



LUND UNIVERSITY
School of Economics and Management

Grade inflation in Sweden

Differences between public and independent schools

Author: Erik Norström

Supervisor: Therese Nilsson

Bachelor Thesis, NEKH02, May 2016

Department of Economics, Lund University

Abstract

In a setting where both schools and students have incentives to demand higher grades than may be warranted, there is a risk of grade inflation. Shedding light on the observed grade inflation in Swedish schools, this study investigates possible grading differences between independent and public compulsory schools by comparing their grades to national test scores. Theory and previous studies suggest that factors such as a school's socioeconomic composition influence grading. This study therefore controls for school characteristics when comparing independent and public schools. The results from cross-school regressions, using data on English, Mathematics and Swedish from 2015, indicate that there is no statistically significant difference between the two education providers for any of the three subjects. The results are robust to various specifications, but the model explains little of the variation across schools. An insight is therefore that the reason deviations vary is possibly school-specific policies or individual teachers' grading practices. These factors are probably determined independently of education provider and school characteristics.

Keywords: compulsory school, grade inflation, independent schools, public schools

Contents

Abstract	1
1 Introduction.....	4
2 Background	6
3 Theoretical foundations	9
4 Previous research	11
5 Method and data.....	14
5.1 Research approach	14
5.2 Data	16
5.3 Variables	18
5.4 Descriptive statistics	20
6 Results.....	22
6.1 Sensitivity analysis.....	24
7 Discussion	27
8 Conclusion	30
References.....	32

List of Figures

Figure 1. Pisa scores and grades in compulsory school.....	5
Figure 2. Independent and public compulsory schools.....	7

List of Tables

Table 1. Descriptive Statistics.....	21
Table 2. Regression results	23
Table 3. Regressions with municipality dummies	25

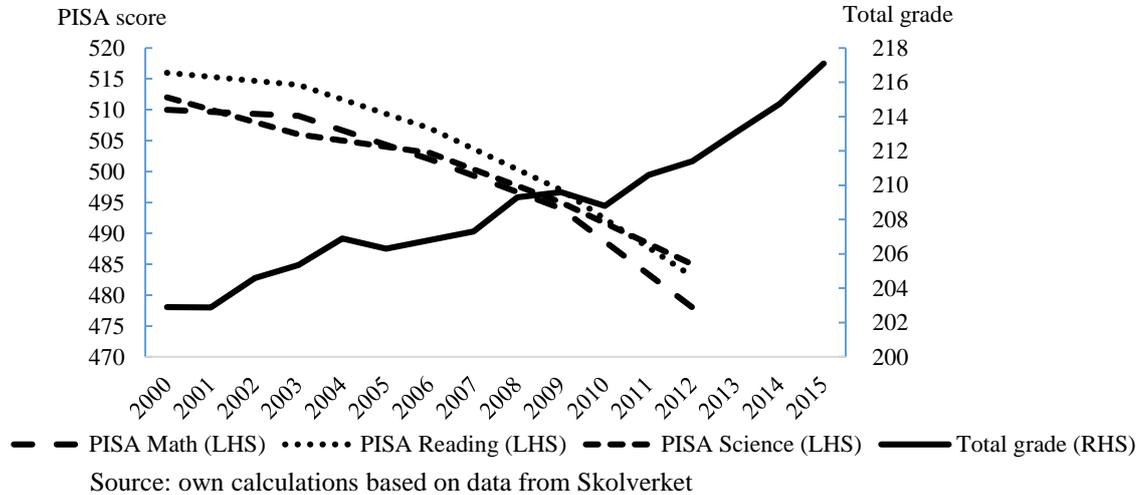
1 Introduction

Since the early 1990s, grades in Swedish compulsory schools have been increasing while the students' knowledge have been decreasing. The decline knowledge can be seen in all ages and across both low- and high-achieving students (Holmlund, Häggblom, Lindahl, Martinson, Sjögren, Vikman & Öckert, 2014). For example, Swedish students now rank about average compared to students in other OECD countries in the PISA (Programme for International Student Assessment) tests of 15-year olds. At the same time, grade differences between Swedish schools are increasing, raising concerns that grading practices vary across schools. Differences in grading practices would have serious consequences because students attending more generous schools would gain an advantage when applying to further education.

Generous grading can also result in students learning less (Vlachos, 2011).

Grade inflation is defined as a decrease in the knowledge needed for a certain grade (Skolverket, 2012). Although no measure of students' knowledge is perfect, figure 1 exemplifies the fall in performance and rise in grades. The graph displays PISA test scores and Swedish students' average total grades in compulsory school from the year 2000 onward. PISA tests are conducted every three years and Sweden's results have worsened at every occasion, contrasting the consistent increase in grades.

Figure 1. Pisa scores and grades in compulsory school



Several explanations for this dismal development have been proposed. Reforms in the early 1990s saw the expansion of independent schools, and increased competition for students, something that may have caused grade inflation across the educational system. The adoption of a criterion-referenced grading system, also in the 1990s, likely contributed (Holmlund et al. 2014).

At the individual level, studies have found that girls, low-achieving students and non-natives are more likely to receive a higher grade than their national test result (Lindahl, 2007; Vlachos, 2010). With socioeconomic factors controlled for, evidence on possible grading discrepancies between independent and public schools is mixed. For example, Vlachos (2010) finds that independent schools inflate their grades less than public schools, while Skolverket (2016) finds that independent schools are more lenient graders in English than public schools.

This thesis brings new evidence on grade inflation in Swedish compulsory schools. The purpose of the paper is to identify possible grading discrepancies between independent and public compulsory schools. Therefore, the research question is whether, conditional on school characteristics, independent and

public compulsory schools set grades that have equal net-deviations from national test scores. Not rewarding the same performance on the national tests with the same grades would be evidence of dissimilar grading standards.

The cross-school analysis, using data from 2015, reveals no statistically significant difference between the two education providers. The findings are in line with previous research, which suggest that the difference in grade inflation, if any, between public and independent compulsory schools is small.

