UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master Degree Project in Economics

Do Swedish upper Secondary Schools Compete through Grade Inflation?

Samuel Pedersen

Supervisor: Johan Stennek
Master Degree Project No. 2016:160
Graduate School

Do Swedish upper secondary schools compete through grade inflation?

September 27, 2016

Author: Samuel Pedersen

Supervisor: Johan Stennek

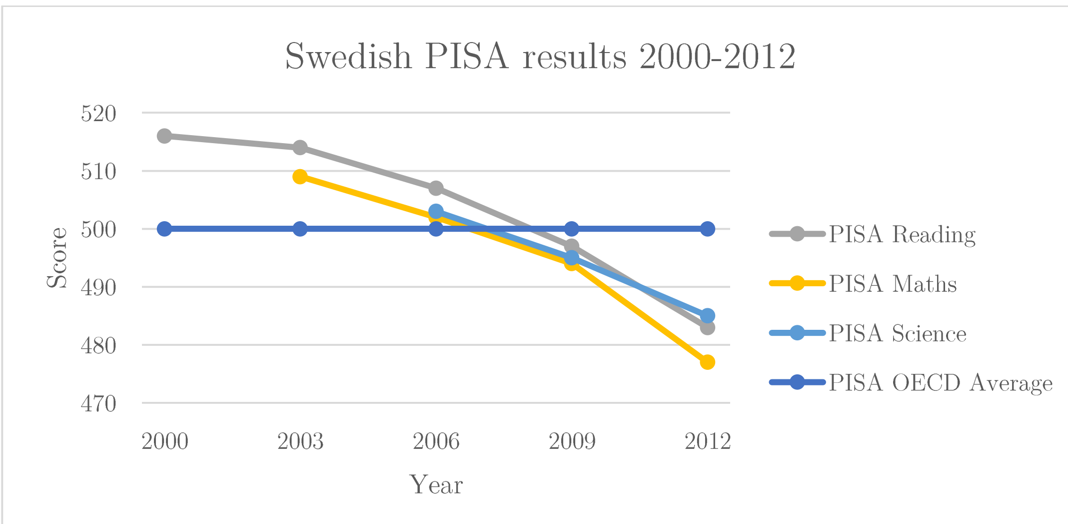
Abstract

This paper studies the relationship between school competition and grade inflation in Swedish upper secondary schooling. It contributes to the literature by introducing a new measure of grade inflation not previously used in relation to competition, as well as a new measure of school competition. The results indicate that competition does not reliably explain grade inflation but that there are large differences in grade inflation between public and non-public schools and between for-profit and non-profit schools. The differences remain even when controlling for competitive environment.

1. Introduction

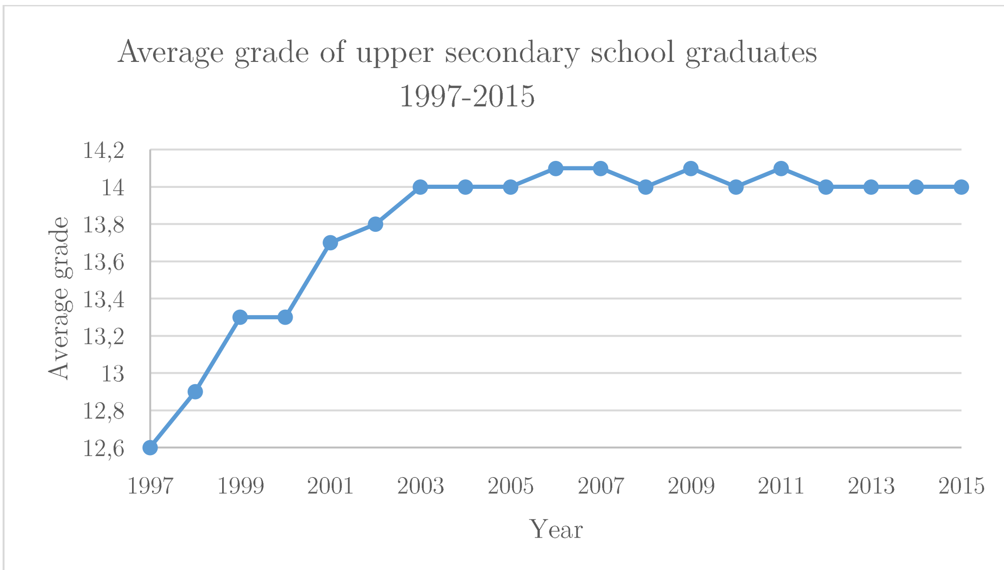
Grade inflation is a significant problem in the Swedish educational system. For the past decades there has been a negative trend in Swedish student performances in international standardized student achievement tests such as PISA and TIMSS.

Figure 1: Swedish PISA results 2000-2012



The PISA study tracks performances of 15 year olds in grade 9. Results from the study, seen in figure 1, shows how the academic performance of Swedish students has moved from being above average among the OECD countries, to below average. Not only has the average score decreased, but the share of students reaching the highest scores has also decreased greatly (OECD 2015). In the TIMSS study, another large international comparative study of student performances, the results of Swedish students in the final year of upper secondary schooling dropped greatly between 1995 and 2008 in magnitudes similar to those found in the PISA study (Skolverket 2009b).

Figure 2: Average grade 1997-2015



During the same period that Swedish students' results plummeted in international comparisons, average grades of graduating upper secondary school students experienced a different development. Figure 2 shows the development of average grades for upper secondary school graduates between 1997 and 2015. After a large increase between the years 1997 and 2003, average grades have remained more or less constant until today. In addition to this, the share of students graduating upper secondary school with the maximum average grade points has increased steadily during the past two decades (Vlachos 2010). The different trends of average grades compared to results in international tests indicate that the value of grades may have decreased during the last 20 years.

Grade inflation is important for a variety of reasons. High quality schools are greatly valued by households, indicated by research showing that housing prices are strongly positively correlated with different school quality attributes, especially test scores and grades (Black and Machin 2011). While school quality is difficult to observe for an individual household, average grades are one of the more visible quality measures. When grade inflation varies across schools and areas it becomes more difficult for households to observe which schools that offer the highest quality education. The results of choosing a poor quality school are significant, as changing schools often leads to negative outcomes for the student. There are also no possibilities for individual households to make schools accountable for poor quality education. This makes self-regulation in schooling in terms of grade inflation problematic. It also makes regulation by governmental actors difficult, as comparisons of student outcomes between schools that does not take into account grade inflation bears the risk of having biased results, as shown by Hinnerich and Vlachos (2016). It also makes the admission process to university studies less efficient as it distorts the system in favor of students with inflated grades. While being a problem of fairness and efficiency, grade inflation also has a larger economic significance for individuals. Diamond and Persson (2016) argue that students who get their test score grade inflated at age 15 on average receive higher grades in secondary school, have higher chances of completing secondary school on time and receive higher wages at age 23. This means that the long term effects of grade inflation are significant and negatively impacts the students who do not receive inflated grades.

