Does gender and ethnic background matter when teachers set school grades?
-Evidence from Sweden

Erica Lindahl

Uppsala University
Preliminary draft – please do not quote
October, 2006

Abstract

Discrimination against girls, as well as foreign born students, is a hot topic both in the public debate and in the academic world. Schools and teachers are often said to be a source of stereotypes that harms girls. However, today girls outperform boys in school in many subjects in many countries. The first hypothesis tested is if girls and foreign-born teachers are more generously rewarded in final grades compared to boys and others, respectively. This hypothesis is also related to the gender and ethnical background of the teacher; do they act differently in this respect? In addition this paper tests if girls and foreign-born students perform better when the teacher is of the same gender or ethnical background. A unique data set from Sweden, including both final grades and comprehensive national test grades – where both oral and written abilities are tested - as well as information about both students and teachers, makes it possible to use a difference in difference strategy. The results suggest that in English, Swedish and Mathematics, both girls and foreign born students are more generously rewarded in final grades than boys and others, respectively. Male teachers are more generous than female teachers in Mathematics and foreign-born teachers are more generous than non foreign-born teachers in English, when controlling for school fixed effects. In Mathematics, girls and foreign-born students perform better on the test when the share of female and non foreign-born teachers increases at the school.

Erica Lindahl
Ph D student
Department of Economics
Uppsala University
Box 513
SE 751 20Uppsala
Tel:+46-18-4712257
Fax: +46-18-471 14 78

1 Introduction

Girls outperform boys with respect to grades. This is true for grade 9 students (15 years old) in Sweden in all subjects except for sports and health education (press release from the Swedish Agency of Education 15th December 2005). The gender gap in favour of girls in final grades is an international phenomenon; in UK, girls perform better than boys in most subjects at the age of seventeen (EOC, 2006). A majority of teachers in younger-aged schools are female and the feminization of the teacher profession has been raised as an explanation for the increasing gender gap in school performance (Holmlund and Sund, 2005). Several studies show that female teachers benefit girls (Dee, 2005; Bettinger E. P. and Long B. T., 2005; Ehrenberg *et al*, 1994). That is, girls perform better when the teacher is female.

A new alternative explanation for the gender gap in school performance is that girls are more generously rewarded in final grades than boys relative to their skills and performance. This reasoning could explain the increased gender gap in final grades that has emerged during the last years and it could be motivated by a concern about *not* discriminating girls. In the Education literature, many authors claim that boys are positively discriminated in the sense that teachers devote more time to boys than to girls (Einarsson, 1981; 1983; Bruzelius *et al.* 1981; Hultman, 1981). Scholars, in Education, claim that the middle-age school is a source of stereotype beliefs that harm girls (Bernard (1979), Dusek and Joseph (1983) and Tiedemann (2000).

Discrimination against minorities, such as foreign born, is a serious (and motivated) concern in the public debate. The school is an important arena in this context. Evidence from US shows that racial pairings of teachers and students influence how teachers allocate their time in the classroom (Ferguson, 1998 and Casteel, 1998). Efforts to recruit teachers with a foreign background are often motivated by the claim that those teachers benefit students with foreign background; teachers (unconsciously) allocate more time to students with a similar ethnical background. Dee (2004) shows that students perform better with a teacher of the same ethnical background. A common conclusion from this discussion and the similar one about gender is that effort is needed to guarantee that neither girls nor foreign-born students become discriminated in school. With respect to final grades it is possible that teachers compensate girls and foreign-born students in final grades in their effort to not discriminate these groups.

This study investigates how final grades are affected by the gender and ethnical background of the student and the teacher. Several hypothesises is tested. The first is if girls and foreign-born students are more generously rewarded in final grades compared to boys and others, respectively. The reason is a concern about not treat them unfairly. This hypothesis is also related to the gender and ethnical background of the teacher; do teachers act differently in this respect depending on gender and ethnical background? If female teachers benefit girls because they give girls more attention (time) in the class room, it is reasonable to believe that female teachers are less generous than their male colleagues when setting final grades on girls. That is, if girls are more generously rewarded than boys in final grades because of a concern about not discriminating girls, the *a priori* expectation is that female teachers (who benefit girls) are less generous than their male colleagues in this respect. A similar argument is applicable with respect to ethnical background. The argument to recruit teachers from an ethnical background in minority is that this should benefit students from the same ethnical background – students perform better with a teacher of the same ethnical background. Thus, if foreign-born students are more generously rewarded because of a concern about not discriminate those, the degree of compensation in final grades with respect to foreign-born students is probably related to the ethnical background of the teacher. In line with this argument, the *a priori* expectation is that foreign-born teachers are less generous in this respect than non foreign-born teachers.

Several studies from different countries show that students benefit from having a teacher of the same gender or ethnical background (Dee, 2005; Bettinger and Long, 2004; Ehrenberg *et al*, 1994; Dee (2001). However, no study presents supportive evidence for this hypothesis by using Swedish data. Thus, this hypothesis is also tested. Several explanations for this hypothesis have been raised in the literature; teachers serve as role models, teachers treat students differently depending on gender (ethnical background) of the student and of them selves. Identify the underlying mechanism is a difficult task and no paper to my knowledge has succeeded with that. This paper also tests the common (total) effect on student's performance of having a teacher of the same gender or ethnical background. However regardless of the underlying mechanism it is reasonable to believe that this positive effect (if it exists) work in the opposite direction of the hypothesis that teachers compensate students they believe they discriminate in one way or another; female (foreign-born) teachers benefit girls (foreign-born students) and thereby female teachers (foreign-born) have less incentive to compensate girls in final grades.

The gender (ethnical background) bias effect is identified by comparing results of comprehensive national tests – where both written and oral abilities are tested - with final grades from grade 9 in Sweden. In Sweden, finale grades are set by the teacher alone. When setting the grade, the teacher should take all available information about the student's knowledge and ability in the subject into account but the assessment should be made in relation to the goals stated in the course syllabus. The Swedish Education Agency formulates the criterions for different grade steps - which goals that has to be attained for a specific grade. In the national directive about final grades, it is clearly stated that diligence, feeling for order and attention in the class room should not be rewarded in the final grades.

In Sweden we have national tests in the core subjects Swedish, English and Mathematics. The aim of the tests is to help teachers to assess students according to the national stated goals and to enable fair and equivalent grades across the country. Teachers in Swedish, English and Mathematics *shall* use the result of the national (compulsory) examinations as assistance in determining the student's final grade. The tests are comprehensive; both oral and written abilities are tested and they are constructed by academic specialists from different universities in Sweden. The tests are enclosed with rigorous correcting instructions but it is the teachers in the school who accomplish the tests and correct them. Thus, they are not blind tests, even though teachers are encouraged to correct tests of their colleagues' students. Both the final grade and test result are graded according to the same metrics: fail, pass, pass with distinction and pass with special distinction. In addition to test results and final grades, the data used include both the gender and the ethnicity of the student as well as information about the gender and ethnical background composition of the teacher staff at the students' school.

Access to both test results and final grades of each student in each subject (Swedish, English and Mathematics) makes it possible to use a difference in difference strategy and it is possible to distinguish the same sex (ethnical background) effect from the gender (ethnical background) bias. The former is assumed to affect the final grade and the test result equally and it is therefore removed in a difference in difference strategy. The effect of having a teacher of the same sex (ethnical background) is tested for separately by estimating a model with only test result as dependent variable and an interaction term of the gender (ethnicity) of the teacher and the student among the explanatory variables.

The data used stems from the Swedish Agency for Education' register about students in grade 9 and a similar register about teachers. The data covers six years, from 1999 to 2005. The sample sizes are large; 120 000 in Swedish and English and almost 280 000 in Mathematics. With respect to foreign-born students the corresponding sample sizes are 20 000 and 82 000. The subjects studied in this study are those in which the Swedish school system provides national standard tests: Swedish (in which Swedish as second language is included), Mathematics and English.

The results suggest the following. On average, both girls and foreign-born are more generously rewarded in final grades, compared to test grades, than boys and non foreign-born students respectively[1]. This approach cannot exclude that part of this effect stems from student behaviour that might differ between genders and students with different ethnic background, respectively. However, the effect varies somewhat depending on the composition of the teacher staff, suggesting that part of the effect identified is a compensation effect or positive discrimination. In Mathematics, it is found that both girls and foreign born students perform better on the test when the share of female and foreign-born teachers, respectively, increases at the school.