The outline of this thesis is as follows. Section two describes the fundamental reforms of the Swedish school system that took place in the 1990s and led to the rapid expansion of independent schools. Section three establishes the theoretical foundations for why schools have incentives to mark generously and section four reviews related studies on the issue. The method and data are discussed in section five. The results are presented in section six and discussed in section seven together with implications for future research. Finally, the conclusions are presented in section eight.

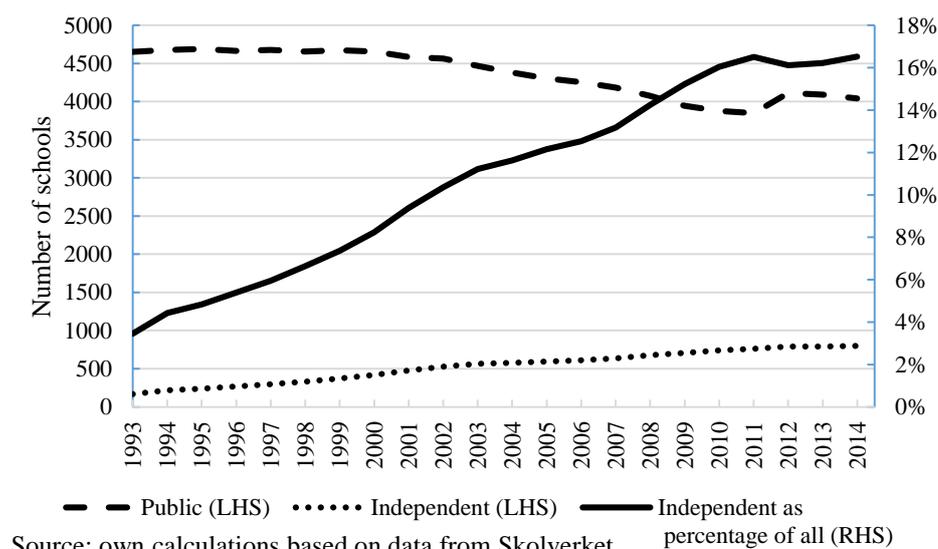
2 Background

The Swedish school system was fundamentally reformed in the 1990s. It went from being highly centralized to very decentralized. For example, the authority for primary and secondary education was transferred from the central government to the municipalities. The municipalities now run the public schools. The system can be described as a quasi-market: education is publicly financed, but families can choose freely among both public and independent schools in their municipality. The municipalities have to fund independent and public schools on a per-student basis. These funds may differ across municipalities, but within them, they are the same for both independent and public schools (Björklund, Clark, Edin, Fredriksson & Krueger, 2005).

These reforms led to a rapid expansion in the number of independent compulsory schools. The number of schools went from around 50 in 1990 to over 700 in 2009. At the same time, the number of public schools decreased from over 4500 to roughly 4000. As a result, the total number of schools have remained approximately the same. During that period, the share of ninth grade students in independent compulsory schools has risen from around zero to eleven percent (Holmlund et al. 2014).

Figure 2 illustrates the boom in the number of independent schools. In 2014 almost a sixth of all Swedish compulsory schools was run by an independent education provider.

Figure 2. Independent and public compulsory schools



The spread of independent schools varies across the country (Holmlund et al. 2014). The independent schools have predominantly increased their market share in urban areas, while there are still several smaller municipalities without any independent school. Compared to the population in rural areas, the population in urban areas is more educated, has a higher income, and has a larger share of people with a foreign background. These demographics are reflected in the socioeconomic composition of students in independent

schools: compared to the average public school, the average independent school has a larger proportion of students with a foreign background and students with university educated parents who have high income. It is not possible to distinguish whether this is due to independent schools actively targeting these groups or higher demand by educated parents and families with a foreign background (Holmlund et al. 2014).

In an overview of independent education providers, Skolverket (2014) notes that independent compulsory schools have, on average, slightly fewer students than public schools. Another fact is that independent schools are a heterogeneous group. For example, their operational form varies. Of all operational forms, limited companies are the most common, educating over sixty percent of all students attending independent schools. Schools run as foundations, non-profit associations and cooperatives share the rest of the market roughly equally. Cooperates are more common in rural municipalities while limited companies and foundations have a larger market share in urban municipalities. Common to all of these different forms is that the proportion of parents with post-secondary education is higher than in public schools. Independent schools are also different from public schools with regards to teachers: 86 percent of teachers in public schools have a degree in education compared to 70 percent in independent schools (Skolverket, 2014).

There are differences in admission criteria between public and independent schools. Public schools must give priority to students in the nearby area before allocating slots to other applicants, whereas independent schools can use three criteria: sibling priority, queue date and proximity. No school is allowed to accept students on the basis of previous school results. In theory all families have the same opportunity to choose school, but there are differences in practice. Firstly, all students do not have multiple schools near their home and may not be able to move to a neighbourhood close to a preferred school. Secondly, some families may not have all the information necessary to make

an informed school choice. Finally, to attend popular independent schools, families must often queue for several years. Hence, to attend these schools, parents must be active and place their children in the queue (Holmlund et al. 2014).

3 Theoretical foundations

As mentioned previously, the 1990s reforms fundamentally changed the Swedish school market. There are arguments for why these reforms may be beneficial, but also for why they may have unintentional effects. Since funding is linked to enrolment, a market mechanism is created. Hence, failing schools will be outcompeted while popular schools will receive more funds (Ahlin, 2003). Competition from independent schools and benchmarking against successful schools can therefore improve the quality of education and lead to a more efficient use of organizational resources (Bergström & Sandström, 2001).

However, evaluating reforms using grades as the measure of output can be misleading when there are few external controls preventing schools from setting higher grades than are warranted. Due to the absence of truly objective measures of students' achievements and the value-added by schools, having high average grades can be a way for schools to attract prospective students. As a result, schools have an incentive to inflate their grades. If the likelihood and cost of being detected is low, it is optimal for the individual school to set too high grades (Wikström & Wikström, 2005).