Grade inflation became a problem in the early 90’s when the Swedish educational system went through a series of reforms with the intention of making it more efficient and more flexible to households’ demands. This was done by allowing households to choose schools to a larger extent than before. A school voucher system was introduced, allowing for non-public schools to compete more or less freely with public schools by basing schools' budgets on the number of students enrolled each semester. These non-public schools go by different names in the literature, but I will refer them as independent schools. The intention was that the system would give more resources to schools that were efficient providers of schooling services sought after by households. At the same time, schools that were inefficient producers would suffer reduced enrollment number and thus reduced budgets. The same reform that introduced the voucher system shifted the political and financial responsibility over primary and secondary education from the national level to the municipal level. Additionally, the grading system was changed. From the early 60's to the early 90's a norm reference grading system was used, where every class was to be graded according to the normal grade distribution with the average performance in the school indicating the mean. This was changed to a criterion based grading system, where students are graded according to certain performance achievement criteria specified on a national level. In the norm referenced grading system there was no possibility of grade inflation and there would have been small incentives for students striving for high grades to congregate at certain schools as the probability of achieving higher grades then would be lower. This coupled with the fact that only a handful of private schools existed meant that there was relatively little competition between schools in Sweden before the reforms. Starting in the early 00's, a massive expansion of independent schools occurred, increasing competition greatly. Before the reform, less than a percent of all upper secondary school students attended an independent school (Holmlund et al. 2014). In 2015 the figure was 26 percent. This massive expansion of independent schools makes Sweden an interesting area of study in terms of school competition. Sweden is also currently the only country in the world to allow for-profit schools to freely compete with publicly owned schools.

This paper asks the question whether if in a voucher system schools respond to competition by inflating grades, in order to attract more students and to avoid losing students to neighboring schools. I contribute to the existing literature by introducing an index of competition that takes into account sizes of schools and the distances between them. The method was originally used for estimating school competition in relation to student achievement outcomes by Misra and Chi (2011) and Misra et al (2012). I change the methodology slightly to make it better fit my data. Previous studies of competition and grade inflation in Sweden use measures of competition that only takes into account the share of students in non-public schools per municipality. This bears the risk of underestimating competition between public schools as well as missing variations in competitive climates within municipalities. In my model I treat public and non-public schools alike, motivated by the design of the voucher system, as it puts all schools under the same rules and budget constraints. Furthermore, I also use a measure of grade inflation that has not been used in relation to competition before, namely the difference between scores on national standardized tests that are graded locally compared to when re-graded centrally by teachers employed by the Swedish Schools Inspectorate (in Swedish: Skolinspektionen).

2. Literature review

The link between school competition and grade inflation has been studied in Sweden (Wikström and Wikström 2005, Vlachos 2010) and the US (Walsh 2010) with evidence pointing in various directions. Hinnerich and Vlachos (2016) uses grade inflation in the form of inflated grades on national standardized tests, which is possible because the tests are graded locally by the class teacher in Sweden. They show that grade inflation has a large influence on grade setting in Sweden and that it varies significantly between public and independent schools. Without taking into account grade inflation, independent schools appear to produce better educational achievements for students than public schools. After correcting for grade inflation the effects are instead negative, which shows that any comparison between schools that does not take into account grade inflation risks having the results compromised by omitted variable bias in the form of grade inflation.

Wikström and Wikström (2005) was the first publication to examine the relationship between grade inflation and competition in Sweden. Their measure of grade inflation is the difference between an upper secondary school student's grade points in the final year of upper secondary school and the student's score on SweSAT, a national aptitude test that is used when applying for university studies. They find no effect of competition on grade inflation, but they do find that independent schools inflate grades heavily compared to public schools. A weakness of their methodology is that their sample only contains data for individuals who graduated in 1997, a year when competition from independent schools was very minor compared to now. Only 9 out of 290 municipalities had at least one independent school in 1997 compared to 103 municipalities in 2015. Unlike Wikström

(2005), Vlachos (2010) finds no significant differences in grade inflation between independent schools and public schools when examining grade inflation in the years 2003-2008. He argues that the differences were largest in the turn of the century but since then have disappeared. Looking at grade inflation in both primary and secondary school, Vlachos uses three measures of grade inflation and finds significant positive effects of competition on grade inflation in all cases. The estimated magnitudes are small but underestimated, according to Vlachos. Walsh (2010) studies grade inflation in response to competition in the US, and finds a small effect of increased competition on grade cut-off points but no actual effect on grades. Walsh's explanation is that school administrators may lower grade cut-off points as a response to competition, but teachers are likely to readjust their grading to leave grades unchanged.

A potential weakness of several previous studies on school competition in Sweden (Sandström and Bergström 2005, Vlachos 2010, Böhlmark and Lindahl 2015) is the variable used to measure competition. All three studies use the share of students in independent schools per municipality as a measure, which is problematic for a number of reasons. It does not take into account competition between public schools, which was strengthened considerably following the 1990's reforms as all schools' budgets went from being fixed to being determined by the number of students enrolled each school year. If the entry of independent schools coincides with the exit of public schools in a municipality, it does not necessarily mean that competition is increasing and this is not reflected in the measure. Furthermore, a problem with using municipalities as markets is that they rarely function as single markets. Municipalities vary greatly in size leading to some municipalities consisting of several markets and some markets spreading over several municipalities. In the city of Stockholm, the market exceeds the Stockholm municipality while in larger, rural municipalities such as in northern Sweden all students are not able to commute throughout the entire municipality. In a 2016 report, Swedish National Agency for Education argues that there are 83 separate markets for upper secondary schooling in Sweden (Skolverket 2016b), compared to the 290 used in Wikström and Wikström (2005) and Vlachos (2010).

Public and independent schools are in theory identical actors in the education market under the Swedish school voucher system, as they receive the same budgets and are subject to the same regulation. However, Hoxby (2003) argues that the incentives of public and private actors in voucher system education do not align. This difference is most apparent between public schools and for-profit independent schools. Hoxby argues that the profit maximization motive of the for-profit school makes it so that these schools only provide quality high enough to attract the optimal amount of students, while keeping costs minimal. An extension to this theory in regards to grade inflation is that it is in the for-profit schools' interest to inflate grades, in order to increase quality without increasing costs. Unlike for-profit schools, public schools are not interested in profits, but they do have incentives to keep quality at such a level that they do not lose students to competitors. Thus they also have incentives to inflate grades in response to increased competition.

3. Methodology

3.1. Data

The school level data on grade inflation, number of students per school and other school characteristics is taken from the SiRiS database, provided by the Swedish National Agency for Education (Skolverket). The distances between schools used for creating the competition index were created using geographical coordinates for every upper secondary school in Sweden. I use two different measures of grade inflation, the first being the share of students per school and year who received a different grade on the course than they did on the national standardized test for the course. The national standardized tests are only taken in the subjects English, Swedish and mathematics. It is common for schools to both give lower and higher course grades to students, although higher grades are far more prevalent. The most common grade step that is inflated is from a failing grade to a passing grade (Skolverket 2016a). I take the difference between the share of students who received a higher grade in the course than on the exam and those who received a lower grade in the course of the exam. The resulting sum tells us the share of students per school and year that have received inflated grades. There are 6327 observations for the measure, where each observation is a school i in year t. Each observation contains the results of one cohort of test takers. The average level of grade inflation using this measure is 0.02 for English, 0.26 for mathematics and 0.16 for Swedish. Data for this measure is available for the spring term for the years 2012-2015 for all schools in Sweden with a few reservations. For schools with fewer than 10 students the measure is not reported. There is also a portion of schools that have not reported the measure to Skolverket despite it being an obligation. It is not clear why some schools choose not to report these statistics, and it is also unclear how this will impact the results. Another source of missing data comes from students that have not yet received either a course or test grade.

The test/course grade measure of grade inflation is susceptible to several sources of bias. Firstly, teachers are not required to give the same result on the student's national test grade and course grade. The test should however serve as a guideline as it covers most of the course's contents, and is a strong indicator of the student's level of knowledge in the course. Some level of deviation between the grade on the test and the course can be explained by factors such as whether different teachers value certain elements of the course differently and whether students who fail the test are given extra teaching resources in order to reach the course's knowledge criteria before the end of the semester. This means that not all schools with positive measures of grade inflation necessarily are involved in grade inflation. Another possible source of bias is that the grades on the standardized tests may be inflated as well. It is likely that a school that inflates course grades more also inflates test grades more. If this is the case it would lead to underestimation of grade inflation in this measure. It is difficult to estimate how large a share of students that would receive a higher course grade than test grade with no grade inflation.

