
<td>(0.0549)</td>
</tr>
<tr>
<td>Quadric time trend</td>
<td>0.0251***</td>
<td>(0.00706)</td>
<td>0.0232***</td>
<td>(0.00503)</td>
</tr>
<tr>
<td>Public</td>
<td>0.0244</td>
<td>(0.0968)</td>
<td>-0.270***</td>
<td>(0.0620)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0173***</td>
<td>(0.00107)</td>
<td>-7.782***</td>
<td>(0.000897)</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.49748</td>
<td>0.49748</td>
<td>0.57844</td>
<td>0.57844</td>
</tr>
<tr>
<td>Log simulated likelihood</td>
<td>12290.2799</td>
<td>12290.2799</td>
<td>20115.7183</td>
<td>20115.7183</td>
</tr>
<tr>
<td>Observations</td>
<td>7770</td>
<td>7770</td>
<td>13040</td>
<td>13040</td>
</tr>
<tr>
<td>Number of schools</td>
<td>1768</td>
<td>1768</td>
<td>2036</td>
<td>2036</td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All non-binary variables were de-meaned and are in logarithm form. \( \sigma_i^2 \) inefficiency. † autocorrelation in the model.
Teachers per students had a negative and significant coefficient in all models. These results were not intuitive and unexpected and were further discussed in the discussion chapter. The ratio of teachers with a university degree in pedagogy had positive but insignificant estimated effects in models 2 through 4, and was only significant at the 90% level in model 1. The indicator variable for public schools had an insignificant effect in models 1 through 3 and a significant negative effect on inefficiency in model 4. The estimates for the models including the variable for parental education were positive while the estimated effects in the models excluding this variable were negative. The ratio of students with parents with post-secondary education had positive and significant effect when included in the models.

Table 3: Mean technical efficiency per model and year

<table>
<thead>
<tr>
<th>Scholastic Year</th>
<th>Model 1</th>
<th>Model 2†</th>
<th>Model 3</th>
<th>Model 4†</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006/2007</td>
<td>0.9821691</td>
<td>0.9766814</td>
<td>0.9805468</td>
<td>0.9745605</td>
</tr>
<tr>
<td></td>
<td>(0.0067684)</td>
<td>(0.0109386)</td>
<td>(0.0079355)</td>
<td>(0.0130488)</td>
</tr>
<tr>
<td>2007/2008</td>
<td>0.9834225</td>
<td>0.9792661</td>
<td>0.9821195</td>
<td>0.9765933</td>
</tr>
<tr>
<td></td>
<td>(0.0059689)</td>
<td>(0.0105779)</td>
<td>(0.0070301)</td>
<td>(0.0131547)</td>
</tr>
<tr>
<td>2008/2009</td>
<td>0.9836697</td>
<td>0.9803035</td>
<td>0.9827628</td>
<td>0.9778452</td>
</tr>
<tr>
<td></td>
<td>(0.0053755)</td>
<td>(0.0095243)</td>
<td>(0.0059856)</td>
<td>(0.0116766)</td>
</tr>
<tr>
<td>2009/2010</td>
<td>0.9825386</td>
<td>0.9795807</td>
<td>0.9818714</td>
<td>0.9768935</td>
</tr>
<tr>
<td></td>
<td>(0.0064955)</td>
<td>(0.0089833)</td>
<td>(0.0067511)</td>
<td>(0.0108313)</td>
</tr>
<tr>
<td>2010/2011</td>
<td>0.9836632</td>
<td>0.9810956</td>
<td>0.9826006</td>
<td>0.9773338</td>
</tr>
<tr>
<td></td>
<td>(0.0061451)</td>
<td>(0.0086465)</td>
<td>(0.0070959)</td>
<td>(0.0115008)</td>
</tr>
<tr>
<td>2011/2012</td>
<td>0.9832798</td>
<td>0.9811216</td>
<td>0.9819938</td>
<td>0.9774226</td>
</tr>
<tr>
<td></td>
<td>(0.0066653)</td>
<td>(0.0087741)</td>
<td>(0.0076806)</td>
<td>(0.0129861)</td>
</tr>
<tr>
<td>2012/2013</td>
<td>0.9820096</td>
<td>0.9800352</td>
<td>0.9816395</td>
<td>0.9773582</td>
</tr>
<tr>
<td></td>
<td>(0.0065402)</td>
<td>(0.0097254)</td>
<td>(0.0070449)</td>
<td>(0.013521)</td>
</tr>
<tr>
<td>2013/2014</td>
<td>0.9790166</td>
<td>0.9772309</td>
<td>0.9782432</td>
<td>0.9741376</td>
</tr>
<tr>
<td></td>
<td>(0.0086102)</td>
<td>(0.0106029)</td>
<td>(0.0091172)</td>
<td>(0.0145472)</td>
</tr>
<tr>
<td>2014/2015</td>
<td>0.9661381</td>
<td>0.9641634</td>
<td>0.9649245</td>
<td>0.9606045</td>
</tr>
<tr>
<td></td>
<td>(0.0183771)</td>
<td>(0.0203383)</td>
<td>(0.0190752)</td>
<td>(0.0231536)</td>
</tr>
<tr>
<td>2015/2016</td>
<td>0.9773565</td>
<td>0.975908</td>
<td>0.9763677</td>
<td>0.9741688</td>
</tr>
<tr>
<td></td>
<td>(0.0152896)</td>
<td>(0.0172411)</td>
<td>(0.0169568)</td>
<td>(0.0190547)</td>
</tr>
<tr>
<td>2006/2016</td>
<td>0.9794271</td>
<td>0.9775548</td>
<td>0.9783962</td>
<td>0.9747402</td>
</tr>
<tr>
<td></td>
<td>(0.0119001)</td>
<td>(0.0129838)</td>
<td>(0.0126725)</td>
<td>(0.0155506)</td>
</tr>
</tbody>
</table>