The incentives to inflate grades are likely stronger in independent schools. Since independent schools are smaller than public schools, they cannot utilize economies of scale to the same extent as public schools. Moreover, they cannot charge tuition fees, and their revenue is therefore to a large extent the

student vouchers they receive. This makes them more sensitive to changes in demand than are public schools (Wikström & Wikström, 2005).

Vlachos (2011) provides a different view compared to the previous arguments. He notes that competition from independent schools not necessarily will cause public schools to increase their quality. For example, it could be the most dissatisfied students who first switch from a public school to a newly opened independent school, reducing pressure for quality improvement in the public school. In addition, if families value factors other than academic quality, such as wellbeing, generous marking, or social composition, it is likely that the outcome of school competition will reflect these other aspects rather than academic quality. Furthermore, since revenue is not linked to quality but to the number of students enrolled, some for-profit schools' strategy could be to keep costs low and attract families not capable of evaluating the schools' educational quality.

Further objections to the arguments that competition will increase school productivity are given by MacLeod and Urquiola (2012). Families cannot observe the value-added by a specific school as student achievement is determined jointly by the school and its students' abilities. Hence, families will use alternative measures when choosing schools. These measures can be related to both input and output. Examples of such measures are selectivity, better peers and absolute achievement (average grades or test results). In a Swedish context, where the only admission criteria for upper secondary school is the final grades from compulsory school, this would imply that families value a student's grades more than his or her national test results. Schools also have incentives to mark generously as that is a much cheaper way to signal high quality compared to improving actual quality, by for example developing better teaching methods. Policy makers, on the other hand, have other goals. They are more concerned with how much knowledge students accumulate in school and that a certain level of achievement is rewarded with the same

grades in all schools. This reasoning is supported by Vlachos (2010). He highlights that students and society have different goals that are not easily reconciled. Students act in self-interest and want to, given effort, maximize their grades since these are used in the application to upper secondary school. Society on the other hand wants grades to be a precise indicator of the students' true knowledge. If the system leads to generous marking, there is a risk that students work less hard and learn less, which is negative for society. Furthermore, if the extent of generous marking differs across schools, students attending a school with strict marking will be disadvantaged when applying to further education.

4 Previous research

The work most recent and similar to this thesis is a study by Skolverket (2016). The authors investigate whether there are any systematic differences between public and independent education providers with regards to discrepancies between national test results and grades. As a measure of discrepancies, they use net-deviations, defined as the percentage of students receiving higher grades than their national test results minus the percentage of students receiving lower grades than their test results for that subject. For example, if a certain school has 30 percent of students receiving a higher grade, 50 percent receiving the same grade, and 20 percent receiving a lower grade, the net-deviation for that school for that subject is 10 percent. They find that for the language subject Swedish, public schools have a higher net-deviation, while for English the opposite is true. For Mathematics they find no difference between public and independent schools. However, they point out that these differences may not be a result of dissimilar grading standards, but of other factors, such as the students' socioeconomic background—gender, foreign background or parents' education level—or the students' overall

academic achievement, measured as their total grade. Regressing the net-deviations on type of education provider and these other factors, they find no statistically significant difference between public and independent schools in Swedish and Mathematics, but for English the result is still significant for the years 2013, 2014, and 2015.

Studying an earlier time period, 2003 to 2008, using individual level data, Vlachos (2010) finds no statistically significant difference between independent and public schools, except for 2008. For that year, students at independent schools had a lower difference between national test results and grades than students at public schools, after socioeconomic factors were accounted for. Still, the difference was slight.

On the individual level, the way teachers assess students depends on the students' gender and ethnicity. Lindahl (2007) finds that girls and non-natives are more favourably graded than boys and natives, given national test scores. Yet she points out that the result for non-natives may be partly explained by the fact that non-natives are overrepresented among poor performing students, who tend to be assessed favourably. She also gives a different explanation, namely that teachers might be afraid to discriminate against these two groups, overcompensating them as a result. Moreover, there is also evidence of relative grading in schools. Relatively high-achieving students in poor performing schools tend to receive higher grades than might be warranted (Fredriksson & Vlachos, 2011).

There are also studies on whether increased competition among schools has caused the observed general grade inflation. The results are mixed. For example, Sandström and Bergström (2005) find no grade inflation in Mathematics due to competition. Studying the upper secondary school level, Wikström and Wikström (2005) compare students' upper secondary school

grades with results on SweSAT¹ to determine the effects of competition between public and independent schools on grade inflation. Their main hypothesis is that independent schools inflate grades more than public schools. Using individual level data, controlling for factors such as the students' final grades from compulsory school, foreign background, gender, parental income and education, they find little evidence of competition causing grade inflation. However, their main hypothesis is confirmed: "The results suggest that independent schools seriously engage in grade inflation" (Wikström & Wikström, 2005, p. 317). The authors consider the independent schools' inflated grades a result of the fact that independent schools are new and small. In order to utilize increasing returns to scale, they need to signal their academic quality as a means of attracting prospective students. Public schools, on the other hand, are already large and have no such incentives. In fact, they may operate under decreasing returns to scale. The authors suggest that for future research, school size effects could be an important factor to study as grade inflation may be more easily prevented in larger schools, which have better opportunities to apply internal control mechanisms.

Another finding in Wikström and Wikström (2005) is that foreign-born students in public schools received more favourable grades than their SweSAT scores predicted. The authors hypothesize that media coverage of the high number of failing students may cause teachers to avoid failing students in the core subjects Mathematics, Swedish, and English. By doing so, they inflate foreign-born students' grades as this group is overrepresented among low-achieving students.