The second measure of grade inflation is the difference between the results on the national standardized test when first graded locally by a school teacher compared to when re-graded centrally by teachers employed by the Swedish Schools Inspectorate. Similarly, to the previous grade inflation measure this data is openly available in the SiRiS database. The external teachers' evaluations provide a more objective view of the students' results, and as a result this measure of grade inflation will be less susceptible to bias of various forms. It is measured as the share of students per school that received a significantly higher score on the test when it was graded locally compared to when it was re-graded centrally by the Swedish Schools Inspectorate, significantly different meaning at least two grade steps difference on the six level scale A-F. The schools included in the sample are randomly selected from the pool of schools up for inspection in the following year. Data for this measure is only available once for the schools in the sample, collected either in 2014 or in 2015. The advantage of this measure of grade inflation compared to the previous is that it is less likely to suffer from the same bias as the previous measure, namely that both the test and course grade are set locally. The disadvantage of this measure is that there are far fewer observations and as can be seen in table 1 the characteristics of the schools in this sample vary slightly from the average school in Sweden. Smaller independent schools appear to be overrepresented in the sample, which may be problematic if these schools behave differently than public schools in response to competition. When analyzing this measure of grade inflation Hinnerich and Vlachos (2016) find significant variation in test grading practices between independent and public schools. The distributions of both measures of grade inflation for all three subjects English, mathematics and Swedish can be seen in appendix A.

The index I create in order to measure school competition is calculated in the following way

$$School\ competition\ index_{i,t} = \frac{1}{\ln(E_{i,t})} \Sigma \left(\frac{\ln(E_{j,t})}{d} \right)$$

where $E_{i,t}$ is the number of students enrolled in school i and $E_{j,t}$ is the number of students enrolled in each neighboring school j within 60 kilometers. d is the distance between school i and each school j. The measure takes into account the size of school i, distance between school i and all competitors j as well as size of competitors. I weight each competitor's size by the distance between the schools because the further away a competitor is situated the smaller is the probability that the two schools are competing over the same students. The distance limit where a school is no longer considered a competitor is set at 60 kilometers, motivated by the argument that this distance corresponds to around one hour of commuting in each direction. The average student is unlikely to be willing to travel daily further than this to school. The same distance limit was used in Misra and Chi (2011), which is the first paper to use a similar method to measure competition. Their competition index was created in the following way:

$$School\ competition\ index\ (Misra\ and\ Chi\ 2011)_{i,t} = \frac{1}{E_{i,t}} \Sigma \left(\frac{E_{j,t}}{d^2} \right)$$

There are two differences between the two indexes. Firstly, I use logarithmic sizes instead of linear. The reason for this is that there are large differences between size schools in my sample, with the smallest school in the sample having 17 students and the largest school having 2492 students enrolled. As there are such large differences in enrollment numbers, effects of size may be overestimated when using the linear effect. The second change I make is to remove the squared term from the distance parameter. I do this because while Misra and Chi's model was used to estimate competition in both primary and secondary schools, I only use it for upper secondary school. I believe it to be realistic that upper secondary school students are more capable of commuting longer distances to school compared to primary school children, and thus less sensitive to distances when choosing a school. $E_{i,t}$ is placed in the denominator of the formula as the model assumes that smaller schools will face stronger competitive pressure as their overall budget will suffer more relative to larger schools from losing students due to competition.

Figure 3: Competition index 2012-2015

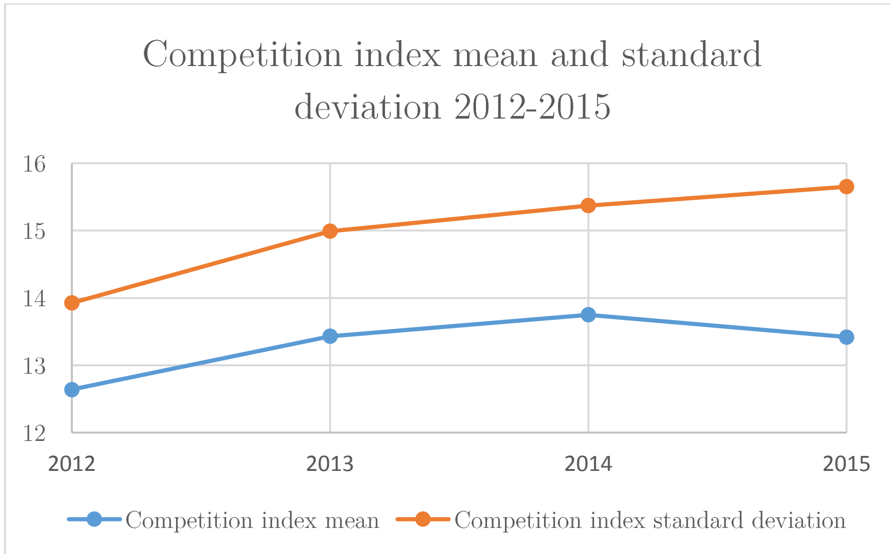


Figure 3 shows the development of the competition index over time. The average competition value increases steadily from 2012 to 2014 until there is a decrease in 2015. The reason for the decrease is unclear but could be explained by oversaturation in the market for schooling.

Figure 4: Competition index and two measures of grade inflation

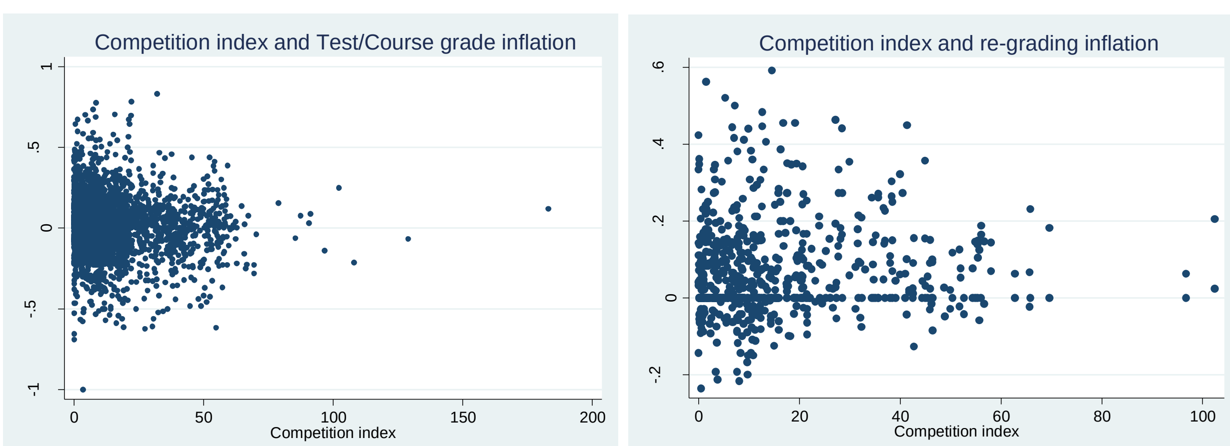
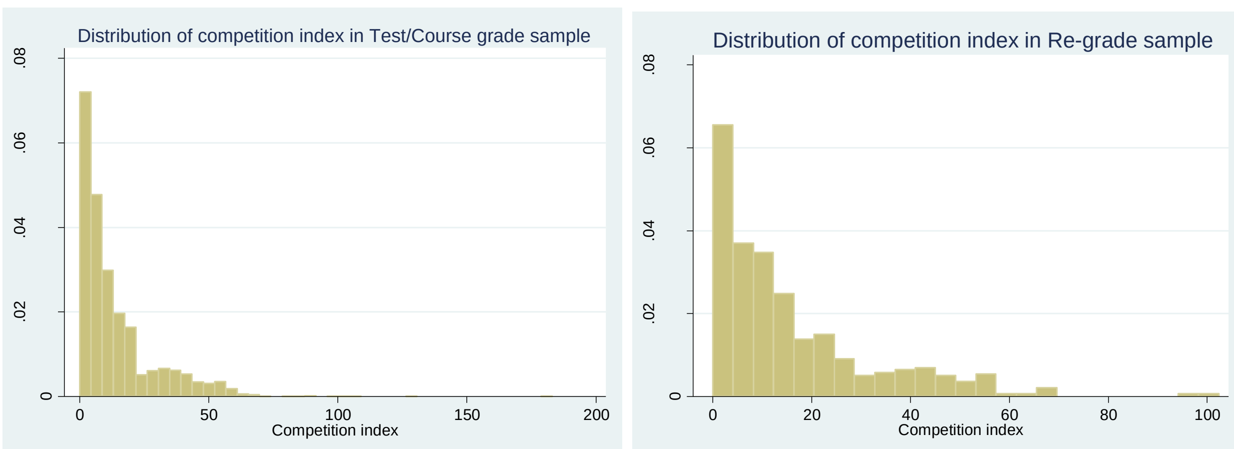