Note: Standard errors in parenthesis. † autocorrelation in the model.
The coefficients for fraction of students with Swedish backgrounds were positive and significant in all four models. Reform or the control for the time before and after the reformations of the Swedish school system were significant in all models but switched sign from negative to positive in the models where “% Higher education” was excluded. The changes in the reporting of the grades had a positive and significant estimated effect in all models. All models shared similar estimated time trends in efficiency. Negative linear time trends for inefficiency were found. These were significant in models 2, 3 and 4 but not in model 1. For the quadric time trends then the coefficients were positive and significant for all model specifications.

In Table 3 the estimated average technical efficiency from model 1-4 is presented separated for each year. This average was above 96% for all models and years. A level of technical efficiency of 96% mean that the subject achieved 96% of the possible output estimated by the best practice frontier. In this case, then the school could raise its output with 4 percentage points without requiring additional resources.

6. Discussion
The negative coefficients for the teacher ratio were unexpected. These results could be due to a possible two-way causality where schools with a higher fraction of students with special needs or with other difficulties get assigned additional teaching resources. It could also be due to schools with few teachers using more efficient educational methods. A third possible explanation could be that some schools allocate their resources to other inputs than teachers and that the productivity of these resources is affecting both the students’ results but also the resources available for hiring teachers. This could for example be investments in equipment like computers, something that could both be costly and which possibly affects the learning of the students. It would have been interesting to compare the relative cost efficiency of teachers and computer equipment, but this goes beyond the purpose of this thesis.