The literature seems to suggest that schools use grade averages to attract students, but when choosing schools, do parents care about average grades or the value-added by the different schools? To determine that, Black and Machin (2011) study whether school quality has any effect on house prices. In

¹ Högskoleprovet in Swedish

their comprehensive literature review of studies mostly from the U.S., but also from Europe, for example Norway, they find that school quality does affect house prices in the nearby area. However, this was only the case when school quality was measured as the school's absolute levels of test scores. "In our view a not unreasonable benchmark summary of the magnitude of the average causal impact is that a one standard deviation increase in test scores raises house prices by around 3%" (Black & Machin, 2011, p. 515). Jacob and Lefgren (2007) use a different approach and examine at the school level which teachers different parental groups prefer. They find substantial differences across schools. Parents in socioeconomic disadvantaged schools valued the teachers' ability to raise student academic achievement more highly, while parents in socioeconomic advantaged schools valued the teachers' ability to improve student satisfaction to a larger extent. However, they also report that within a certain school, parents from different socioeconomic groups seemed to have the same preferences regarding teachers' abilities.

5 Method and data

5.1 Research approach

Systematic grading differences between independent and public compulsory schools can be studied in three ways, which have their own strengths and weaknesses (Vlachos, 2010).

Firstly, one can test how compulsory school graduates with a certain grade perform in upper secondary school, and determine whether there are any differences between graduates from independent and public compulsory

schools. The method's main limitation is that it assumes that independent and public upper secondary schools provide the same value-added to all students. In addition, this method requires individual level data, which is not readily available.

Secondly, one can test the relationship between grades in subjects that have national tests and grades in subjects without them, for example art and music. It is possible that students with aesthetic interests and talent self-select into independent schools, and naturally receive higher grades in the practical subjects as a result. It is also possible that independent schools are more generous graders in all subjects, concealing any difference between the two types of subjects. For these reasons the third method is not perfect either.

Thirdly, one can test discrepancies between a school's average grades and national test results. Vlachos considers using this difference between a subjective and a more objective measure of knowledge as the most intuitive way to test for dissimilar grading standards between independent and public schools. Admittedly, this approach has its own drawbacks. The national tests do not test all aspects of the curriculum, so some deviations are warranted. A more serious critique is that grading of the national tests are done at the schools, often by the same teacher who decides the student's final grade. External controls indicate that teachers mark these tests generously too (Skolinspektionen, 2015). As expected, schools where teachers do not mark their own class are less generous than schools where teachers mark the class that he or she normally teaches. Also, the weight given to national test scores as a basis for the final grade likely vary between schools. For these reasons this method probably underestimates the difference in grade inflation between schools. Although Vlachos uses individual data, I believe this measurement is also the most intuitive when using school-level data.

Just as it is a flawed approach to use only average grades to determine which education provider is the best, testing whether independent and public schools

set the same grades given test scores by simply comparing their average deviations is naïve. To illustrate, if non-native students are assessed more favourably by all schools but are overrepresented in public schools, this would give a false impression that public schools are more generous graders.

With the method's limitations in mind, I construct the following model for each of the three subjects English, Mathematics and Swedish for the academic year 2014/2015. Although there are also national tests in the social and natural sciences, those subjects are not included as each school is only tested in a few of the subjects (Skolverket, 2016).

$$\begin{aligned} \text{Netdev} = & \beta_0 + \beta_1 \text{independent} + \beta_2 \text{girls} + \beta_3 \text{foreign background} \\ & + \beta_4 \text{size} + \beta_5 \text{total grade} + \beta_6 \text{highed} + \beta_7 \text{degree} + \varepsilon \end{aligned}$$

In this model, the net-deviations from national test results are a function of the type of education provider, the share of girls and students with a foreign background in the class and their average total grade, their parents' level of education, school size, the teachers' education level, and an error term. The error term is assumed to have an expected value of zero for all values of the explanatory variables. More formally, this means that

$$E(\varepsilon | \text{independent}, \text{girls}, \text{size}, \text{foreign background}, \text{total grade}, \text{highed}, \text{degree}) = 0.$$

If this condition holds, an ordinary least squares (OLS) estimate will be unbiased (Wooldridge, 2009). Hence, OLS is used to estimate the model.

5.2 Data

All data is collected from Skolverket's information system SIRIS. SIRIS contains several data sets describing the schools' performance and composition in terms of teacher and student characteristics. SIRIS is primarily intended to be used not by researchers but by for example politicians,

journalists, and schools (SIRIS, 2016). The data is collected by Statistics Sweden (SCB). Still it is apparent that the data is not meant for researchers. The data sets had to be “cleaned” and some transformations be made before merging the different data sets in Stata using each school’s unique school id.

There are a few problems with the data. The data on net-deviations is only provided for schools where at least fifteen students both took the test and received a final grade. Additionally, any value on *highed* that is constructed using data from fewer than ten students is censored, regardless of the percentage. The reason this variable is chosen instead of one covering only the ninth grade students, which is not censored, is that I found inconsistencies when inspecting the data on the latter: a substantial fraction of schools had percentages adding up to a lot less than 100 percent. As urban areas have larger schools and a more educated population, the sample is ultimately slightly biased towards large and urban schools. This kind of sample selection is a variant of exogenous sample selection and reduces the sample size, but does not lead to a biased OLS estimate, as opposed to endogenous sample selection (selection based on the dependent variable), which leads to a biased OLS estimate (Wooldridge, 2009). The sample size ranges from 1305 for Swedish to 1370 schools for English. Swedish has the fewest observations probably because some schools have a large proportion of students taking Swedish as a second language, a separate course, reducing the sample size. As the total number of schools providing education for ninth grade students is around 1680, the data covers approximately 80 percent of all schools.

However, a closer inspection of the data reveals that not all schools seem to be active, as they have not reported any scores. Among the rest of missing schools, 117 have five students or fewer in ninth grade, and 87 have six to ten students. As these extremely small schools may be fundamentally differently organised than larger schools, it could be justified to not include them. In addition, these schools’ net-deviations would be extremely sensitive to

individual cases and therefore not that useful for determining systematic deviations from national test scores.