Figure 4 shows the relationship of the two main independent variables and the dependent variable in two sample groups. From a glance there does not appear to be an obvious relationship between the two variables.

Figure 5: Distribution of competition index in both samples



From the distribution of the competition index in the two sample groups shown in figure 5 we can see than the re-grade sample group appears to contain a lower share of values that are zero or close to zero. This indicates the possibility that rural schools are underrepresented and smaller schools overrepresented in the re-grade sample group.

Table 1: Summary statistics

	Test / Course grade sample group					Re-grading sample group					All schools				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Test / Course grade difference											.	.	.	.	.
English	3135	0.020	0.176	-1	0.833	.	.	.	.	.	.	.	.	.	.
Mathematics	1526	0.262	0.2	-0.357	1	.	.	.	.	.	.	.	.	.	.
Swedish	3019	0.156	0.193	-0.75	1	.	.	.	.	.	.	.	.	.	.
Re-grade difference											.	.	.	.	.
English	.	.	.	.	.	270	0.059	0.109	-0.216	0.444	.	.	.	.	.
Mathematics	.	.	.	.	.	125	0.001	0.016	-0.1	0.1	.	.	.	.	.
Swedish	.	.	.	.	.	272	0.112	0.145	-0.235	0.591	.	.	.	.	.
Competition index	6323	13.8	15.36	0	183.1	667	16	16.82	0	102.4	22707	13.2	15.3	0	183.1
Number of students	6327	306.1	242.8	19	2492	669	248.2	172.3	24	1235	23150	326.8	272.4	1	2492
Share women	6125	50.4	20.07	4	100	639	47.4	21.70	0.05	0.98	22204	49.40	19.30	0.04	1
Share of students with foreign background	5684	23.2	15.31	2	95	587	24.1	15.51	0.04	0.94	20722	23.80	16.20	0.02	1
Share of students with university educated parent	6303	47.7	16.67	8	93	665	44	17.68	0.1	0.91	22678	47.30	16.40	0.04	0.94
Independent dummy	6327	0.41	0.492	0	1	669	0.56	0.5	0	1	24197	0.34	0.48	0	1

Table 1 shows the summary statistics of the two sample groups as well as for all upper secondary schools in Sweden. The sample in the mathematics subject is smaller in both sample groups. The average test/course grade difference varies greatly between subjects, with the average inflation in English being 1.9% while it is 26.1% for mathematics. This large difference is partly due to the rare occurrence of teachers setting lower grades on the course than on the test in the mathematics subject. In Swedish and English around ten percent of students annually receive lower grades on the course than the on test, while the number is around two percent for mathematics. For the re-grade measure there are also large differences between the subjects with Swedish having the highest level of grade inflation with an average of 11% while the number is only 0.1% for mathematics. The difference can most likely be explained by the nature of the subjects, as questions on the mathematics test are less ambiguous and open for interpretation, increasing the likelihood that the original graders and the re-graders arrive at similar grades. The average value of 11% for Swedish means that there are large differences between how local teachers view the results compared to the re-grading teachers. The third category, *All schools,* shows the difference between the two samples and all schools, in order to see if the samples are representative of the population. We can see that the two samples consist of smaller schools, in terms of number of students, than the country average. Largely because of this the two samples also have larger averages in the competition index. It should be noted that the values of the competition index variable do not have meaning by themselves, and are only relevant in relation to the other values. The re-grade sample consists of 15 percent more independent schools than the test/course grade sample, and 22 more than the country average. The share of students who are women, of foreign background and with university educated parents appears to be quite consistent across the samples. Overall, smaller independent schools appear to be slightly overrepresented in both samples compared to the whole country, especially in the re-grade sample group. The reason could be that independent schools are subject to more oversight by the schooling authority, or that they target urban schools to a higher degree.

3.2. Empirical specification

$$\text{Test/Course } grade\ inflation_{i,t,c} = \propto_i + B_1 Competition\ index_{\,i,t} + B_2 Independent_i + B_3 Competition\ index_{i,t} * Independent_i + \gamma X_{i,t} + \delta Year_t + \epsilon_{i,t}$$

I use the above model to estimate the effects of competition on grade inflation for one of the two measures of grade inflation I use, namely test/course grade inflation. *Test/Course grade inflation* represents one of two measures of grade inflation, reported for course *c*, school *i* in year *t* with the value of the shares ranging from -1 to 1. Course *c* is one of three courses in Swedish, English and mathematics. These are the first courses in either subject at the upper secondary school level. I use these specific courses as they are available for both measures of grade inflation, and because the national standardized test in these courses are taken by all upper secondary school students due to the first courses in either subject being mandatory. This reduces the risk of sampling bias in the estimates. *Independent* is a dummy variable indicating whether the school is owned and run by a public or independent actor. I include this due to results from Wikström and Wikström (2005) and Hinnerich and Vlachos (2016) indicating that independent schools inflate grades more than public schools. I include the interaction term *Competition index*Independent* for the same reason. If the slope of B_1 significantly differs between independent and public schools it will appear in the B_3 estimate. *X* is a vector of school level variables including number of students enrolled at each school, share of students with foreign origin, share of students that are female and share of students with highly educated parents. These demographic variables are reported by the Swedish National Agency for Education as they are known to be correlated with student performance, and I also suspect them to be correlated with grade inflation. Share of students with foreign origin might influence grade inflation, especially in the subject Swedish. Share of students with highly educated parents might be an indicator of socioeconomic status which in turn may influence grade inflation. Share of women is an important determinant as women on average receive more inflated grades in all subjects (Skolverket 2016a). *No.Students* is the number of students enrolled at each school per year, which I include as there may be a difference between smaller and larger schools in terms of grade inflation. *Competition index* is the measure of competition I use for the model, explained in the data section. *Year* is the variable used to capture time effects. In the OLS estimation it is a series of year dummy variables, and in the panel data estimation it is a continuous time trend variable.

$$\text{Re-grade inflation} = \propto_i + B_1 Competition\ index\ _i + B_2 Independent_i + B_3 Competition\ index_i * Independent_i + \gamma X_i + \epsilon_i$$

I use the above model to estimate the relationship between competition and the second measure of grade inflation, namely re-grade inflation. This data set is cross sectional and the time component of the above model is therefore omitted. Other than this and the different dependent variables the models are identical.

One of the weaknesses of the model is the lack of data on teacher characteristics. Skolverket (2009a) argue that variation in how different types of teachers grade tests and courses is a more important determinant of grade inflation than competition. Another weakness is the short time period for which the data on grade inflation is available, which is 2012-2015 for the test/course measure and only one year for the re-grade measure.