The estimated technical efficiencies suggest that Swedish secondary schools on average operates relatively close to their best practice frontier. These results should be interpreted with precautions as the estimated inefficiencies were affected by the within-transformations of the data. If relative efficiencies were constant over time then the levels of technical efficiency could be over-estimated. The results could be interpreted as Swedish secondary schools being
homogenous in terms of technical efficiency or that the differences in technical efficiency were rigid over time and obscured by the school fixed effects. These potential explanations are not mutually exclusive. High average levels of technical efficiency coincide with the results of Skolverket (2005). This thesis had levels of average technical efficiencies 4 – 6.4 percentage points higher than those estimated in Skolverket (2005). This could be a sign of the individual levels of technical efficiency being obscured by the considerations made for producer heterogeneity. Another explanation could be the deterministic method used in Skolverket (2005), this approach likely yields lower levels of efficiency as these models leave no room for random errors. It is important to remember that one can only make statements regarding the relative level of efficiency, not the absolute, when analysing productivity using SFA models.

Positive relationships between school competition and student performance were found in the regressions. These estimations agree with previous studies made on education in Sweden like Bergström and Sandström (2005), Skolverket (2005) and Böhlmark and Lindahl (2015). This correlation could be explained in various ways. One plausible explanation is that competition creates stronger incentives for schools to raise the quality of their education to become more appealing to potential students. An alternative explanation is that school competition induces grade inflation as schools might use high grading as a cheap way to attract potential students. The evidence for whether schools use grade inflation to raise their productivity was inconclusive in this thesis. Two out of three coefficients had the sign that was expected. The negative relationships between fraction of students receiving final grades higher than their results on the standardised tests in math and average grade were unexpected. This could be a sign of low explanatory power of this variable or of the flaws in this instrument for measuring grade inflation. The fraction of final grades set higher than the students’ results on the standardised test in both Swedish and English had positive effects on the average grades. This could be an indication of schools using inflated grades to raise their productivity levels. The fact that the size of the competition effects was relatively unaffected by the inclusion of these variables could be interpreted as this effect being robust for grade inflation. One way to test the validity of the benefits of competition against the possibility of grade inflation as explanations is to study the future achievements of the students graduating from the schools. This would require individual level data and patience as the last cohort included in this study graduated 2016. A third possible explanation is that independent schools could have preferred to establish themselves in areas with socioeconomically strong populations so that these results was induced by selection bias. That the size of the competition effect increased when the control for parental
education was included could be interpreted as evidence against this potential explanation. Another argument against this explanation is that the socioeconomic status of an area is likely persistent over time and thus should the benefits from being able to choose location at least partly be reduced by the usage of de-meaned data. What can be stated from the results is that within municipality school competition between public and independent schools seem to have a positive effect on the average grades of the graduating 9th grade students.

When it comes to the relative efficiency of public and independent schools then the coefficient was insignificant in all but one model. In model 4 the coefficient was negative and significant, which implies that public schools could on average be less inefficient than the independent ones. The results in models 1 through 3 agree with the inconclusive results of previous studies. One could argue that other factors than the type of the principal organiser are more relevant for school performance.

That the coefficients for the fraction of teachers with a pedagogical degree were insignificant at the 95% level in all models was unexpected. One possible explanation is that this is an inadequate measure of teacher quality as it does not control for whether teachers tutor in the subjects for which they have qualifications or how experienced they are. Other explanation could be that qualified teachers require higher standards of their students when grading them. The data used in this thesis was insufficient for making any conclusive statements regarding this.

6.1 Concluding remarks
Swedish schools seem to be relatively similar regarding efficiency according to the results of this thesis. The results point toward potential positive effects from local school competition on the productivity of upper secondary schools in Sweden and that the type of the principal organiser of a school could be of little relevance for its level of technical efficiency.
References


Wibe, S. (2002). Leder konkurrensen från friskolor till högre kvalitet i undervisningen?. Ekonomisk debatt, 4, 244-55.

Appendix 1: Tables

Table A1.1: Results from Hausman tests

<table>
<thead>
<tr>
<th></th>
<th>$\chi^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>458.21</td>
<td>0.0000</td>
</tr>
<tr>
<td>Model 2</td>
<td>480.90</td>
<td>0.0000</td>
</tr>
<tr>
<td>Model 3</td>
<td>543.63</td>
<td>0.0000</td>
</tr>
<tr>
<td>Model 4</td>
<td>374.08</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table A1.2: Summary statistics of estimated average inefficiencies for model 1 through 4 with random effects and with random effects set to zero.