The paper is organized in the following way. Section 2 discusses the raised hypothesises and related literature. Section 3 explains the Swedish school system and section 4 presents data. Section 5 presents some descriptive statistics and section 6 discusses the identification strategy. The results are presented in section 7 and section 8 discusses the results and concludes the paper.

2 Background and related literature

The gender gaps in favour of girls in school performance have got attention in the Economics literature. The majority of teachers in the middle-school, especially in readings, are female. Holmlund and Sund (2005) raise the feminization of the teacher profession as an explanation for the gender gap; girls may perform better when the teacher is female. By using Swedish data, they did not find any clear supportive evidence for their hypothesis. A few other studies, using data from other countries, present evidence of teachers serving as role models (Dee, 2005; Bettinger and Long, 2004 and Ehrenberg *et al*, 1994). In the same spirit, Dee (2001)

[1] Except for non Nordic-born students in English.

tests on US data if the same ethnical background of the teacher and the student matter for the student's performance. He finds supportive evidence. To my knowledge, no study has tested this hypothesis on Swedish data.

International comparisons of test results, concludes that there are gender differences in school performance all over the world. In general, girls are significantly better in readings (languages) while boys (on average) perform better in Mathematics (NAEP, 1999 and PISA 2003). In problem solving performance – a cross disciplinary competence – there is no significant difference between the genders (PISA study 2003). In Sweden, girls outperform boys significantly in readings and problem solving but not in Science and Mathematics, according to PISA 2003 (in Mathematics boys significantly outperform girls).

Gender comparisons in finale grades, which are set by the teacher, show that girls outperform boys in almost *all* subjects in Sweden. Evidence from UK shows a similar pattern (EOC, 2006; Skolverket, 2005). Thus, the gender gap in final grades is not reflected in test results. In Education, the gender gap between the difference in test results and teacher assessments have been investigates by several papers. Studies from different countries show that in general teachers assess girls higher than boys, irrespective of written exams (Frazer et al., 1975; Arnold 1968; Hess et al 1969; Peck 1071, Zazzo 1982). Emanuelsson and Fischbein (1986) have reviewed the gender gap in both final grades and test grades in Sweden between 1968 and 1983. They conclude that girls are on average better rewarded than boys in final grades irrespective of their relative performance on corresponding tests. This is true in a historical perspective (from 1968- 1983) as well as over the students' age. The test results used in Emanuelsson and Fischbein (1986) are not as comprehensive as the test results used in this paper. The most important difference is that these tests do not contain any oral part, which might disfavour girls.

Discrimination against girls and women as well as against minorities such as foreign-born is a hot topic both in the public debate and in academic research. The school is an important arena in this context. Several studies using Swedish data show that boys are more active and dominating in the class room and that teachers devote more time to boys than to girls (Einarsson, 1981; 1983; Bruzelius et al. 1981; Hultman, 1981). Ferguson (1998) and Castee (1998) shows that in US racial pairings of teachers and students influence how teachers

allocate their time in the classroom. Efforts to recruit teachers with a foreign background are often motivated by the claim that those teachers benefit students with foreign background. In contrast to the gender gap in school performance, foreign born students in Sweden perform below Swedish-born in school, both with respect to tests and final grades (Swedish Agency of Education, 2005). Thus, the common feature between girls and foreign-born is that there might be a concern by teachers about not discriminating these groups.

The idea that girls are positively discriminated has empirical support in a study by Lavy (2004). He tests for the existence of gender stereotype beliefs and discrimination on Israeli high school students by using a blind score and a non blind score test. The blind score tests and the non blind score tests are available in nine subjects – four in humanities, one in mathematics and four in science. In contrast to his preconception, he finds that the bias is against male students. The extent of the bias varies by subjects and test scores and ranges from 5 to 25 percent of the standard deviation of the blind score distribution. On average, girls outperform boys in all subjects, except in English in the blind score, in both the blind and the non blind test. Thus the bias against male student, on average, increase the gender score gap.

3 The Swedish School system

National objectives and guidelines for public education are laid down by the Swedish Parliament and the Government. The National Agency for Education shall evaluate, follow up and supervise the public schools. The Agency should also draw up and take decisions on grade criteria and general recommendations.

The teacher is alone responsible for assessing the grade. However, they are obliged to follow national directives and the Swedish Education Agency formulate the criterions for different grade steps; which goals that has to be attained for a specific grade. In the curriculum for the Compulsory School System (Lpo, 94) it is stated that grades should reflect skills and knowledge in the subject in accordance to the goals stated in the course syllabi. This means that grades should not reflect attention in the classroom, diligence, ambition, home works and work during lesson as long as it is not a prerequisite for attaining the goals (as in the case of laboratory work) (“Bedömning och betygsättning”).

To enable equivalent and fair grades across the country, national tests are performed during the spring semester in grade 9[2]. The tests in languages (Swedish and English) measure writing and reading abilities as well as listening comprehension and verbal interaction. The tests in Mathematics include problem solution, calculation ability as well as an oral part testing Mathematical reasoning. The tests are corrected at the school level and are graded according to nationally stated correcting instructions. Teachers are encouraged to not correct their own students' exams but they are allowed to do so. The tests are enclosed with rigorous instructions to the teachers how to correct and grade them in accordance with national standards. The interpretation of the correcting directives might differ between teachers. PRIM-gruppen 2002-2003[3] (the educationalists constructing the test in mathematics) investigated the reliability of the correcting instructions in Mathematics and the result shows that when five different teachers assess the same solution suggestion, the difference is on average 4 credits out of 60.

Teachers *shall* use the nationally approved examinations to assist in determining the grade to be given in grade 9 ("Likvärdig bedömning och betygsättning"). However, in the individual case, the teacher is free to set the final grade differently from the test grade. The reason is (of course) that the student might be low performing on the test day due to occasional conditions. Further, the teacher should take into consideration all available information about the student's knowledge and ability in the subject. The tests can not guarantee to capture *all* goals stated in the course syllabi, although the aim of the tests is to measure, as comprehensive as possible, the student's ability and knowledge in the subject. However, it is clear from the national directives that the tests should make an important basis for the teacher's assessment of their student's knowledge and skills in the subject. The tests are intended to serve as assistance for equivalent and fair grades.

The Swedish grade system has the following ordinal scale: Pass (P), Pass with Distinction (PD) and Pass with Special Distinction (PSD). A student can also fail to achieve the goals and is then reported as Failed (F). The aim of the Swedish school is that all student should attain at least "Pass" in the main subjects Swedish, Swedish as a Second Language, Mathematics and English. The grade "Pass" is required for being qualified to continue to upper secondary

[2] Swedish is tested in February, English in Mars and Mathematics in May.
[3] www.skolverket.se/sb/d/502/a/1977

school. In year 2005, 89.2 percent of all students in grade 9 were qualified to upper secondary school level.

4 Data

The data used stems from the Swedish Agency for Education's register about students in grade 9 (åk 9 registret) and a similar register about teachers (lärarregistret). Data on test scores is available in English and Swedish for the years 1999-2001 and for year 2005. In Mathematics, test scores are available between years 1999-2005, except for 2001[4].

For students with another language than Swedish as mother tongue and who are assessed to not be able to follow the ordinary course in Swedish, a special course is offered: Swedish as second language. Swedish as second language has its own course syllabus and students taking this course are graded according to the goals stated in this course syllabus. The grades in Swedish as second language are therefore not comparable with grades in the normal course "Swedish". In this paper Swedish as second language and Swedish are treated as one subject. The focus is on the difference between the final grade and the test grade and it should therefore not matter which course the student has taken[5].

The information about teachers is available at the school level and is matched with the student information via a school code. Thus, we can only draw conclusions about the effect of an increase in the percentage unit of a certain kind of teachers at the school level on students' performance. However, from a policy perspective this variable is relevant; it is easier to target teacher resources to schools rather than to individuals within schools. In addition, with information at the school level, we do not have to bother about selection of teachers within schools. The information about teachers includes teachers' formal training, how much he or she works (full time or part time), gender and age.