Another issue has to do with the school level control variables. They are included as I believe them to be important determinants of grade inflation, but there is also a possibility the variables causing endogeneity through reverse causality. This could happen if schools not only compete to attract more students or not to lose students, but instead to attract a certain type of student that is less demanding of resources. Women, children of highly educated parents and non-immigrants are groups that all receive higher average grades than their opposites. If schools compete for these types of students using grade inflation as a tool it will lead to bias in the estimates due to reverse causality. Even if this is the case, it is unclear how large the effects would be and whether it would challenge the legitimacy of the results. I will report the estimation with and without the school level control variables to see whether they have a significant effect on the results.

4. Results

Table 2: Pooled OLS estimation using test/course grade inflation

	English			Mathematics			Swedish		
VARIABLES	(1) Pooled OLS Test/Course	(2) Pooled OLS Test/Course	(3) Pooled OLS Test/Course	(4) Pooled OLS Test/Course	(5) Pooled OLS Test/Course	(6) Pooled OLS Test/Course	(7) Pooled OLS Test/Course	(8) Pooled OLS Test/Course	(9) Pooled OLS Test/Course
Competition index	-0.00111***	-0.00158***	-0.00153***	-0.00218***	-0.000111	4.70e-06	-0.000232	-0.000240	-0.000315
	-0.00036	(0.000443)	(0.000440)	(0.000507)	(0.000566)	(0.000538)	(0.000352)	(0.000404)	(0.000376)
Independent	0.0161*	0.00988	0.00323	0.0334**	0.0629***	0.0501***	-0.00350	0.0125	0.0255**
	-9.78E-03	(0.0108)	(0.0107)	(0.0162)	(0.0172)	(0.0162)	(0.0106)	(0.0122)	(0.0114)
Comp. Index * Indep.	1.32E-05	0.000263	0.000190	0.000856	-0.000545	-0.000626	-0.000114	-0.000558	-0.000439
	-0.000468	(0.000515)	(0.000511)	(0.000683)	(0.000685)	(0.000647)	(0.000482)	(0.000514)	(0.000481)
Control variables	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year controls	No	No	Yes	No	No	Yes	No	No	Yes
Observations	3060	2,598	2,598	1,525	1,434	1,434	3,017	2,571	2,571
R-squared	0.0082	0.034	0.051	0.018	0.066	0.165	0.001	0.004	0.131
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1									

Table 2 presents the results of the OLS estimation of the test/course grade inflation measure. The first column of each subject does not control for time effects or school level characteristics, the second column controls for school level characteristics but not time and the third column controls for both factors. Before including school controls competition is statistically significant at the 1 percent level for two out of three subjects. Interestingly the relationship is negative. When including school and year controls the negative effects disappear from the mathematics estimation but remains in the results for English. The value of R-squared is much lower in English than in the other two subjects, indicating that grade inflation in English differs from the other subjects. The dummy variable *Independent* is statistically significant for mathematics and Swedish when using the control variables and year controls. The effects are positive and quite large, and the interpretation of the results is that independent schools on average give 5.01 percentage points higher grades in mathematics and 2.55 in Swedish. This is a large effect as the average grade inflation for mathematics is 15.6% and 26.1% for Swedish. The interaction variable *Competition index*Independent* is statistically insignificant in all three specifications, meaning that the slope of *Competition index* does not differ significantly between public and independent schools in the sample. As it is impossible to control for all omitted factors through control variables, there is a strong likelihood that the estimates suffer from endogeneity through omitted variable bias both of the time-variant and time-invariant variety. Because of this I should be careful when interpreting the results.

Table 3: Panel data estimation using test/course grade inflation

	English		Mathematics		Swedish	
	(1)	(2)	(4)	(5)	(6)	(7)
	Fixed effects	Random effects	Fixed effects	Random effects	Fixed effects	Random effects
VARIABLES	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course
Competition index	0.000756	-9.89e-05	0.00153	-0.000197	0.000900***	-0.00127*
	(0.00104)	(0.000316)	-0.0032	-0.000668	(0.000336)	(0.000646)
Comp. Index * Indep.	0.000905	-0.000542	-0.00589*	-0.000725	-0.00106	-1.98e-05
	(0.00190)	(0.000560)	-0.00327	-0.0009	(0.000741)	(0.000510)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3060	2,598	1,525	1,434	3,017	2,571
R-squared	0.0082	0.034	0.018	0.066	0.001	0.004
Number of schools	910	910	543	543	917	917
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1						

Table 3 shows the result of the panel data estimation of the test/course grade measure using fixed effects and random effects. I use standard errors clustered by the school markets created by Skolverket (2016b). In order to determine whether fixed or random effects is the best estimator I turn to the Hausman test, which tests whether the estimates of the two methods are significantly different. If they are different, then random effects are inconsistent and fixed effects is preferred. If they are not, random effects is the most efficient estimator. The *Independent* variable is omitted in the fixed effects estimation as it does not vary over time. The resulting p-values of the Hausman test are 0.28 for English, 0.03 for mathematics and 0.22 for Swedish. This means that random effects is inconsistent and we should instead rely on fixed effects for all subjects other than mathematics. In addition to this, the assumptions supporting the random effects estimator are not suitable for this estimation, as it is likely that the fixed effects are correlated with the explanatory variables. For these reasons I believe fixed effects to be the best estimator for the estimation. The fixed effects estimation in column 5 is the only estimation to provide positive effects of competition that are significant. The magnitude of the effects are however very small, as an increase of competition by one standard deviation would lead to an increase in grade inflation of 0.002 percentage points. The negative effects estimated using OLS can no longer be seen. The results of the fixed effects estimation are more reliable than that of the OLS estimation. The control variables do control for certain school level characteristics as well as any linear time trend, but it is likely that there are other time-variant omitted variables that could be leading to endogeneity in the estimates. One example of an important time-variant effect that is not controlled for is teacher characteristics that change over time.

Table 4: Re-grading estimation using OLS

	English		Mathematics		Swedish	
	(1)	(2)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	OLS	OLS	OLS
VARIABLES	Re-grade	Re-grade	Re-grade	Re-grade	Re-grade	Re-grade
Competition index	0.000215	-0.000157	-4.28e-05	-3.58e-05	2.66e-05	-0.000355
	(0.000354)	(0.000435)	(5.52e-05)	(8.09e-05)	(0.000648)	(0.000764)
Independent	0.0581***	0.0607***	-0.00292	-0.00717**	0.0988***	0.0936***
	(0.0176)	(0.0183)	(0.00413)	(0.00344)	(0.0233)	(0.0280)
Comp. Index * Indep.	9.52e-05	0.000321	-9.16e-06	8.67e-05	-0.000282	-0.000246
	(0.000650)	(0.000634)	(9.87e-05)	(9.23e-05)	(0.000864)	(0.000927)
Control variables	No	Yes	No	Yes	No	Yes
Observations	270	228	125	108	272	229
R-squared	0.075	0.147	0.017	0.049	0.101	0.152
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1						

Table 4 presents the results of the OLS estimation of competition on grade inflation in the re-graded tests. The difference before and after adding control variables is very small for English and Swedish. The coefficient for *competition index* becomes negative when including controls, but remains insignificant in all three specifications. For mathematics the effects of competition go from insignificant to being significant, although negative. The negative effect of a school being independent is not found for any other subject using any measure of grade inflation. The reason for it could be that the sample of the mathematics subject is very small, containing 108 observations of which only 12 has a value other than zero. The distribution can be seen in figure A5 in appendix A. Because of the small variation in the mathematics variable is it is very likely that the negative coefficient for mathematics suffer from bias and therefore are not trustworthy.