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>7043</td>
<td>0.0210373</td>
<td>0.0126363</td>
<td>0.0061831</td>
<td>0.250764</td>
</tr>
<tr>
<td>Model 1*</td>
<td>7043</td>
<td>0.0210373</td>
<td>0.0126363</td>
<td>0.0061768</td>
<td>0.250764</td>
</tr>
<tr>
<td>Model 2</td>
<td>11205</td>
<td>0.0229673</td>
<td>0.0139068</td>
<td>-0.1504016</td>
<td>0.2574406</td>
</tr>
<tr>
<td>Model 2*</td>
<td>11205</td>
<td>0.0229673</td>
<td>0.0139068</td>
<td>-0.1504281</td>
<td>0.2574405</td>
</tr>
<tr>
<td>Model 3</td>
<td>7770</td>
<td>0.0221066</td>
<td>0.0135006</td>
<td>-0.0034593</td>
<td>0.2726704</td>
</tr>
<tr>
<td>Model 3*</td>
<td>7770</td>
<td>0.0221066</td>
<td>0.0135006</td>
<td>-0.003476</td>
<td>0.2726704</td>
</tr>
<tr>
<td>Model 4</td>
<td>13040</td>
<td>0.0259425</td>
<td>0.0167013</td>
<td>0.0059407</td>
<td>0.2965075</td>
</tr>
<tr>
<td>Model 4*</td>
<td>13040</td>
<td>0.0259425</td>
<td>0.0167013</td>
<td>0.0059532</td>
<td>0.2965075</td>
</tr>
</tbody>
</table>

Note: * random effects set to 0.

Table A1.3: Results from Wooldrige tests

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.794</td>
<td>0.3731</td>
</tr>
<tr>
<td>Model 2</td>
<td>6.605</td>
<td>0.0103</td>
</tr>
<tr>
<td>Model 3</td>
<td>1.986</td>
<td>0.1590</td>
</tr>
<tr>
<td>Model 4</td>
<td>8.091</td>
<td>0.0045</td>
</tr>
</tbody>
</table>
Appendix 2: Stata code

clear
import excel "C:sers\Johan\Desktop\Data D\PanelD1506.xlsx", sheet("PanelD1506")
firstrow clear
set more off

*Installation of packages
net install sfpanel, all from(http://www.econometrics.it/stata)
ssc install outreg2
ssc install center
findit xtserial
net sj 3-2 st0039
net install st0039

*Destringing data
destring Students_Total, replace force
destring Average_Merit1, replace force
destring FrUPD, replace force
destring Students_Teacher, replace force
destring FrHigh_Sv, replace force
destring FrHigh_Ma, replace force
destring FrHigh_En, replace force
destring FrPostG, replace force
destring FrSwe, replace force

*Defining panel
xtset School_Code Year, yearly

*Discarding rows without data on output
drop if missing(Average_Merit1)
**2062 observations were dropped due to a missing output data.