Unfortunately, the information about teachers' formal training is vague. The information is given by a code indicating if the teacher is trained within a group of subjects. The subject groups relevant in this study are: 1) Mathematics and Science (used for teachers in

[4] For the years 1999 to 2002, test scores from 150 randomly chosen schools are available. From 2003 until 2005 all schools reported test results. However, the information is incomplete in some years for certain subjects.
[5] The definition of foreign-born in this paper is not Nordic-born. However, it could be the case that a student has another mother tongue than Swedish but still is born in Sweden. A student can therefore be Swedish-born but taking Swedish as second language.

Mathematics), 2) Social Sciences and Swedish (used for teachers in Swedish) and 3) Swedish and languages (used for teachers in English).

Finally, I also have register data from Statistics Sweden about family background variables, gender and country of birth. The interest in this paper is foreign-born students and teachers. In Sweden many inhabitants are born in the other Nordic countries. Those people speak Swedish well, look Swedish and know the country well. Foreign-born is therefore defined as those born in a non Nordic country. In the following we use foreign-born as those born in a non Nordic country.

The sample sizes we end up with when information about teachers as well as students are matched are with respect to gender around 120, 000 in Swedish and English and around 280 000 in Mathematics. Unfortunately, country of birth is only available for the years 1999 to 2003. With respect to foreign-born students, the sample size for Swedish and English are around 20, 000 per subject and for Mathematics around 82, 000.

5 Descriptive statistics

Table 1 presents summary statistics on final grades and test grades for all, female, male and foreign-born students. In Swedish and English girls perform significantly better than boys with respect to both final grades and test result. In Swedish, the difference is largest; almost two grade points in test result difference. In English, the gender difference according to the test is small, although statistical significant: less than one half grade point. In Mathematics there is no statistical significant difference between test results while the difference in final grades is statistical significant and in favour of girls. Throughout, the gender gap is larger with respect to final grades than test results. With respect to foreign-born students and others, Nordic born students outperform foreign-born significantly in all subjects both with respect to final grades and test results.

Figure 1 to 6 present histograms for final grades and test grades in all three subjects. Female and male students as well as Foreign-born and others are presented separately. The distribution of final grades and test grades and the deviations between the two grades differ between the subjects. In Mathematics, a large share of all students score fail on the test but a much smaller fraction receive “F” as final grade (see Figure 3 and 6). In the other subjects,

failing is not that common, neither on the test nor as final grade. In English (see Figure 5), the difference between the measures is largest for the grades “P” and “PD”. In Swedish, the largest difference instead is with respect to “PSD”. It is important to note that the subject specific distribution is similar in final grades and in test grades. This is an indicator that the final grade and the test grade measure about the same outcome, even though final grades in general are more generous.

The final grades are set after the test grade. Thus, the teacher can only deviate from “F” on the test by raising the final grade while a deviation from “PSD” on the test implies a lower final grade than test grade. A larger share of the boys compared of the girls fail on the test in Swedish and English while the opposite is true with respect to the highest test grade. However, still a larger share of the girls receives a higher final grade than test grade compared to boys in English and Swedish. In Mathematics there is no gender difference with respect to test results. Foreign-born students score fail on the test in a higher degree in all subjects compared to others. This group also receives higher final grades than test grades in all subjects.

Table 2 presents the share of female and foreign-born teachers for the relevant subject categories. In Languages around 84 percent of the teachers are female. In Social science the corresponding share is 55 percent while in Mathematics and Science, the share is around 45 percent. Foreign-born teachers constitute 6 percent of the colleagues in Science and Maths and around 9 percent in Languages and Swedish. In Social Sciences and Swedish only 3 percent of the teachers are foreign-born.

Table 3 shows the difference between average test grades and final grades for all students, only females, only males and only foreign-born students. With respect to all students, final grades are, on average, significantly higher than test grades in all three subjects. In Swedish and Mathematics this is also true for all three sub groups. In English, the pattern changes when only considering boys. For boys, average test grades are significantly higher than average final grades. Thus, in English boys on average receives a lower final grade as compared with the average test score.

Table 4 presents the distribution of the absolute grade difference between final-grade minus test grade. The large majority of student receives the same final grade as test grade; more than

70 percent in all subjects. Looking at the percentage share of the students who receive one step lower finale grade than test grade, this share is in all subjects larger for boys than for girls. In English and Mathematics it is almost twice as large. With respect to foreign-born students, the share receiving one step lower final grade than test grade is in all subjects lower in comparison with Nordic-born students. In the next section these differences are estimated more precisely.

6 Identification Issues

In this section I discuss identification issues with respect to the gender bias effect. The same reasoning is also applicable for the ethnical background case but is left out for the ease of exposition.

The ideal design for testing for gender bias is to compare two groups where the only difference is that the control group has been assessed anonymously while the treated group has been assessed by a teacher knowing the gender of the student. The set up used in this paper differ from this case in two important ways. Firstly, the tests are not blind; in most cases the teacher knows the student's identity while assessing him or her. Secondly, the two assessments (final grade and test grade) are not based on the same information set. Final grades are a subjective assessment by the teacher based on all information about the student the teacher has access to. Test results, on the other hand, are a measure of the student's knowledge at one single occasion. In the following we discuss these two deviations. We start by the first.

The teacher may know the student's identity while correcting the test. Especially in the verbal part it is obvious that the student's gender is revealed. This deviation from the ideal case may bias the result. If test grades are affected by the gender and ethnicity, this approach would capture the lower bound.

Are final grades and test results comparable? According to the national directives, final grades should reflect knowledge and ability in the subject, only. Home works and good behaviour should not be rewarded. At the same time, teachers shall take all available information about the student into account and home works and good behaviour is a mean for the students to prove their competence in the subject. What teachers *should* reward in final grades is behind

the scope of this paper. The interest in this context is *how* they in general interpret the national directives about final grades. The fact that final grades overall are more generous than test grades suggest that students can compensate a bad result by proving their competence in other ways. More boys than girls fail on the test and more girls than boys score the highest grade. Thus, if additional information about the student's performance is included in the final grade, it seems easier for boys on average to improve through good behaviour; the teacher can only deviate from "F" on the test by raising the final grade while a deviation from "PSD" implies a lower final grade than test grade. However, if girls (on average) do their home works better than boys they have more occasions to prove their knowledge and ability in the subject to the teacher. Thus, if girls behave better than boys in this respect, it could explain why girls are better rewarded in final grades than boys. We cannot control for student behaviour. Thus we cannot exclude that part of the effect identified stems from gender differences in student behaviour. However, we can investigate if the generosity in final grades varies depending on teacher characteristics. The hypothesis is that girls are more generously rewarded in final grades as a compensation for not giving them as much attention as boys. It is also known that the same gender of the teacher and the student benefit the student. Several explanations have been raised in the literature. One possible explanation is that the gender of the teacher affects how the teacher allocates her time among girls and boys. Thus, if the hypothesis about more generous rewarding of girls than boys is true, it is reasonable to believe that male teachers reward girls more than boys in final grades. In order to check this, a model that interact female students with the share of female teachers at the school is also estimated.

With respect to foreign-born students and others it is also possible that there are differences in student behaviour that might explain the result but I have not found any theoretical support for such differences. Foreign-born students are included in this analysis because they - as well as girls - might be regarded as discriminated in the class room. However, related to teacher characteristics, it is reasonable to believe that non foreign-born teachers are less generous with respect to non-foreign-born students because foreign-born teachers benefit foreign-born students.

Finally, in order to investigate if the same sex and ethnical background of the teacher and the student matter in this data set, this effect is estimated separately with test result as dependent variable.

6.1 Econometric Model

Very few students fail in a subject in the sense that they do not attain "P" as final grade. Few also receive (and also should according to the directives) the grade "PSD". Thus, the step from F to P is larger than the step from "PD" to "PSD", thus the grade system is ordinal and hence ordered response methods are appropriate. However, the estimated coefficients from ordered logit or probit models give only directions of effects. Magnitudes of the coefficients are possible to get by calculating partial effects at various values of the explanatory variables, but the information is difficult to summarize. Linear models are easy to interpret and are good estimates of the average effect (Wooldridge,). I will therefore focus on linear regressions. Estimation results of ordered probit models (with marginal effects) give the same qualitative results and are presented upon request.