5. Robustness

In order to determine how the competition index I created compare to the measure of school competition used in other school competition studies (Hoxby 1994, Vlachos 2010, Böhlmark and Lindahl 2015) I estimate the model presented in table 2 and 3 again, this time replacing test/course grade inflation with the share of students enrolled in independent schools per municipality as the variable measuring competition.

Table 5: Share of independent students and grade inflation using OLS

	English			Mathematics			Swedish		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Pooled OLS	Pooled OLS	Pooled OLS	Pooled OLS	Pooled OLS	Pooled OLS	Pooled OLS	Pooled OLS	Pooled OLS
VARIABLES	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course
Share independent	-0.0610***	-0.0943***	-0.0949***	-0.0998***	0.00225	0.00668	0.00715	-0.000322	-0.0124
	(0.0216)	(0.0245)	(0.0242)	(0.0354)	(0.0388)	(0.0360)	(0.0243)	(0.0274)	(0.0252)
Independent	0.00960	0.00395	-0.00686	0.0808***	0.0994***	0.0858***	-0.00937	0.0155	0.0312*
	(0.0145)	(0.0171)	(0.0170)	(0.0265)	(0.0290)	(0.0275)	(0.0155)	(0.0188)	(0.0179)
Share indep.*Indep.	0.0198	0.0366	0.0448	-0.0920	-0.128*	-0.129*	-3.24e-05	-0.0434	-0.0392
	(0.0381)	(0.0441)	(0.0439)	(0.0689)	(0.0717)	(0.0682)	(0.0407)	(0.0476)	(0.0445)
Control variables	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year controls	No	No	Yes	No	No	Yes	No	No	Yes
Observations	3,058	2,596	2,596	1,525	1,434	1,434	3,014	2,568	2,568
R-squared	0.003	0.028	0.045	0.018	0.068	0.167	0.000	0.003	0.129

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 5 presents the results of the pooled OLS estimation when replacing the competition index with the share of students in independent schools per municipality. The signs and statistical significance of the parameters are almost identical to those in table 2, except for the interaction term, which becomes statistically significant at the 10 percent level in the mathematics estimation with control variables. The largest difference between these results and those in table 2 is the magnitude of the effects. The effects of increasing *Share independent* by one standard deviation is almost two times larger than those produced by the competition index, in all three specifications. A similar difference in magnitude of the estimates can also be seen for the *independent* estimate.

Table 6: Share of independent students and grade inflation using panel data

	English		Mathematics		Swedish	
VARIABLES	(1) Fixed effects Test/Course	(2) Random effects Test/Course	(4) Fixed effects Test/Course	(5) Random effects Test/Course	(6) Fixed effects Test/Course	(7) Random effects Test/Course
Competition index	0.0595	-0.0880***	-0.124	-0.0187	-0.00435	0.0158
	(0.176)	(0.0327)	(0.286)	(0.0462)	(0.272)	(0.0366)
Comp. Index * Indep.	-0.0146	0.0553	-0.0379	-0.0970	0.0212	-0.0697
	(0.228)	(0.0614)	(0.405)	(0.0778)	(0.237)	(0.0702)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3060	2,598	1,525	1,434	3,017	2,571
R-squared	0.0082	0.034	0.018	0.066	0.001	0.004
Number of schools	910	910	543	543	917	917
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1						

Table 6 shows the results of the panel data estimation using the *share independent* measure of competition. Again, fixed effects is preferred to random effects. The estimates are insignificant in all fixed effects estimations, which is the same result as when estimating the model with the competition index variable. The results in both table 5 and 6 very closely resemble those in the original specifications. This gives me confidence in the accuracy of the competition index as a measure of school competition. The advantages of the competition index compared to using share of independent school students per municipality is that the latter measure does not allow for variation in levels of competition within municipalities, and also does not include competition between public schools. The competition index allows for both these factors which should make it a more accurate measure of competition than share of independent students per municipality.

The relationship between competition and grade inflation might be nonlinear. Due to the large number of zero values in both the dependent and independent variables I am no able to estimate a logarithmic functional form, as it would lead to missing values and thus omit a significant portion of the sample. Instead I add a quadratic competition term to the specification, motivated by evidence from other industries indicating that the strongest effects of competition occur in the early stages of competition, meaning for example when a market changes from having one actor to two or three actor. As each additional competitor enters the market, the effects of competition diminish (Bresnahan and Reiss 1991). If we adapt this argument to the education sector we could expect to see a quadratic relationship with diminishing effects of competition. When running the model with a squared competition term the effects of competition change slightly in magnitude, but the estimates do not become statistically significant. The results of the estimation are presented in appendix B.

6. Extension: Analysis of independent schools by owner type

A natural follow-up after finding that there are significant differences in grade inflation between independent schools and public schools is to examine how independent schools differ. I do this by making a distinction between for-profit and non-profit independent schools, based on owner type. The by far most common owner type of independent schools in Sweden is joint stock companies. I categorize these schools as for-profit as it expressed goal of these organisations is to extract profits to the owners. Non-profit schools are generally owned by foundations or non-profit associations. 89% of independent schools in the sample are for-profit schools and 11% non-profit.

I estimate the same model as in the original specification, except that I replace the *independent* variable and the *Competition index*independent* interaction term with *For-profit* and *Non-profit* dummy variables as well as *Competition index*For-profit* and *Competition index*Non-profit* interaction terms. As I am interested in the coefficients of binary variables, fixed effects is not appropriate and I will therefore use OLS for both measures of grade inflation.

Table 7: Competition and test/course grade inflation by owner type using OLS

	English	Mathematics	Swedish
	(1)	(2)	(3)
	Pooled OLS	Pooled OLS	Pooled OLS
VARIABLES	Test/Course	Test/Course	Test/Course
Competition index	-0.00152***	-0.000112	-0.000333
	(0.000440)	(0.000544)	(0.000377)
For-profit	0.00448	0.0523***	0.0288**
	(0.0111)	(0.0173)	(0.0118)
Non-profit	-0.00376	0.00999	-0.0136
	(0.0318)	(0.0329)	(0.0342)
Comp. Index*For-profit	6.42e-05	-0.000438	-0.000539
	(0.000530)	(0.000704)	(0.000505)
Comp. Index*Non-profit	0.000904	-0.000553	0.000601
	(0.000925)	(0.000916)	(0.000944)
Control variables	Yes	Yes	Yes
Year controls	Yes	Yes	Yes
Observations	2,598	1,434	2,571
R-squared	0.051	0.168	0.131

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 8: Competition and re-grade inflation by owner type using OLS

	English	Mathematics	Swedish
	(1)	(2)	(3)
	OLS	OLS	OLS
VARIABLES	Re-grade	Re-grade	Re-grade
Competition index	-0.000157	-3.10e-05	-0.000225
	(0.000444)	(8.20e-05)	(0.000766)
For-profit	0.0633***	-0.00755**	0.0933***
	(0.0190)	(0.00354)	(0.0270)
Non-profit	0.0365	-0.00145	-0.0857
	(0.0460)	(0.00387)	(0.0791)
Comp. Index*For-profit	0.000214	8.02e-05	-0.000385
	(0.000686)	(8.64e-05)	(0.000944)
Comp. Index*Non-profit	0.00124	1.41e-06	0.00276
	(0.00137)	(0.000168)	(0.00203)
Control variables	Yes	Yes	Yes
Year controls	Yes	Yes	Yes
Observations	2,598	1,434	2,571
R-squared	0.051	0.168	0.131

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

In both table 7 and 8 the coefficients of the interaction terms are insignificant. This means that there is no significant difference between for-profit and non-profit schools in the slope of the competition index variable. The beta estimates of the dummy variable *For-profit* are statistically significant in the same way *independent* was significant in the previous estimations. The similarity between the results is expected as the for-profit schools constitute 89% of independent schools in the sample. The coefficient for *Non-profit* is however insignificant in all specifications using both measures of grade inflation. This means that there is a significant difference between the two groups in the intercept, indicating that for-profit schools give higher grades on average in the sample even when controlling for competition.