5.3 Variables

This section presents the variables most used in this study. I define them, motivate their inclusion and describe the expected outcome. Although additional variables are included in subsequent robustness tests, those variables are not essential to the study and are not treated here.

Net-deviation is the dependent variable used in all regressions. It measures a school's grading deviation from its national test score. Assuming that national test scores are a more objective indicator of student achievement, net-deviation is a measure of the grade inflation for a particular subject. Like Skolverket (2016), I define net-deviation as:

% of students with a higher grade than their national test score –
% of students with a lower grade than their national test score.

This measure does not account for the extent of deviation for a particular student. For example, it does not differentiate between a student who is raised from a C to an A and one that is raised from C to B. However, almost all discrepancies are only one step on the scale (Skolverket, 2016), making this only a minor issue. Nor does the measure account for the gross deviations from test results. For example, schools with no deviations and schools where the same number of students receives a higher and lower grade will both have a net-deviation of zero.

Independent is the explanatory variable of main interest. It is a dummy variable that takes the value 1 if the school is run by an independent education provider and 0 for public schools. If the estimated coefficient on *independent* is significantly different from zero, public and independent schools have

different net-deviations that cannot be explained by other factors, such as the schools' socioeconomic composition.

Foreign background is defined as the percentage of ninth grade students who are either born abroad or have parents whom are both born abroad. As previous research has indicated that students with a foreign background may be assessed more favourably, the estimated coefficient is likely to be positive.

Highed is defined as the percentage of the school's students who have at least one parent with university education. This measure is used instead of one covering only the students in ninth grade. Previous studies suggest that schools with a favourable socioeconomic composition should feel less pressured to set high grades as the parents in these schools prefer teachers who improve student satisfaction. The coefficient on *highed* is therefore hypothesised to be negative.

I also control for the percentage of ninth grade students who are girls. This is an important variable to control for since studies indicate that girls receive more favourable grades conditional on their national test scores. The share of girls in the class should therefore be positively related to the school's net-deviation.

School size is measured as the number of students in ninth grade. The logarithmic value is used in the regressions. Small schools ought to be more financially sensitive to fluctuating demand. They therefore have stronger incentives to inflate grades in order to attract prospective students. In addition, larger schools may have better internal control mechanisms, such as cross-grading by other teachers. The coefficient on *size* is therefore likely negative.

Total grade is the average total grade score for the students in ninth grade. It ranges from 0 to 320. To illustrate, if all students achieve the highest grade possible, A, in all 16 subjects, the variable takes the value 320. As the literature suggests that teachers avoid failing students, and that high-achieving

students in poor performing schools are graded generously, it should be schools with overall low-achieving students that have a larger net-deviation. This effect is also to some extent purely mechanical: students who failed the test can only be raised and A-students only lowered.

Degree is defined as the percentage of teachers at the school who holds a degree in education. I have not found any previous study testing the effect of having a larger proportion of teachers with a degree in education. Nor have I found any compelling theoretical arguments for why it should affect a school's net-deviation. Yet the measure may capture differences in schools' organization and focus. For example, schools that employ fewer educated teachers may be schools specialising in teaching practical subjects.

5.4 Descriptive statistics

Table 1 presents descriptive statistics of the variables used in this study. The grade inflation problem appears to be most severe in Mathematics—the average net-deviation is almost 37 percent. In Swedish and English, the average net-deviations are 17.8 and -3.8 percent, respectively. Although these large differences are certainly interesting, this paper is more concerned with the difference between independent and public schools rather than the absolute level of net-deviations. For this reason, a discussion of why the absolute levels differ among the three subjects is outside the scope of this paper and will not be treated.

Table 1. Descriptive Statistics

VARIABLES	English mean (sd)	Mathematics mean (sd)	Swedish mean (sd)
Net-deviation English	-3.802 (17.59)		
Net-deviation Mathematics		36.71 (19.95)	
Net-deviation Swedish			17.77 (17.01)
Independent schools (%)	21.4	21.1	21.5
Girls (%)	48.51 (9.248)	48.51 (9.219)	48.36 (9.256)
Students in ninth grade	67.94 (37.54)	68.10 (37.49)	69.13 (37.73)
Foreign background (%)	23.12 (21.89)	23.05 (21.82)	20.11 (17.50)
Total grade average	225.4 (25.53)	225.5 (25.53)	226.8 (24.39)
At least one parent with higher education (%)	52.88 (16.46)	52.83 (16.40)	53.80 (15.96)
Teachers with a degree (%)	83.69 (11.61)	83.69 (11.62)	83.92 (11.49)
Students per teacher	12.40 (2.37)	12.42 (2.37)	12.46 (2.37)
Number of teachers	26.28 (11.40)	26.30 (11.43)	26.24 (11.40)
Students in the school	328.51 (166.49)	329.14 (166.93)	330.00 (168.43)
Observations	1,370	1,365	1,305

6 Results

Table 2 presents the results of the baseline estimation. The low R-squared implies that the model explains little of the variation in net-deviations. Independent schools appear to have the same net-deviations as public schools. As in Skolverket (2016), the estimated coefficient is positive for English but negative for Swedish. Still, these results are not statistically significant. The share of girls in the class does not seem to explain much variation between schools, other than in Mathematics. The school size can explain some variation in Swedish. The sign is not negative as expected, but positive. The estimated coefficient is 2.4, implying that a one percent increase in the number of ninth grade students increases net-deviation by 0.024 percentage points. This is a minimal effect and bears little economic relevance.

As anticipated, schools with a larger share of students with a foreign background have a larger net-deviation. For English, where the effect is strongest, the *ceteris paribus* effect of a one percentage point increase in the share of students with foreign background is a 0.18 percentage point increase in a school's net-deviation. It is not surprising that foreign background does not affect a school's net deviation in Swedish. A large proportion of foreign students takes Swedish as a second language, which is a separate course. The students' total grade also matters. A one-point increase in the average total grade is associated with a 0.15 and 0.14 percentage point increase in a school's net-deviation in English and Swedish, respectively. The share of students whose parents have high education is strongly negatively related to net-deviation in all the three subjects. So is the share of teachers with a degree, with Swedish as an exception.