A crucial challenge when using a linear model with an ordinal outcome variable is to transform the ordinal grade to an appropriate numeric scale. The Swedish National Agency for Education summarizes all final course grades into a total Grade Point Average (GPA). This GPA is the most important selecting instrument for higher education. The values used for getting the GPA are 0, 10, 15 and 20. These values are also used in this paper for both final grade and test grade. The distributions of both grade variable are somewhat skewed in the sense that higher grades are more frequent than lower.

6.1.1 Estimating the gender bias (ethnical background) effect

Two outcomes variables per students make it possible to use a difference in difference estimation approach. The following model is estimated for each subject separately:

$$G_{ijt} = \beta_i + \beta_1 fem_{ijt} + \beta_2 TG_{ijt} + \beta_3 Interaction_{ijt} + \eta_j + \varepsilon_{ijt} \quad (1)$$

where G_{ijt} is individual *i*'s grade in year *t* for type of grade *j*. *fem* is a dummy for female student and *TG* is a dummy for the type of grade. *TG* equals 1 (0) if the grade corresponds to final grade (test grade). The interaction term consists of the gender and the grade type dummy

and finally the last term (η_t) captures year effects[6]. By using a difference in difference strategy, individual ability in the subject as well as school fixed effects is removed as long as they have the same effect on final grades and test grades. As Lavy (2004) – who estimate the existence of discrimination by Israeli high school teachers – points out, the advantage of estimating model (1) is that it reveals the estimates of the other parameters.

Model *(1)*, is a basic model. In order to check the model specification, two extended models are also estimated. First, school dummies are added. The reason is that the policy to set final grades may differ between schools. This argument is most relevant with respect to foreign-born students. It might be the case that schools having many foreign-born students also are schools in which the policy to set final grades differ from other schools. The final grade policy is associated with the social status of the school which in turn is associated with many foreign-born among the students. If this is the case, the social status is driving the result that foreign-born students are discriminated, not foreign-born students[7].

Secondly, we add the following explanatory variables in order to increase the precision of my estimates: the student's month of birth, the teachers' mean age and, in the regression about ethnicity, the share of female teachers at the school. Earlier literature (Fredriksson and Öckert, 2004) has shown that children born late in the year perform poorer than their peers born in the beginning of the year. It is possible that absolute age influence the student's behaviour relative to her peers which in turn affect the teacher's subjective evaluation of the student. In this paper month of birth also captures those who are one or two years older (younger) than the majority[8].

6.1.2 The gender effect and teacher characteristics

In this section we test if the gender bias effect differs depending on teacher characteristics. I estimate the model 3 above but I use the difference between final grade minus test grade as

[6] The data used is of a cross section type. However, the sample covers several years and some schools are observed several times. A year specific dummy is therefore also included to capture wich year the individual and the school is observed. Including year dummies imply that we do not have to bother about changes in grade policy (grade inflation) over time. Wikstrom and Wikstrom (2004) claim that grade inflation occurs in Sweden during the 1990s.

[7] If low status schools are associated with a non generous grade policy and many foreign born students, we overestimates the effect (in absolute meaning) of being foreign born in a model specification without school dummies.

[8] January the year the majority (around 97 percent) is born is the reference. Born one year earlier in January gives -12 and born in January one year later gives +12 etc.

dependent variable (TG – dummy and the interaction term concealed). The share of female (foreign-born) teachers at the school is added as well as an interaction of the dummy for girl and the share of female teachers. The information about teachers is at the school level. Thus, the variation used stems from a variation across years. In Mathematics we observe, the same schools across 6 years. In English and Swedish, we observe the same schools across 4 years. When we use school fixed effects, the variation at the school level stems from 342 schools in English, 311 in Swedish and 2541 schools in Mathematics, with respect to gender. With respect to ethnic background, the variation is: 32 schools in Swedish, 34 in English and 918 in Mathematics. Thus, with respect to ethnical background, the sample sizes in English and Swedish are extremely small.

6.1.4 A same sex (ethical background) effect

A same sex teacher effect - that girls perform better in the subject with a female teacher – is tested for by estimating a model with only test grade as dependent variable, otherwise exactly as above. This effect is identified under the assumption that the assessment of the test is not influenced by the student's gender and ethnical background. The variable of interest is the interaction term which capture if female students perform better with a larger share of female teachers in the subject. The same sex teacher effect is identified under the assumption that the assessment of the test is not influenced by the student's gender.

7 Results

The results are presented in tables 5-7.

7.1 Interaction of Gender (Foreign-born) and Type of Grade (Table 5)

The interaction of gender and type of test is highly statistical significant in all subjects. Thus, girls are better rewarded than boys in terms of final grades when controlling for gender and type of test. The effect in Mathematics (0.364) corresponds to about 76 percent of a standard deviation of the grade difference distribution. Remember that the highest grade, "PSD", is given 20 points. The steps from "P" to "PS" and from "PS" to "PSD" correspond to 5 credit points each. Roughly speaking, around one girl in each class (of 30 students) gets one step higher grade than she deserves according to the test result[9].

[9] In a class with 30 students and 15 girls, the average grade difference is 0,364 higher for girls compared to boys. That means that girls' total grade credit is 5.565 (15*0,364) higher than boys'. That corresponds to the situation that one female student gets one step higher grade (between "P" and "PS" or "PS" and "PSD" or half the step between "F" and "P") due to gender.

This model specification also reveals how gender and type of grade (TG) is correlated with grade (the latter can be both final grade and test grade). In English, the coefficient in front of the type of grade dummy is negative when we control for girl and the interaction between girl and type of grade. Thus, as shown in the descriptive part, boys get on average a lower final grade than test grade in English. In Mathematics, the gender difference disappears when controlling for type of test and the interaction term; reflecting the no difference between genders in test result. In Swedish the gender gap is largest (compared to the other subjects) with respect to type of grade but the interaction term is smallest (compared to the other subjects). Thus, the gender bias effect is larger in the subjects in which the gender gap in the test results is smaller.

Foreign-born students on average significantly underperform compared to others in all subjects. The TG dummy is significant and positive in all subjects and in Swedish and Mathematics it is about the same size and sign as the corresponding estimates in the gender model. The interaction term, which is of main interest, is positive and significant in Swedish and Mathematics but not in English. Thus, except for in English, both girls and foreign-born students are better rewarded in final grades than boys and non foreign-born, respectively. The effect sizes are about the doubled size with respect to ethnical background compared to gender.

The results seem robust. Adding school dummies (column 2) and also covariates (column 3) change neither coefficients nor precision significantly in any estimated model.

7.2 The gender effect and teacher characteristics (Table 6)

The interaction term between the share of teachers and the student with respect to gender, is only significant in Mathematics. The reason is probably the sample size (which is more than twice as big in Mathematics compared to the other subjects). The effect is negative and has the following interpretation: a percentage unit increase of the share of female teachers in Mathematics and Science, affect female students' final grade negatively with around 0.134 credit points (test grade unchanged). The result suggests that, in Mathematics, male teachers are more generous than female teacher to girls in final grades. The result holds when school dummies and additional covariates are included. The robust result is in line with the difficulty to claim that school dummies should be included in the gender case. The effect size of the interaction term (in Mathematics) corresponds to around 4 percent of a standard deviation of

the grade difference distribution. The interpretation is the following. If the share of female teachers in Mathematics (and Science) increases with *10 percentage units* at the school, on average girls would get 13.5 credit points higher final grade[10]. This result is significant; it corresponds to (more than) a grade change from "Fail" to "Pass" or from "Pass" to "Pass with Special Distinction".

Interaction between foreign-born teachers and students shows another pattern. In English the effect becomes larger and more precise when school dummies and covariates are added. Foreign-born teachers seem to self select to schools with a deviating final grade policy. Thus, controlling for the school is important when estimating the effect of foreign-born teachers. The coefficient in front of the interaction term in English ends up positive and significant at the 5 percent level. The effect of the interaction term in English (0.134) corresponds to 30 percent of a standard deviation in the grade difference distribution. However, it is important to remember that this result is estimated from a variation of 34 schools, only. The effect size is about the same size as with respect to gender. However, it is important to have in mind the small variation in schools.