7. Discussion and conclusion

The results showed that effects of competition on grade inflation varies across subjects but is insignificant on average. For the English subject, the effects were negative and significant at the one percent level in the OLS estimation but insignificant in the fixed effects estimation. For mathematics the effects were insignificant in all specifications. For Swedish the effects were insignificant using OLS but later became positive and significant at the one percent level when using fixed effects. For the re-grade inflation measure competition was insignificant for all subjects in all specifications. The overall effect of competition on grade inflation can therefore be considered to be weak and mostly insignificant.

The effects of the variable *Independent* could only be estimated using OLS. The variable was strongly positively correlated with grade inflation for two out of three subjects using both measures of grade inflation. This result is supported by previous findings by Wikström and Wikström (2005) as well as Hinnerich and Vlachos (2016). The estimate for *Independent* was statistically significant at the 1 percent level for two out of three subjects in the test/course grade estimation and for all subjects in the re-grade inflation estimation. The magnitude of the effects of *Independent* were estimated to be between 2.5 and 5.0 percentage points for the test/course grade inflation measure and between 6.0 and 9.4 percentage points for the re-grade inflation measure. As the mean values of the test/course measure is between 0.019 and 0.261 and the re-grade inflation mean value is between 0.00 and 0.11, the effect of a school being independent appears to have a very strong impact on grade inflation, especially for the re-grade measure. However, due to the potential endogeneity in the results as they come from OLS estimation, which we know to suffer from omitted variable bias, we should be careful before drawing any strong conclusions from the estimates.

For-profit and non-profit schools do not significantly differ in terms of how competition affects their level of grade inflation. However, we see a significant difference in the intercept very similar to the difference between public schools and independent schools. This may indicate that non-profit independent schools are more similar to public schools than for-profit schools in their behavior concerning grade inflation. One theoretical explanation of this could be the importance of the profit motive as a driving force of grade inflation, as profit seeking schools will compete more intensely in order to attract students. However, as the model is estimated using OLS there is omitted variable bias in the results, meaning we should be cautious of interpreting the results too far. Further research on the importance of the profit motive may give additional insights into this relationship. The large increase of for-profit independent schools in the past two decades makes this a highly relevant question for the future of Swedish education.

One weakness of the model I use is that it assumes that schools themselves determine grade inflation when in reality it is the teachers that set both test and course grades. Skolverket (2009a) argue that teachers' varying grading practices are more important than differences between schools as an explanation of grade inflation. However, as I do not have data that includes information about teachers I cannot include this in the model. If teachers that are likely to inflate grades are more prevalent in certain types of schools, such as independent schools, or in different areas of Sweden it will lead to bias in the estimates. Adding teacher characteristics to the model would be a great addition that might explain some of the differences found between public and independent schools.

In conclusion, the results presented in this paper suggests that the effect of competition on grade inflation is on average insignificant. If true, it means that schools do not compete through inflating grades and that grade inflation is determined by other factors. There is also weak evidence indicating that independent schools give much more inflated grades than public schools do, even after controlling for competitive climate. This finding is in line with results from Wikström and Wikström (2005), Walsh (2010) and Hinnerich and Vlachos (2016). Furthermore, for-profit independent schools appear to give more inflated grades compared to non-profit independent schools even when controlling for competitive environment. This is a result that has not been found in previous literature.

References

Black, S., & Machin, S. (2011). Housing valuations of school performance. *Handbook of the Economics of Education*, *3*, 485-519.

Bresnahan, T. F., & Reiss, P. C. (1991). Entry and competition in concentrated markets. *Journal of Political Economy*, 977-1009.

Diamond, R., & Persson, P. (2016). *The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests* (No. w22207). National Bureau of Economic Research.

Hinnerich, BT., & Vlachos, J. (2016). *The impact of upper-secondary voucher school attendance on student achievement. Swedish evidence using external and internal evaluations* (No. 2016: 2). Stockholm University, Department of Economics.

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*. Institutet för arbetsmarknads-och utbildningspolitisk utvärdering (IFAU).

Hoxby, C. M. (1994). *Do private schools provide competition for public schools?* (No. w4978). National Bureau of Economic Research.

Hoxby, C. M. (2003). School choice and school productivity. Could school choice be a tide that lifts all boats?. In *The economics of school choice* (pp. 287-342). University of Chicago Press.

Misra, K., & Chi, G. (2011). Measuring public school competition from private schools: A gravity-based index. *Journal of Geographic Information System*,*3*(4), 306.

Misra, K., Grimes, P. W., & Rogers, K. E. (2012). Does competition improve public school efficiency? A spatial analysis. *Economics of Education Review*,*31*(6), 1177-1190.

OECD (2015): *Improving schools in Sweden: An OECD perspective,* retrieved from http://www.oecd.org/edu/school/improving-schools-in-sweden-an-oecd-perspective.htm [Accessed 14 Aug. 2016]

Skolverket (2009a): *Likvärdig betygssättning i gymnasieskolan?* Rapport 338

Skolverket (2009b): *TIMSS Advanced 2008: Svenska gymnasieelevers kunskaper i avancerad matematik och fysik i ett internationellt perspektiv.* Rapport 336

Skolverket (2016a): *Skillnaden mellan provresultat och kursbetyg i gymnasieskolan 2015.* Dnr 2016:00318

Skolverket (2016b): *Förändringar i skolmarknadens geografi mellan 2009 och 2014.* Rapport 439

Vlachos, J. (2010). Betygets värde. En analys av hur konkurrens påverkar betygssättningen vid svenska skolor. *Konkurrensverket uppdragsforskningsrapport*, *6*.

Walsh, P. (2010). Does competition among schools encourage grade inflation? *Journal of School Choice*, *4*(2), 149-173.

Wikström, C., & Wikström, M. (2005). Grade inflation and school competition: an empirical analysis based on the Swedish upper secondary schools. *Economics of education Review*, *24*(3), 309-322.