The results are presented with robust standard errors because a White-test revealed non-constant variance of the error term, resulting in invalid t-

statistics and p-values (Wooldridge, 2009). The heteroscedasticity-robust standard errors correct this and make p-value inference possible.

A possible issue when estimating individual coefficients is multicollinearity, which occurs when independent variables are highly correlated. To detect possible multicollinearity, the variance inflation factor (VIF) is calculated in Stata for all variables. Although arbitrarily chosen, a VIF above ten is sometimes considered a problem (Wooldridge, 2009). As the VIF for no variable is above three, multicollinearity is not considered an issue when interpreting the results.

Table 2. Regression results

VARIABLES	English	Mathematics	Swedish
	Net-deviation	Net-deviation	Net-deviation
Independent	1.076 (1.423)	-0.669 (1.808)	-1.157 (1.552)
Girls	0.066 (0.056)	0.122* (0.063)	-0.038 (0.055)
Logsize	-0.353 (0.844)	0.995 (1.032)	2.408*** (0.884)
Foreign background	0.177*** (0.025)	0.053* (0.031)	0.046 (0.030)
Total grade	0.151*** (0.029)	-0.047 (0.039)	0.144*** (0.032)
High education	-0.313*** (0.043)	-0.127** (0.054)	-0.200*** (0.047)
Degree	-0.154*** (0.047)	-0.108** (0.054)	-0.074 (0.047)
Constant	-14.564* (8.616)	51.981*** (10.179)	-6.621 (8.386)
Observations	1,370	1,365	1,305
R-squared	0.129	0.044	0.033

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.1 Sensitivity analysis

One possible issue with a multiple regression is the risk of omitted variable bias, which occurs when an independent variable that is correlated with both an independent variable and the dependent variable is left out from the regressions. This violates the assumption that the error term has an expected value of zero for all values of the explanatory variables, leading to biased estimates (Wooldridge, 2009). To check the robustness of my findings, the model is also estimated using various specifications and combinations of independent variables.

To permit differences across municipalities, I use the baseline specification but include municipality dummies. The reasons for doing so are that municipalities are education providers and can have different grading policies, or strategies to help students who failed the national tests. Also, theory suggest that competition, which is probably stronger in urban areas, can cause grade inflation. Hence, including municipality dummies can mitigate possible omitted variable bias arising from, for example, the fact that schools face different levels of competition.

Table 3 presents the results of this alternative specification. (Municipality dummies are not shown.) For English and Swedish, the results are no different than those of the baseline model. For Mathematics, *independent* is significantly different from zero on a ten percent level. On the five percent level, which is the one normally considered statistically significant (Westerlund, 2005), the major difference between the two models is that parents' education is no longer statistically significant in explaining net-deviations in Mathematics.

Table 3. Regressions with municipality dummies

VARIABLES	English Net- deviation	Mathematics Net- deviation	Swedish Net- deviation
Independent	-0.419 (1.758)	-3.678* (2.141)	-1.231 (1.852)
Girls	0.043 (0.063)	0.090 (0.072)	-0.018 (0.062)
Logsize	-0.865 (1.086)	1.003 (1.247)	3.024*** (1.064)
Foreign background	0.145*** (0.039)	0.029 (0.048)	0.058 (0.045)
Total grade	0.133*** (0.036)	-0.059 (0.052)	0.143*** (0.038)
High education	-0.303*** (0.066)	-0.047 (0.078)	-0.190*** (0.068)
Degree	-0.156*** (0.058)	-0.184*** (0.071)	-0.037 (0.060)
Constant	4.923 (9.715)	64.458*** (13.039)	-11.384 (11.083)
Observations	1,370	1,365	1,305
R-squared	0.307	0.269	0.281

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

As the share of teachers with a degree only captures the quality of teachers, I also test whether including the student per teacher ratio alters the results. The variable is defined as the total number of students at the school divided by the total number of teachers. The ratio can reflect variations in class size, and having a high ratio is a way of keeping costs low, as hiring teachers is expensive. Adding the student per teacher ratio to the baseline regressions does not alter the results. For Swedish and English, the variable is not significant and does not change the estimates on other variables. In Mathematics, the variable is significant on the ten percent level and the coefficient estimate is 0.49, which means that an additional student per teacher

increases net-deviation by roughly half a percentage point. Including the variable also makes *foreign background* non-significant and *girls* significant on 5 percent level, but with no change in the estimated coefficients. Most importantly, the estimate on *independent* does not change.

In the baseline regressions the logarithmic number of students in ninth grade is used as a measure of school size. An alternative measure of school size is the logarithmic total number of students at the school. Using this variable does not alter the results in any meaningful way. Yet another variable to use is the logarithmic number of teachers employed at the school. Doing so changes no estimate on *independent*. However, the school size becomes significant for English at the ten percent level, with a coefficient estimate of -1.97. For Mathematics, *girls* is significant at the five percent level with the same coefficient estimate. For Swedish the results are the same.

I also test whether omission of the variable *degree* alters the results. The reason for doing so is that it is the only variable that the schools are in direct control of; schools choose which teachers to hire but not which students to admit. The *ceteris paribus* effect of *independent* on net-deviations may not be that relevant for policy as independent schools also, on average, employ less educated teachers. Although leaving out significant variables that are correlated with other independent variables results in omitted variable bias, it can be interesting to test how doing so changes the results. Omission of *degree* results in *independent* being significant for English on the five percent level, with a coefficient estimate of 3.06. This implies that, in English, if all other explanatory variables are held constant, independent schools have net-deviations that are 3.06 percentage points larger than those of public schools. This is likely the result of independent schools having less educated teachers, as *degree* is negatively related to net-deviation. Still, the grade inflation problem is not that apparent in English, as the average net-deviation is -3.8 percent. As a result, this finding has little policy relevance.