The differences in estimates depending on teacher staff characteristics indicate that part of the gender (ethnical background) bias effect stems from teacher behaviour rather than the possibility that other qualities – good behaviour - are rewarded in final grades. However, the effects from the interaction of teacher and student are smaller. Thus, this evidence does not exclude that final grades reflect good behaviour which might be correlated with girls or foreign-born students.

7.3 A same sex or ethnicity effect (Table 7)

A same sex or ethnical background teacher effect is statistical significant in Mathematics. In Mathematics, girls seem to perform better with a female teacher and foreign-born students perform better with a foreign-born teacher (significant at 1 and 5 percent level respectively). The sample size is largest in Mathematics and this could be the reason why an effect only is captured in this subject. The effect of 0.233 credit points with respect to gender corresponds to 4.5 percent of the standard deviation from the test grade distribution in Mathematics. The effect with respect to ethnical background is 0.973, which is 20 percent of the standard

[10] Again assuming a class with 30 students of which 15 are female.

deviation of the test grade distribution. If we extrapolate the result with respect to gender the estimated effect has the following interpretation. A *10 percentage units* increase in the share of female teachers at the school positively effect girls on average with more than 2 credit points. That corresponds to a half grade step between "Pass" and "Pass with Distinction". The interpretation of the corresponding effect for foreign-born students and teachers is almost 10 credit points – an increase from "Fail" to "Pass".

8 Discussion of results

The conclusion from the results is that girls and foreign-born students on average get a higher final grade than boys and others, respectively, that perform equally well on the test. The effect identified could stem from better behaviour of girls and foreign-born students. However, the identified effect varies depending on the gender and ethnical background composition of the teacher staff, which suggests that not all of the effect could be explained by the fact the final grades also reflect god behaviour. In Mathematics, girls and foreign-born students perform better with a same sex and ethnical background teacher, respectively. Thus, same sex and ethnical background of the teacher matter for the student's performance.

In the next section an additional robustness control with respect to the gender and ethnical background bias effect is presented. Finally, a summary discussion of the results concludes the paper.

8.1 Statistical Discrimination with respect to gender (ethnical background) bias

On average, girls perform better than boys according to both test grades and final grades. Lavy (2004) discusses the phenomenon of statistical discrimination motivated by the average superior performance of girls. If teachers are influenced by the higher performance of girls in general, boys and girls will receive different final grades even if they perform at the same level. This might be the case in English and in Swedish especially, subjects in which girls on average perform significantly better than boys. The reasoning about statistical discrimination is also applicable in the analysis about ethnical background.

I follow Lavy (2004) and estimate the gender effect in final grades for a subgroup; schools in which boys on average outperform girls on the national tests, only. This sub sample consists of around 3 percent of the individuals in Swedish, 34 percent in English and 50 percent in

Mathematics. I use the extended model (model 3)with results presented in table 5. With respect to foreign-born students the direction of possible statistical discrimination is not evident. According to the tests, foreign-born students on average perform below the average of Nordic-born. Thus, I estimate the corresponding model for schools in which foreign-born students on average outperform others. This sub sample consists of 31 percent of the individuals in Swedish, 41 percent in English and 20 percent in Mathematics. The results are presented in Table 8.

The effect of being a female student remains highly significant and the magnitude of the coefficient increases in all subjects. Thus, the *fact* that girls on average perform better seems not *per see* explain the result. However, the *belief* they do can still exist. Such belief may arise if, for example, teachers believe that girls perform worse under pressure, as is the case in a test situation. If this is true or not is irrelevant for that type of statistical discrimination.

With respect to country of birth the pattern changes. In all subjects, the coefficients in front of the dummy change sign (becomes negative and non significant). Thus, when, foreign-born actually outperform others, they are not more generously rewarded. This dramatic change suggests that when teachers know that this group in general performs above the average, they are not more generous in final grades to this group compared to others.

8.2 Final Conclusion

To be written

References

Bernard M. E. (1979) “Does Sex Role Behaviour Influence the Way Teachers Evaluate Students?” *Journal of Educational Psychology, 71*pp. 553-562

Bettinger E. P. and Long B. T. “Do Faculty Serve as Role Models? (2005) “The Impact of Instructor Gender on Female Students” *The American Economic Review, forthcoming*

Bruzelius *et al.* 1981 Report 4:1981 from the Department of Subject Methodology and theory (Swedish) at the university of Lund. Language and Sex Project.

Casteel C. A (1998), “Teacher-Student Interactions and Race in Intergrated Classrooms”, *Journal of Educational Research 92:* 115-120

Dee T. (2005), “Teachers and The Gender Gaps in Student Acievement”, National Bureau of Economic Research, Working Paper 11660

Dee T. (2001), “Teachers, Race and Student Achievement in a Random Experiment”, National Bureau of Economic Research, Working Paper 8432

Dusek J. B. and G. Joseph (1983), “The Bases of Teacher Expectancies: a Meta-Analysis “ Journal of Educational Psychology, 75, 327-346

Ehrenberg, Ronald G and Goldhaber D and Brewer D (1995) “Do Teachers' Race, Gender, and Ethnicity Matter?” Evidence from NELS88, *Industrialand Labor Relations Review 48(3): 547-561*

Emanuelsson and Fischbein (1986) “Vive la Difference? A study on Sex and Scoolong”, Scandinavian Journal of Educational Research 30, 71-84

Einarsson, 1981; Report 5:1981 from the Department of Subject Methodology and theory (Swedish) at the university of Lund. Language and Sex Project.

Equal Opportunities Commission (EOC) Research Findings (2006) “Gender and Differential Achievement in Education and Training: A Research Review” Research Findings

Fredriksson and Öckert (2004) “Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance”

Ferguson, R.F. (1998), “Teachers’ Perceptions and Expectations and the Black-White Test Score Gap”, in *The Black-White Test Score Gap* C. Jencks and M. Philips, editors, Brookings Institution Press Washington DC, 1998

Holmlund H. And Sund K (2005), “Is the Gender Gap in School Performance Affected by the Sex of the Teacher?”, Preliminary Draft September 6, 2005 SOFI, Stockholm university

Hultman, 1981 Report 6:1981 from the Department of Subject Methodology and theory (Swedish) at the university of Lund. Language and Sex Project.

Lavy V (2004), “Do Gender Stereotypes Reduce Girls’ Human Capital Outcomes?”, Evidence from a Natural Experiment, NBER Working Paper 10678

NAEP (2004) Trends in Academic Progress – Three Decades of Student Performance in Reading and Mathematics, US Department of Education, Institute of Education Sciences, NCES 2005-464

PISA 2003; “Problem Solving for Tomorrow’s World”, First Measures of Cross-Curricilar Competencies from PISA 2003, Programme for International Student Assessment, OECD

Stipek, D (2001), “At What Age Should Children Enter Kindergarten? A Question for Policy Makers and Parents”, *Social Policy Report*, Volume XVI, number 2

Swedish National Agency for Education (skolverket) (2004), Skolverkets allmänna råd 2004

“Bedömning och betygsättning” s.33, SKOLFS 2001:15

Tiedemann, J (2000) , “Parents’ Gender Stereotypes and Teachers’ Beliefs as Predictors of Children’ Concept of Their Mathematical Ability in Elementary School”, *Journal of Educational Psychology*, 92 (1), pp. 144-151

Wikström C and Wikström M (2004), ”Grade inflation and school competition: an empirical analysis based on the Swedish upper secondary schools”, *Economics of education Review* 24:309-322

Wooldridge J M (2002) *Econometric Analysis of Cross section and Panel Data*, the MIT Press

Table 1 Difference in Tests Grades and Final Grades between Gender and between Foreign-born and Others