Appendix A

Figure A1: Distribution of test/course grade inflation for English

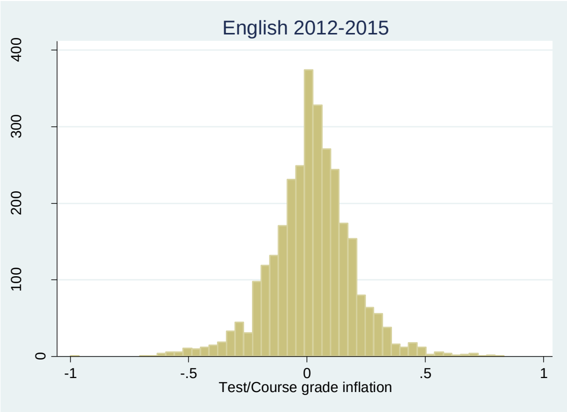


Figure A2: Distribution of Test/Course grade inflation for mathematics

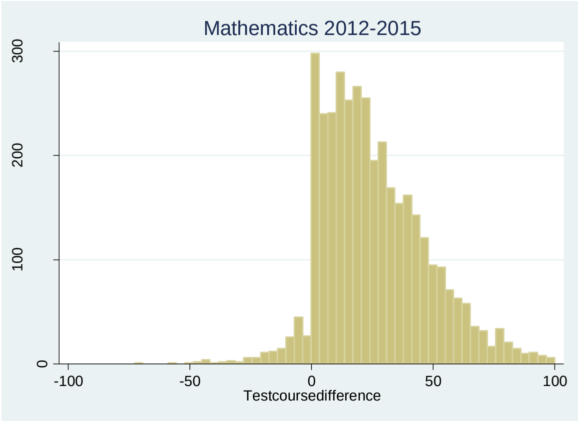


Figure A3: Distribution of Test/Course grade inflation for Swedish

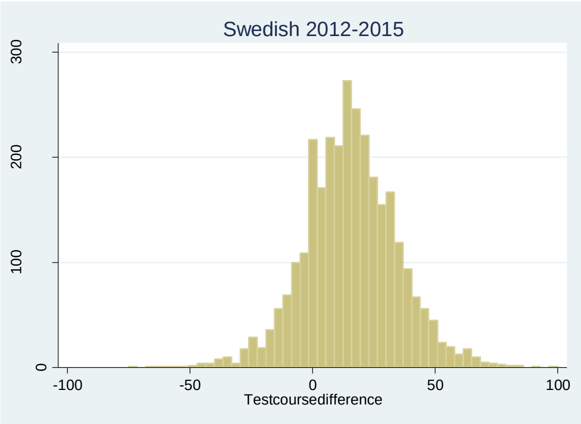


Figure A4: Distribution of re-grade inflation for English

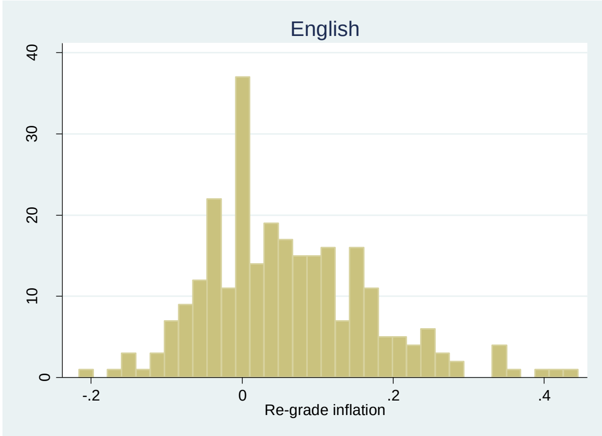


Figure A5: Distribution of re-grade inflation for mathematics

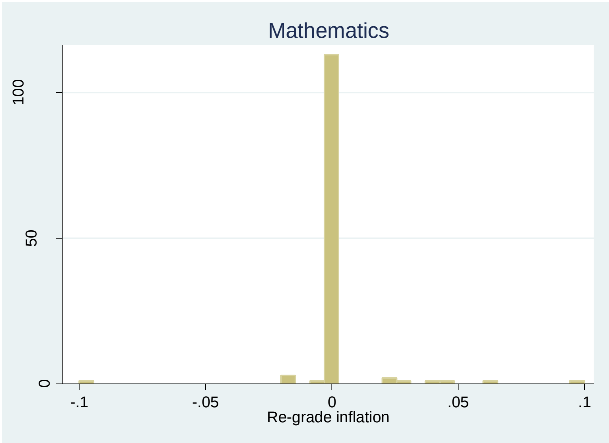
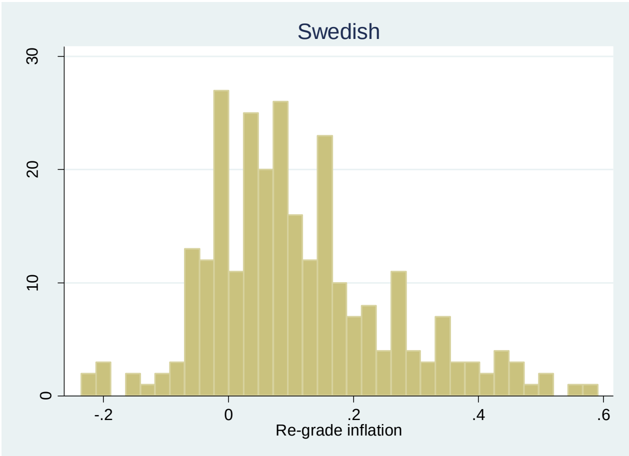


Figure A6: Distribution of re-grade inflation for Swedish



Appendix B

Table B1: Quadratic relationship between competition and test/course grade inflation using OLS

	English	Mathematics	Swedish
VARIABLES	(1) Pooled OLS Test/Course	(4) Pooled OLS Test/Course	(7) Pooled OLS Test/Course
Competition index	-0.00313***	-0.000107	-0.00108**
	(0.000533)	(0.000866)	(0.000485)
Competition index ^2	2.35e-05***	2.25e-06	1.13e-05***
	(6.18e-06)	(1.19e-05)	(4.05e-06)
Comp. Index * Indep.	0.000168	-0.000672	-0.000447
	(0.000472)	(0.000693)	(0.000465)
Independent	0.0111	0.0513***	0.0292**
	(0.0109)	(0.0175)	(0.0114)
Control variables	Yes	Yes	Yes
Year controls	Yes	Yes	Yes
Observations	2,598	1,434	2,571
R-squared	0.057	0.165	0.132

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

When adding the quadratic term *Competition index^2* to the OLS specification using the test/course grade inflation measure the β_1 estimate increases in magnitude by a factor of two for English and becomes negative and significant at the 5 percent level for Swedish. The coefficients of the squared term are significant and positive. This is the opposite effect of what would occur if competition had a diminishing effect on grade inflation. There appears to be significant non-linear effects in the relationship between competition and grade inflation.

Table B2: Quadratic relationship between competition and grade inflation using panel data

	English		Mathematics		Swedish	
	(1)	(2)	(4)	(5)	(6)	(7)
	Fixed effects	Random effects	Fixed effects	Random effects	Fixed effects	Random effects
VARIABLES	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course	Test/Course
Competition index	0.000219	-0.00280***	-0.00147	-0.000354	0.00380	-0.000682
	(0.00194)	(0.000752)	(0.00360)	(0.000851)	(0.00382)	(0.000507)
Competition index ^2	3.61e-06	1.95e-05***	6.59e-05**	3.10e-06	-1.61e-05	7.65e-06**
	(8.56e-06)	(5.34e-06)	(2.94e-05)	(8.86e-06)	(1.68e-05)	(3.43e-06)
Comp. Index * Indep.	-0.000822	0.000119	-0.0118**	-0.000791	-0.000103	-0.000498
	(0.000792)	(0.000388)	(0.00479)	(0.000912)	(0.00213)	(0.000631)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3060	2,598	1,525	1,434	3,017	2,571
R-squared	0.0082	0.034	0.018	0.066	0.001	0.004
Number of schools	910	910	543	543	917	917
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1						

In the panel data estimation, adding the squared term changes the size and sign of the β_1 estimates significantly. The β_1 for Swedish in column 6 was statistically significant at the 1 percent level in the linear panel data estimation, but is now insignificant and much smaller in magnitude. The estimates for the squared competition term is significant in one of the fixed effects estimations, and again the signs of the coefficients are positive, which indicates that the effects of competition are positively quadratic.

Table B3: Quadratic relationship between competition and re-grade inflation using OLS

	English	Mathematics	Swedish
	(1)	(2)	(3)
	OLS	OLS	OLS
VARIABLES	Re-grade	Re-grade	Re-grade
Competition index	-8.72e-05	-6.49e-06	-0.00169
	(0.00102)	(0.000127)	(0.00142)
Competition index ^2	-1.01e-06	-5.92e-07	1.91e-05
	(1.07e-05)	(1.75e-06)	(1.56e-05)
Independent	0.0605***	-0.00759**	0.0998***
	(0.0185)	(0.00348)	(0.0285)
Comp. Index * Indep.	0.000310	0.000104	-0.000230
	(0.000659)	(9.93e-05)	(0.000883)
Control variables	Yes	Yes	Yes
Observations	228	108	229
R-squared	0.147	0.050	0.156
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1			

When adding a squared term to the re-grade measure the magnitude of the β_1 estimates changes significantly, again suggesting that there are non-linear effects between competition and grade inflation in the sample.