7 Discussion

Economic theory and previous studies suggest that factors such as school size and class composition may affect how lenient teachers are when grading their students. The results of this paper vary slightly depending on the exact specification, but in general, type of education provider and socioeconomic factors can explain little of the variation in net-deviations seen among Swedish schools. However, the results appear fairly robust to the inclusion of municipality-specific factors. This is in line with previous work by Vlachos and Skolverket who find that any difference between independent and public schools is small. The results contrast those of Wikström and Wikström who study the upper secondary schools. Possible explanations for this could be that they use the SweSAT as the measure for “objective” achievement and that the study is over ten years old—a lot may have changed since then, for example with the introduction of a new grading scale in 2012/2013.

As the net-deviations appear only weakly related to school and classroom characteristics, the results instead imply that other factors may shed light on why schools set different grades given test scores. For example, some schools’ policy may be to not deviate from national test scores more than a certain amount for a given year; individual teachers assess students differently; and the weight given to test scores as the basis for the final grade can vary across schools and classes within schools. In determining why some schools appear more generous than others, I suggest that future research focus on school-specific policies and individual teachers’ grading practices.

Theory suggests that small schools ought to be more sensitive to demand as the marginal benefit of the individual student is larger in small schools. For this reason, they have stronger incentives to set high grades to keep and attract

new students. As school size was either not significant or had a very small effect, the results refute this theory.

Contrary to previous research and theory, the results of this study imply that net-deviations in English and Swedish increases with students' overall achievement, measured as their total grades. It is not possible to distinguish whether this is due to some schools grading generously in all subjects, or if overall high-achieving students perform well in aspects not captured by the tests. In any case, this refutes the hypothesis that low-achieving students are graded generously.

The results do confirm those of Jacob and Lefgren. To recall, they found that parents in socioeconomic advantaged schools cared more about teachers' abilities to raise student satisfaction than the students' academic achievement, as opposed to parents in disadvantaged schools where the opposite was true. As *Highed* is strongly negatively related to net-deviations in most estimations this could be taken as evidence that parents in socioeconomically disadvantaged schools exert more pressure on teachers, leading to higher grades given performance. Yet caution should be exercised, because of the "all other factors held constant" interpretation of a multiple regression analysis. In reality, the students' total grade and the parents' education level is strongly positively correlated. This means that in practice, when schools attract more parents with high education, student performance will usually also increase. Therefore, from a policy perspective, studying the parents' education level in isolation may be of little interest.

The results imply that the share of ninth grade students with a foreign background is positively related to the net-deviations, at least in English. This confirms the result of earlier studies. As mentioned previously, several students with a foreign background likely take Swedish as a second language, a separate course. Hence, it is not surprising that the variable is never significant for Swedish. Lindahl puts forth the hypothesis that teachers are

afraid to discriminate against foreign students and therefore overcompensate them with higher grades. The results do not support this hypothesis. Students with a foreign background are probably more disadvantaged in Mathematics than in English, as the former is normally taught and examined in Swedish while the latter is not. So if the hypothesis holds, non-natives should be more favourably assessed in Mathematics, not in English as the results of this study indicate.

Previous studies also find that girls are more favourably assessed than boys given national test scores. The paper does not confirm these results. A likely reason is that there is not much variation in the share of girls in ninth grade among the schools. Therefore, whether girls are assessed more favourably is probably better studied using individual level data.

This paper also controls for the share of teachers at the school who holds a degree in education. The variable is significant for English and Mathematics, implying that increasing the share of teachers with a degree results in a lower net-deviation. One must be careful in interpreting this as evidence that increasing teacher qualification will stem grade inflation on the national level. Due to the thin theoretical support for why this variable should be included in the model, caution is needed. Still, the sensitivity analysis suggests that the main results of the paper are not contingent on the inclusion the variable.

I must acknowledge that there are several important limitations of this study. One is that it only covers one year, albeit the most recent. As both Vlachos and Skolverket have noted, which educational provider is more generous varies through time. Therefore, the results of this study will not necessarily be able to predict future development. Another is the weaknesses of the method used. Admittedly, some deviations from test scores are warranted as they are not designed to test every aspect of the students' knowledge. Moreover, using net-deviations as a measure of grade inflation implicitly assumes that schools mark the national tests correctly. In fact, schools that are generous graders can

be strategic by marking the national tests generously, making their net-deviations artificially low. Ultimately, an important policy recommendation is therefore to develop more reliable tests of student achievement, so that schools' grading policies can be better compared, which can also be achieved by marking the national tests centrally.

A general problem with the research focusing on the comparison between independent and public schools is that independent schools are a heterogeneous group. Grouping together limited companies, foundations and non-profit associations hides any differences between these sub-groups. As Swedish schools cannot charge tuition fees to increase revenues, the only way to be profitable is to keep costs low and attract students. As theory and previous studies suggest, high overall grades are important for students when choosing schools. As for-profit schools have stronger financial incentives than other schools, they may find it easier to attract students by setting slightly higher grades than warranted instead of increasing educational quality, which is costly. Future research can therefore focus on the different sub-groups of independent education providers.

8 Conclusion

To shed light on possible grading discrepancies between public and independent education providers, the research question of this thesis is whether, conditional on school characteristics, independent and public compulsory schools set grades that have equal net-deviations from national test scores. The answer to that question is most likely yes, as I find no differences between the two education providers.

Theory and previous studies suggest that factors such as school size and socioeconomic composition can influence grading decisions. The broader implications of the paper are that these factors explain little of the observed variation in discrepancies between national test results and grades among schools. Instead, the answer to why some schools appear more generous than others is possibly found by studying aspects not related to school characteristics, such as school-specific grading policies and individual teachers' grading practices, which are two areas I suggest future research to focus on. A lot of work is still required to understand why Swedish students receive higher grades while their knowledge is decreasing.