Subject	*Observations,*	*Boys*	*Girls*	*t-test*	*Observations*	*Non Foreign-born*	*Foreign-born*	*t-test*
	All students	*Mean (St dev)*	*Mean (St dev)*		*All students*	*Mean (St dev)*	*Mean (St dev)*	
Swedish								
Test Grade	128963	11.284	13.3756	-89.6763	26021	12.42803	10.62525	20.9450
		(4.304476)	*(4.062806)*			*(4.010769)*	*(4.599726)*	
Final Grade	128963	11.9876	14.27642	-1.0e+02	26021	13.1913	11.86564	15.3544
		(4.035887)	*(4.093747)*			*(4.059796)*	*(4.296661)*	
English								
Test Grade	131418	13.02672	13.46605	-18.1146	26880	13.08359	11.90927	13.4728
		(4.483371)	*(4.302382)*			*(4.108968)*	*(4.776337)*	
Final Grade	131418	12.90395	13.77758	-36.1545	26880	13.38427	12.18146	13.6394
		(4.411773)	*(4.34528)*			*(4.160161)*	*(4.808345)*	
Mathematics								
Test Grade	286361	11.21682	11.2363	-1.0084	85451	11.4623	9.04576	41.2038
		(5.153402)	*(5.185077)*			*(4.774344)*	*(5.385708)*	
Final Grade	286361	12.30054	12.68393	-24.2751	85451	12.51259	10.97174	31.2493
		(4.223148)	*(4.226851)*			*(4.035727)*	*(4.32064)*	

Table 2 Share of female and Foreign- born teachers in different subjects (sample basic)

	Share of female			*Share of foreign-born*		
At the school level:	*Obs*	*Mean*	*St Dev*	*Obs*	*Mean*	*St Dev*
Social Science and Swedish	121562	.549265	.3097371	19947	.0248276	.1084718
Language and Swedish	124231	.837848	.4488278	21013	.0897058	.2044835
Science and Mathematics	282422	.4488278	.2424945	82811	.0600209	.1337706

Foreign born teachers are those not born in a Nordic country.

Table 3 Test Grade versus Final Grades

Subject and group of students	*Observations*	*Test Grade*	*Final Grade*	*t-test for difference in means*
Swedish				
All students	128963	12.31248	13.11306	47.6157
		(.0120183)	*(.0117579)*	
Females	63414	13.3756	14.27642	39.3313
		(.0161337)	*(.0162566)*	
Males	65549	11.284	11.9876	30.5289
		(.0168127)	*(.0157636)*	
Foreign-born	2456	2.234935	2.427932	9.5613
		(.0139148)	*(.0146227)*	
English				
All students	131418	13.24263	13.3333	5.2810
		(.0121397)	*(.01214)*	
Females	64585	13.46605	13.77758	12.9471
		(.0169295)	*(.0170983)*	
Males	66833	13.02672	12.90395	-5.0458
		(.0173424)	*(.0170654)*	
Foreign-born	2522	2.4659	2.508327	1.9104
		(.0154087)	*(.0159935)*	
Mathematics				
All students	286361	11.22637	12.48848	101.1261
		(.0096593)	*(.0079034)*	
Females	140378	11.2363	12.68393	81.0781
		(.013839)	*(.0112815)*	
Males	145983	11.21682	12.30054	62.1461
		(.0134879)	*(.0110531)*	
Foreign-born	7452	2.028717	2.270129	20.6141
		(.0085129)	*(.0080423)*	

Table 4a Absolute Grade Difference, Percentage Distribution, Gender

Difference	*Swedish*			*English*			*Mathematics*		
	All	*Females*	*Males*	*All*	*Females*	*Males*	*All*	*Females*	*Males*
-3	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00
-2	0.04	0.05	0.04	0.09	0.07	0.11	0.06	0.06	0.06
-1	4.71	4.13	5.28	8.05	5.84	10.19	3.44	2.64	4.20
0	77.10	75.18	78.96	82.63	82.68	82.57	75.89	74.23	77.49
1	17.88	20.34	15.50	9.15	11.29	7.07	20.38	22.78	18.07
2	0.25	0.29	0.21	0.07	0.09	0.05	0.22	0.28	0.17
3	0.01	0.01	0.00	0.00	0.01	0.00	0.01	0.01	0.01

Table 4b Absolute Grade Difference, Percentage Distribution, Foreign-born versus Others

Difference	*Swedish*			*English*			*Mathematics*		
	All	*Foreign*	*Others*	*All*	*Foreign*	*Others*	*All*	*Foreign*	*Others*
-3	0.01	0.00	0.00	0.01	0.08	0.01	0.00	0.00	0.00
-2	0.04	0.04	0.03	0.09	0.08	0.06	0.06	0.09	0.05
-1	4.71	3.68	4.03	8.05	6.44	6.27	3.44	2.51	4.18
0	77.10	74.09	78.36	82.63	82.25	82.38	75.89	71.02	76.91
1	17.88	21.23	17.39	9.15	11.04	11.17	20.38	25.93	18.65
2	0.25	0.88	0.18	0.07	0.12	0.10	0.22	0.45	0.21
3	0.01	0.08	0.00	0.00	0.00	0.01	0.01	0.00	0.01

Table 5a Interaction of Gender and Type of Grade (TG)
Parameter estimates from OLS estimation, three types of models: *1) Basic, 2) Add school dummies, 3) Add Covariates*

	Swedish			***English***			***Mathematics***		
	1)	*2)*	*3)*	*1)*	*2)*	*3)*	*1)*	*2)*	*3)*
Female	2.092	2.095	2.075	0.440	0.427	0.406	0.019	0.005	-0.028
	(0.029)***	(0.028)***	(0.028)***	(0.030)***	(0.029)***	(0.029)***	(0.022)	(0.022)	(0.021)
TG	0.704	0.704	0.704	-0.123	-0.123	-0.123	1.084	1.084	1.084
	(0.019)***	(0.019)***	(0.019)***	(0.017)***	(0.017)***	(0.017)***	(0.019)***	(0.019)***	(0.019)***
Interaction	0.197	0.197	0.197	0.434	0.434	0.434	0.364	0.364	0.364
	(0.019)***	(0.019)***	(0.019)***	(0.015)***	(0.015)***	(0.015)***	(0.014)***	(0.014)***	(0.014)***
Birth Month			-0.111			-0.111			-0.109
			(0.003)***			(0.003)***			(0.002)***
Mean Age			0.004			0.031			0.004
			(0.018)			(0.019)			(0.012)
Share Female			0.993			0.379			0.558
			(0.636)			(0.564)			(0.389)
Constant	7.237	7.267	7.109	8.176	7.863	6.878	7.341	7.061	7.162
	(1.733)***	(1.554)***	(1.810)***	(1.493)***	(1.263)***	(1.570)***	(1.252)***	(1.311)***	(1.430)***
Observations	257926	257926	257926	262836	262836	262836	572722	572722	572722
R-squared	0.07	0.14	0.16	0.01	0.08	0.09	0.02	0.08	0.09

Standard errors in parentheses are clustered on schools, all models include year dummies, * significant at 10%; ** significant at 5%; *** significant at 1%

Table 5b Interaction of Foreign-born student and type of Grade (TG)
Parameter estimates from OLS estimation, three types of models: *1) Basic, 2) Add School Dummies, 3) Add Covariates*

	Swedish			*English*			*Mathematics*		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Foreign-born	-1.365	-1.305	-1.156	-0.817	-0.725	-0.579	-2.396	-2.169	-1.999
	(0.164)***	(0.156)***	(0.160)***	(0.167)***	(0.155)***	(0.154)***	(0.196)***	(0.168)***	(0.168)***
TG	0.796	0.796	0.796	0.396	0.396	0.396	1.801	1.801	1.801
	(0.042)***	(0.042)***	(0.042)***	(0.039)***	(0.039)***	(0.039)***	(0.073)***	(0.073)***	(0.073)***
Interaction	0.488	0.488	0.488	-0.070	-0.070	-0.070	1.043	1.043	1.043
	(0.146)***	(0.146)***	(0.146)***	(0.085)	(0.085)	(0.085)	(0.150)***	(0.150)***	(0.150)***
Birthmonth			-0.084			-0.084			-0.098
			(0.007)***			(0.007)***			(0.008)***
Teacher Age			-0.190			-0.013			-0.045
			(0.107)*			(0.179)			(0.110)
Share Female			-3.980			2.148			-0.775
			(3.217)			(5.294)			(5.582)
Constant	10.102	9.114	20.306	8.802	7.642	7.225	7.599	6.565	9.374
	(1.097)***	(0.492)***	(4.988)***	(1.519)***	(1.444)***	(8.460)	(1.345)***	(1.596)***	(6.760)
Observations	27568	27568	27568	27568	27568	27568	27568	27568	27568
R-squared	0.02	0.09	0.10	0.10	0.10	0.10	0.12	0.12	0.13