*Discarding implausible students per teacher ratios
**+/- two st.dev, rounded to two decimals, for year for max/min
drop if Students_Teacher <= 6.25 & Year == 2006
drop if Students_Teacher >= 18.85 & Year == 2006
drop if Students_Teacher <= 6.42 & Year == 2007
drop if Students_Teacher >= 16.28 & Year == 2007
drop if Students_Teacher <= 4.25 & Year == 2008
drop if Students_Teacher >= 16.85 & Year == 2008
drop if Students_Teacher <= 6.39 & Year == 2009
drop if Students_Teacher >= 16.91 & Year == 2009
drop if Students_Teacher <= 6.15 & Year == 2010
drop if Students_Teacher >= 18.85 & Year == 2010
drop if Students_Teacher <= 5.67 & Year == 2011
drop if Students_Teacher >= 17.13 & Year == 2011
drop if Students_Teacher <= 4.71 & Year == 2012
drop if Students_Teacher >= 17.69 & Year == 2012
drop if Students_Teacher <= 5.63 & Year == 2013
*2013 Outlier removed, Dorotea reported an average of 10 teachers per student, 0.1. statially lower than all other observations
drop if Students_Teacher >= 18.57 & Year == 2013
drop if Students_Teacher <= 6.30 & Year == 2014
drop if Students_Teacher >= 16.60 & Year == 2014
drop if Students_Teacher <= 6.38 & Year == 2015
drop if Students_Teacher >= 18.02 & Year == 2015
*775 observations with output data were dropped

*Generating the variables used
gen Reform = Year >= 2013
gen Report = Year >= 2014
gen Public = School_Type == "Kommunal"
gen lnAM = ln(Average_Merit1)
gen lnFrUPD = ln(FrUPD)
gen lnFrHEn = ln(FrHigh_En)
gen lnFrHMa = ln(FrHigh.Ma)
gen lnFrHSv = ln(FrHigh.Sv)
gen lnFrS = ln(FrSwe)
gen lnFrPG = ln(FrPostG)
gen t = Year - 2005
quietly generate double tt=t
quietly generate double tt2=t^2
gen FrInd = frind*100 + 10^(-10)
gen lnFrInd = ln(FrInd)
gen TS = (Students_Teacher)^(-1)
gen lnTS = ln(TS)

*Wooldrige tests for autocorrelation
xtserial lnAM lnTS lnFrUPD lnFrInd lnFrHMa lnFrHEn lnFrHSv lnFrS lnFrPG Reform Report
xtserial lnAM lnTS lnFrUPD lnFrInd lnFrHMa lnFrHEn lnFrHSv lnFrS Reform Report
xtserial lnAM lnTS lnFrUPD lnFrInd lnFrS lnFrPG Reform Report
xtserial lnAM lnTS lnFrUPD lnFrInd lnFrS Reform Report

*Hausman tests of models
xtreg lnAM lnTS lnFrUPD lnFrInd lnFrHMa lnFrHEn lnFrHSv lnFrS lnFrPG Reform Report, fe
estimates store fixed1
xtreg lnAM lnTS lnFrUPD lnFrInd lnFrHMa lnFrHEn lnFrHSv lnFrS lnFrPG Reform Report, re
estimates store random1
hausman fixed1 random1

xtreg lnAM lnTS lnFrUPD lnFrInd lnFrHMa lnFrHEn lnFrHSv lnFrS Reform Report, fe
estimates store fixed2
xtreg lnAM lnTS lnFrUPD lnFrInd lnFrHMa lnFrHEn lnFrHSv lnFrS Reform Report, re
estimates store random2
hausman fixed2 random2

xtreg lnAM lnTS lnFrUPD lnFrInd lnFrS lnFrPG Reform Report, fe
estimates store fixed3
xtreg lnAM lnTS lnFrUPD lnFrInd lnFrS lnFrPG Reform Report, re
estimates store random3
hausman fixed3 random3

xtreg lnAM lnTS lnFrUPD lnFrInd lnFrS Reform Report, fe
estimates store fixed4
xtreg lnAM lnTS lnFrUPD lnFrInd lnFrS Reform Report, re
estimates store random4
hausman fixed4 random4

* Demeaning the data to remove fixed effects
by School_Code: center lnAM lnTS lnFrUPD lnFrInd lnFrHMa lnFrHEn lnFrHSv lnFrS lnFrPG