References

Ahlin, Å. (2003). Does School Competition Matter? Effects of a Large-Scale School Choice Reform on Student Performance, Working Paper, no. 2003:2, Department of Economics, Uppsala University, Available online: <http://www.econstor.eu/bitstream/10419/82869/1/wp2003-002.pdf> [Accessed 20 April 2016]

Bergström, F. & Sandström, M. (2001). Konkurrens Bildar Skola: En ESO-rapport om friskolornas betydelse för de kommunala skolorna, Ds 2001:12, Stockholm: Ministry of Finance, Available online: http://eso.expertgrupp.se/wp-content/uploads/2013/08/2001_12-Konkurrens-bildar-skola.pdf [Accessed 25 April 2016]

Björklund, A., Clark, M. A., Edin, P.-A., Fredriksson, P., & Krueger, A. B. (2005). The Market Comes to Education in Sweden: An evaluation of Sweden's surprising school reforms, New York: Russell Sage Foundation, Available online: <http://www.jstor.org/stable/10.7758/9781610440554> [Accessed 22 April 2016]

Black, S. E. & Machin, S. (2011). Housing Valuations of School Performance, in Hanushek, E., Machin, S. & Woessman, L. (eds), *Handbook of the Economics of Education*, Volume 3, Amsterdam: North Holland, pp.485-518, Available online: <http://www.sciencedirect.com.ludwig.lub.lu.se/science/handbooks/15740692> [Accessed 15 April 2016]

Fredriksson, P. & Vlachos, J. (2011). Reformer och Resultat: Kommer regeringens utbildningsreformer att ha någon betydelse? Rapport till

- Finanspolitiska rådet 2011/3, Stockholm: Finanspolitiska rådet, Available online:
<http://www.finanspolitiskaradet.se/download/18.34d8785812fd178ec94800022368/110608+Fredriksson+och+Vlachos.pdf> [Accessed 5 May 2016]
- Holmlund, H., Häggblom, J., Lindahl, E., Martinson, S., Sjögren, A., Vikman, U. & Öckert, B. (2014). Decentralisering, Skolval och Fristående Skolor: Resultat och likvärdighet i svensk skola, IFAU Report, no. 2014:25, Available online: <http://www.ifau.se/globalassets/pdf/se/2014/r-2014-25-Decentralisering-skolval-och-friskolor.pdf> [Accessed 31 March 2016]
- Jacob, B. A. & Lefgren, L. (2007). What Do Parents Value in Education? An empirical investigation of parents' revealed preferences for teachers, *The Quarterly Journal of Economics*, vol. 122, no. 4, pp.1603–1637, Available online: <http://www.jstor.org.ludwig.lub.lu.se/stable/25098884> [Accessed 20 April 2016]
- Lindahl, E. (2007). Comparing Teacher's Assessments and National Test results: Evidence from Sweden, Working Paper, no. 2007:24, IFAU, Available online: <http://www.ifau.se/globalassets/pdf/se/2007/wp07-24.pdf> [Accessed 20 May 2016]
- MacLeod, W. B. & Urquiola, M. (2012). Competition and Educational Productivity: Incentives writ large, Discussion Paper, no. 7063, IZA, Available online: <http://ftp.iza.org/dp7063.pdf> [Accessed 5 April 2016]
- Sandström, M. & Bergström, F. (2005). School Vouchers in Practice: Competition won't hurt you. *Journal of Public Economics*, vol. 89, no. 2-4, pp.351–380. Available online:
<http://www.sciencedirect.com/science/article/pii/S0047272704000404>
[Accessed 2 April 2016]
- SIRIS. (2016). Om SIRIS, Available online:
<http://siriskolverket.se/siris/f?p=SIRIS:11:0> [Accessed 22 May 2016]

- Skolinspektionen. (2015). Ombedömning av Nationella Prov 2014: Processerna spelar roll, Available online: <https://www.skolinspektionen.se/sv/Beslut-och-rapporter/Publikationer/Granskningsrapport/Omrattning-av-nationella-prov/ombedomning-av-nationella-prov-2014/> [Accessed 7 May 2016]
- Skolverket. (2012). Betygsinflation – Betygen och den faktiska kunskapsutvecklingen, Available online: <http://www.skolverket.se/publikationer?id=2804> [Accessed 28 April 2016]
- Skolverket. (2014). Private Actors in Preschools and Schools: A mapping of independent education providers and owners in Sweden, Report, no. 410, Skolverket, Available online: <http://www.skolverket.se/publikationer?id=3356> [Accessed 6 May 2016]
- Skolverket. (2016) Redovisning av Uppdrag om Relationen Mellan Provresultat och Betyg i Grundskolans Årskurs 6 och Årskurs 9, Available online: <http://www.skolverket.se/publikationer?id=3618> [Accessed 4 April 2016]
- Vlachos, J. (2010). Betygets Värde: En analys av hur konkurrens påverkar betygssättningen vid svenska skolor, Uppdragsforskningsrapport, no. 2010:6, Konkurrensverket: Stockholm, Available online: http://www.konkurrensverket.se/globalassets/publikationer/uppdragsforskning/forsk_rap_2010-6_betygens_varde.pdf [Accessed 15 April 2016]
- Vlachos, J. (2011). Friskolor i förändring, in Hartman, L. (ed), *Konkurrensens Konsekvenser: Vad händer med svensk välfärd?*, SNS Förlag: Stockholm, pp.66-110, Available online: http://www.sns.se/sites/default/files/konkurrensens_konsekvenser_pod_2.pdf [Accessed 10 April 2016]
- Westerlund, J. (2005). *Introduktion Till Ekonometri*, Lund: Studentlitteratur

Wikström, C. & Wikström, M. (2005). Grade Inflation and School Competition: An empirical analysis based on the Swedish upper secondary schools, *Economics Of Education Review*, vol. 24, no. 3, pp.309-322,
Available online:
<http://www.sciencedirect.com/science/article/pii/S0272775704000895>
[Accessed 9 April 2016]

Wooldridge, J. M. (2009). *Introductory Econometrics: A modern approach*, 4th edn, Mason, OH: South Western, Cengage