Standard errors in parentheses are clustered on schools, all models include year dummies, * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6a Interaction of share of female teachers in the subject and girl
Parameter estimates from OLS estimation, three types of models: *1) Basic, 2) Add School Dummies, 3) Add Covariates*

	Swedish			*English*			*Math*		
	1	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
Girl	0.183	0.187	0.189	0.416	0.422	0.424	0.426	0.422	0.430
	(0.041)***	(0.041)***	(0.041)***	(0.063)***	(0.064)***	(0.064)***	(0.029)***	(0.029)***	(0.029)***
Share Female	-0.043	-0.115	-0.106	-0.192	-0.243	-0.206	-0.163	-0.146	-0.120
	(0.062)	(0.086)	(0.085)	(0.083)**	(0.123)**	(0.124)*	(0.069)**	(0.094)	(0.097)
Interaction	0.027	0.027	0.026	0.028	0.029	0.027	-0.137	-0.134	-0.135
	(0.064)	(0.064)	(0.064)	(0.075)	(0.075)	(0.075)	(0.058)**	(0.057)**	(0.057)**
Birth Month			0.005			-0.000			0.024
			(0.002)**			(0.002)			(0.001)***
Teacher Age			-0.024			-0.010			-0.024
			(0.016)			(0.016)			(0.012)**
At school			-0.246			-1.135			-0.535
			(0.498)			(0.466)**			(0.383)
Constant	-2.054	-2.043	-0.869	-7.308	-7.376	-6.234	-2.535	-2.674	-1.408
	(1.815)	(1.869)	(2.019)	(1.770)***	(1.902)***	(2.107)***	(2.764)	(2.741)	(2.811)
Observations	121562	121562	121562	124231	124231	124231	282422	282422	282422
R-squared	0.00	0.04	0.04	0.01	0.05	0.05	0.01	0.07	0.07
Schools	311	311	311	342	342	342	3087	3087	3087

Standard errors in parentheses are clustered on schools, all models include year dummies * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6b Interaction with share of foreign-born teachers at the school and foreign-born student
Parameter estimates from OLS estimation, three types of models: *1) Basic, 2) Add School Dummies, 3) Add Covariates*

	Swedish			*English*			*Math*		
	1	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
Foreign-born	0.058	0.059	0.061	-0.012	-0.015	-0.012	0.077	0.059	0.058
	(0.016)***	(0.016)***	(0.017)***	(0.013)	(0.012)	(0.013)	(0.009)***	(0.007)***	(0.007)***
Share Foreign	0.022	0.339	0.348	-0.032	-0.163	-0.196	0.045	0.133	0.118
	(0.045)	(0.076)***	(0.126)***	(0.038)	(0.043)***	(0.064)***	(0.031)	(0.051)***	(0.051)**
Interaction	0.042	-0.011	-0.012	0.075	0.133	0.134	-0.068	-0.008	-0.003
	(0.165)	(0.159)	(0.159)	(0.058)	(0.058)**	(0.059)**	(0.043)	(0.036)	(0.036)
Birth Month			-0.001			-0.001			0.001
			(0.001)			(0.001)**			(0.000)*
Mean Age			-0.001			0.001			-0.014
			(0.012)			(0.012)			(0.004)***
Share Female			-0.355			0.344			0.059
			(0.182)*			(0.278)			(0.113)
Constant	-0.212	-0.321	-0.059	-1.000	-1.029	-1.263	-0.346	-0.376	0.229
	(0.177)	(0.202)	(0.587)	(.)	(0.018)***	(0.548)**	(0.309)	(0.298)	(0.352)
Observations	20099	20099	20099	21161	21161	21161	82614	82614	82614
R-squared	0.00	0.05	0.05	0.00	0.05	0.05	0.02	0.09	0.09
Schools	32	32	32	34	34	34	918	918	918

Standard errors in parentheses are clustered on schools, all models include year dummies * significant at 10%; ** significant at 5%; *** significant at 1%

Table 7a Test of a Same Sex Teacher Effect by using Test Result as dependent variable
Parameter estimates from OLS estimation, covariates from model presented in Table 5

	Swedish	*English*	*Mathematics*
Female	2.127	0.405	-0.129
	(0.062)***	(0.124)***	(0.046)***
Female teachers, subject	0.115	-0.080	0.130
	(0.149)	(0.158)	(0.142)
Interaction	-0.062	-0.035	0.233
	(0.095)	(0.146)	(0.090)***
Month of birth	-0.113	-0.113	-0.121
	(0.003)***	(0.003)***	(0.002)***
Mean teacher age	0.019	0.037	0.016
	(0.028)	(0.023)	(0.017)
Share female teachers	0.778	0.829	0.722
	(0.906)	(0.717)	(0.534)
Constant	6.350	10.707	8.501
	(1.961)***	(1.453)***	(2.040)***
Observations	121562	124231	282422
R-squared	0.15	0.10	0.09
Schools	311	342	3087

Standard errors in parentheses are clustered on schools, all models include year dummies * significant at 10%; ** significant at 5%; *** significant at 1%

Table 7b Test of a Same Ethnic Background Teacher Effect by using Test Result as dependent variable
Parameter estimates from OLS estimation, covariates from model 3 presented in Table 5

	Swedish	*English*	*Mathematics*
Foreign-born	-1.295	-0.753	-1.425
	(0.153)***	(0.145)***	(0.083)***
Foreign teachers in subject	-3.618	1.261	-0.627
	(1.582)**	(0.365)***	(0.581)
Interaction	0.558	-0.215	0.973
	(0.695)	(0.496)	(0.441)**
Month of birth	-0.091	-0.097	-0.099
	(0.007)***	(0.006)***	(0.004)***
Mean teacher age	0.013	0.146	0.111
	(0.076)	(0.071)**	(0.039)***
Share female teachers	1.445	-5.450	-0.554
	(1.681)	(2.328)**	(1.063)
Constant	7.677	8.599	5.569
	(3.858)**	(3.789)**	(2.668)**
Observations	20099	21161	82614
R-squared	0.10	0.10	0.13
Schools	32	34	918

Standard errors in parentheses are clustered on schools, all models include year dummies * significant at 10%; ** significant at 5%; *** significant at 1%

Table 8a Statistical Discrimination, Model 3 from Table 5 but only schools in which boys outperform girls
Parameter estimates from OLS estimation, covariates from model 3 presented in Table

	(1)	*(2)*	*(3)*
	Swedish	English	Mathematics
Female	-0.646	-0.772	-0.987
	(0.079)***	(0.029)***	(0.020)***
TG	0.647	-0.215	0.944
	(0.089)***	(0.027)***	(0.024)***
Interaction	0.513	0.544	0.663
	(0.112)***	(0.026)***	(0.019)***
Birth Month	-0.108	-0.107	-0.111
	(0.015)***	(0.005)***	(0.003)***
Mean Age	-0.690	-0.002	0.029
	(0.005)***	(0.053)	(0.020)
Share Female	0.000	2.037	-0.297
	(0.000)	(1.490)	(0.674)
Constant	42.315	6.882	4.512
	(0.345)***	(2.533)***	(1.019)***
Observations	7510	89594	286008
R-squared	0.12	0.10	0.10

Standard errors in parentheses are clustered on schools, all models include year dummies * significant at 10%; ** significant at 5%; *** significant at 1%

Table 8b Statistical Discrimination, Model 3 from Table 5 but only schools in which foreign-born students outperform others
Parameter estimates from OLS estimation, covariates from model 3 presented in Table

	(1)	*(2)*	*(3)*
	Swedish	English	Mathematics
Foreign-born	0.750	1.088	0.911
	(0.184)***	(0.188)***	(0.257)***
TG	0.873	0.441	2.182
	(0.075)***	(0.071)***	(0.145)***
Interaction	-0.051	-0.145	-0.165
	(0.154)	(0.139)	(0.332)
Birthmonth	-0.106	-0.088	-0.108
	(0.013)***	(0.013)***	(0.022)***
Teacher Age	-1.019	0.189	-1.804
	(0.347)***	(0.282)	(1.681)
Share Female	-11.968	13.553	-12.981
	(10.034)	(5.239)**	(12.135)
Constant	61.784	-9.204	92.009
	(14.081)***	(14.109)	(76.838)
Observations	8520	11352	5476
R-squared	0.10	0.09	0.13