* Frontier models, estimation of (in)efficiencies and generation of tables
  ** Model 1
  sfpanel c_lnAM c_lnTS c_lnFrUPD c_lnFrInd c_lnFrHMa c_lnFrHEn c_lnFrHSv c_lnFrS c_lnFrPG Reform Report, model(tre) usigma(tt tt2 Public) robust
predict TE1, bc
predict IE1, u
predict IE10, u0
outreg2 using rreg1.doc
  ** Model 2
  sfpanel c_lnAM c_lnTS c_lnFrUPD c_lnFrInd c_lnFrHMa c_lnFrHEn c_lnFrHSv c_lnFrS Reform Report, model(tre) usigma(tt tt2 Public) robust
predict TE2, bc
predict IE2, u
predict IE20, u0
outreg2 using rreg2.doc
  ** Model 3
sfpanel c\_lnAM c\_lnTS c\_lnFrUPD c\_lnFrInd c\_lnFrS c\_lnFrPG Reform Report, model(tre)
usigma(tt tt2 Public) robust
predict TE3, bc
predict IE3, u
predict IE30, u0
outreg2 using rreg3.doc

** Model 4
sfpanel c\_lnAM c\_lnTS c\_lnFrUPD c\_lnFrInd c\_lnFrS Reform Report, model(tre) usigma(tt tt2 Public) robust
predict TE4, bc
predict IE4, u
predict IE40, u0
outreg2 rreg4.doc

* Summary statistics for Jondrow et. al (1982) inefficiencies
sum IE1 IE10 IE2 IE20 IE3 IE30 IE4 IE40

* Summary statistics for Battese and Coelli (1988) efficiencies
sum TE1 TE2 TE3 TE4

*Estimating yearly average efficiency for model 1-4
gen TE106 = TE1 if Year == 2006
gen TE107 = TE1 if Year == 2007
gen TE108 = TE1 if Year == 2008
gen TE109 = TE1 if Year == 2009
gen TE110 = TE1 if Year == 2010
gen TE111 = TE1 if Year == 2011
gen TE112 = TE1 if Year == 2012
gen TE113 = TE1 if Year == 2013
gen TE114 = TE1 if Year == 2014
gen TE115 = TE1 if Year == 2015

sum TE106 TE107 TE108 TE109 TE110 TE111 TE112 TE113 TE114 TE115

gen TE206 = TE2 if Year == 2006
gen TE207 = TE2 if Year == 2007
gen TE208 = TE2 if Year == 2008
gen TE209 = TE2 if Year == 2009
gen TE210 = TE2 if Year == 2010
gen TE211 = TE2 if Year == 2011
gen TE212 = TE2 if Year == 2012
gen TE213 = TE2 if Year == 2013
gen TE214 = TE2 if Year == 2014
gen TE215 = TE2 if Year == 2015
sum TE206 TE207 TE208 TE209 TE210 TE211 TE212 TE213 TE214 TE215

gen TE306 = TE3 if Year == 2006
gen TE307 = TE3 if Year == 2007
gen TE308 = TE3 if Year == 2008
gen TE309 = TE3 if Year == 2009
gen TE310 = TE3 if Year == 2010
gen TE311 = TE3 if Year == 2011
gen TE312 = TE3 if Year == 2012
gen TE313 = TE3 if Year == 2013
gen TE314 = TE3 if Year == 2014
gen TE315 = TE3 if Year == 2015

sum TE306 TE307 TE308 TE309 TE310 TE311 TE312 TE313 TE314 TE315

gen TE406 = TE4 if Year == 2006
gen TE407 = TE4 if Year == 2007
gen TE408 = TE4 if Year == 2008
gen TE409 = TE4 if Year == 2009
gen TE410 = TE4 if Year == 2010
gen TE411 = TE4 if Year == 2011
gen TE412 = TE4 if Year == 2012
gen TE413 = TE4 if Year == 2013
gen TE414 = TE4 if Year == 2014
gen TE415 = TE4 if Year == 2015

sum TE406 TE407 TE408 TE409 TE410 TE411 TE412 TE413 TE414 TE415