Standard errors in parentheses are clustered on schools, all models include year dummies * significant at 10%; ** significant at 5%; *** significant at 1%

Figure 1

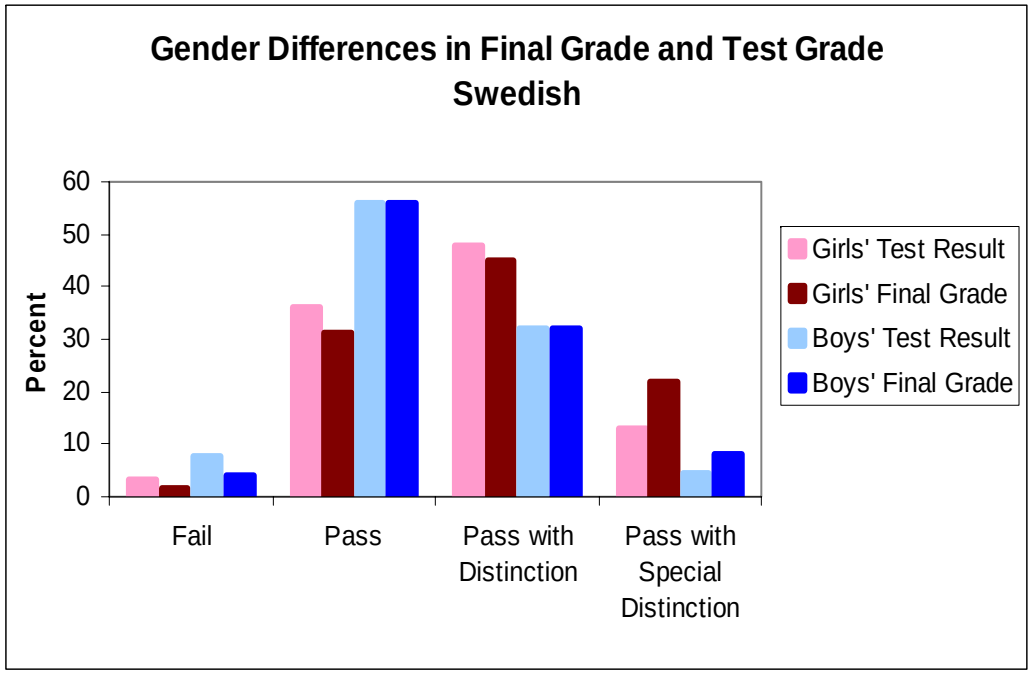

Gender Differences in Final Grade and Test Grade Swedish
Percent
60
50
40
30
20
10
0
Fail
Pass
Pass with Distinction
Pass with Special Distinction
Girls' Test Result
Girls' Final Grade
Boys' Test Result
Boys' Final Grade

Figure 2

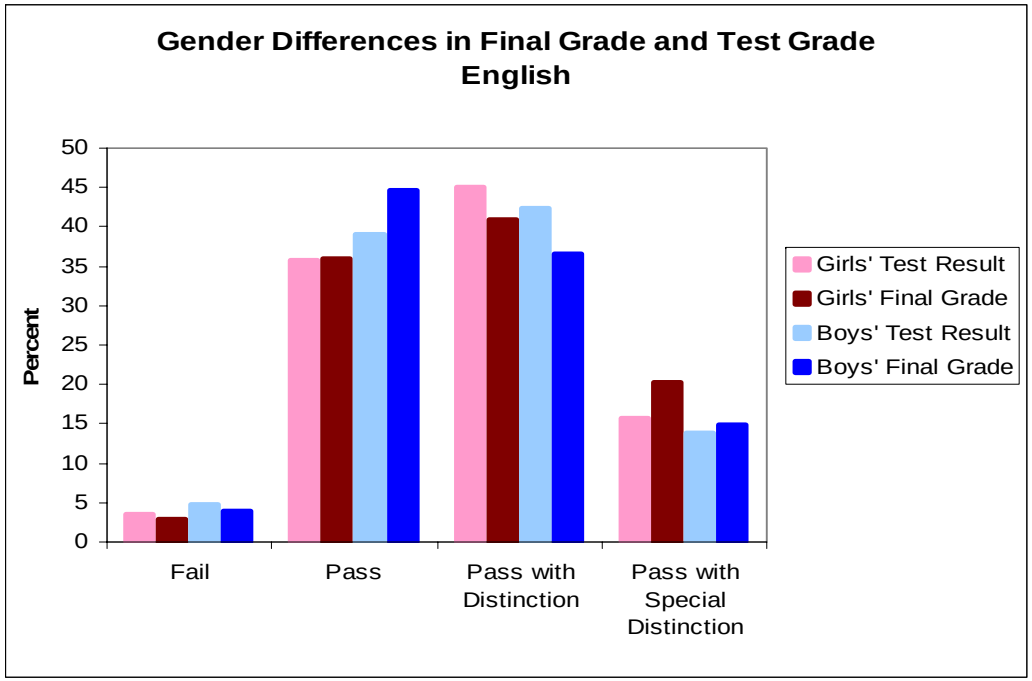

Gender Differences in Final Grade and Test Grade English
Percent
50
45
40
35
30
25
20
15
10
5
0
Fail
Pass
Pass with Distinction
Pass with Special Distinction
Girls' Test Result
Girls' Final Grade
Boys' Test Result
Boys' Final Grade

Figure 3

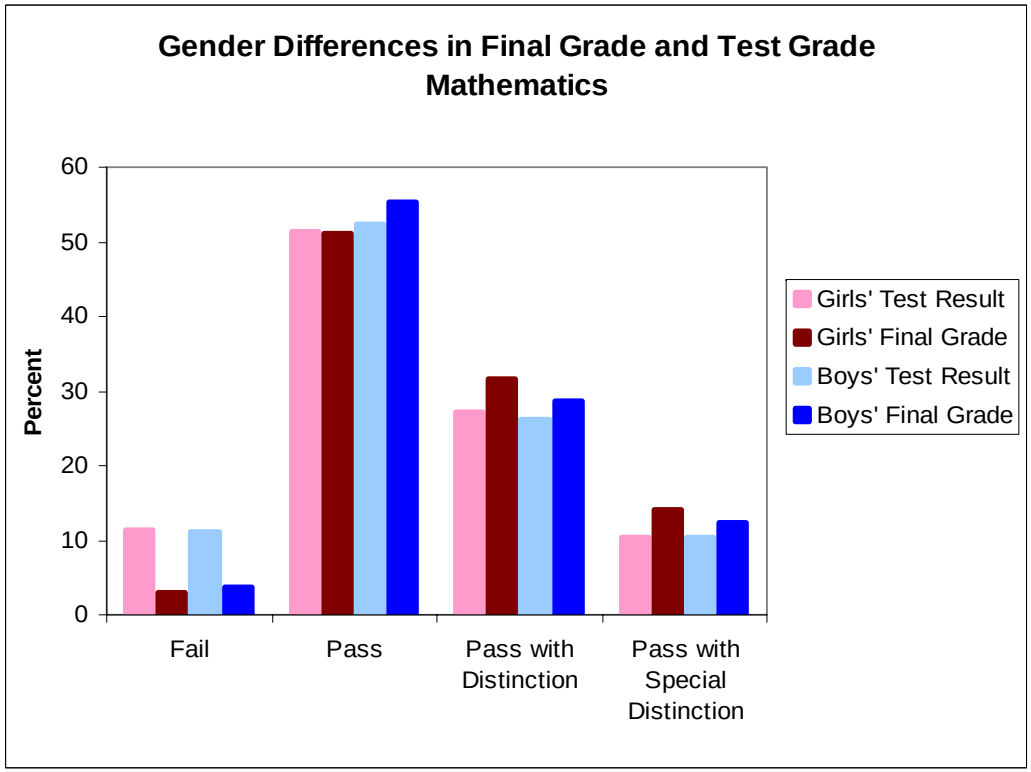

Gender Differences in Final Grade and Test Grade Mathematics
Percent
60
50
40
30
20
10
0
Fail
Pass
Pass with Distinction
Pass with Special Distinction
Girls' Test Result
Girls' Final Grade
Boys' Test Result
Boys' Final Grade

Figure 4

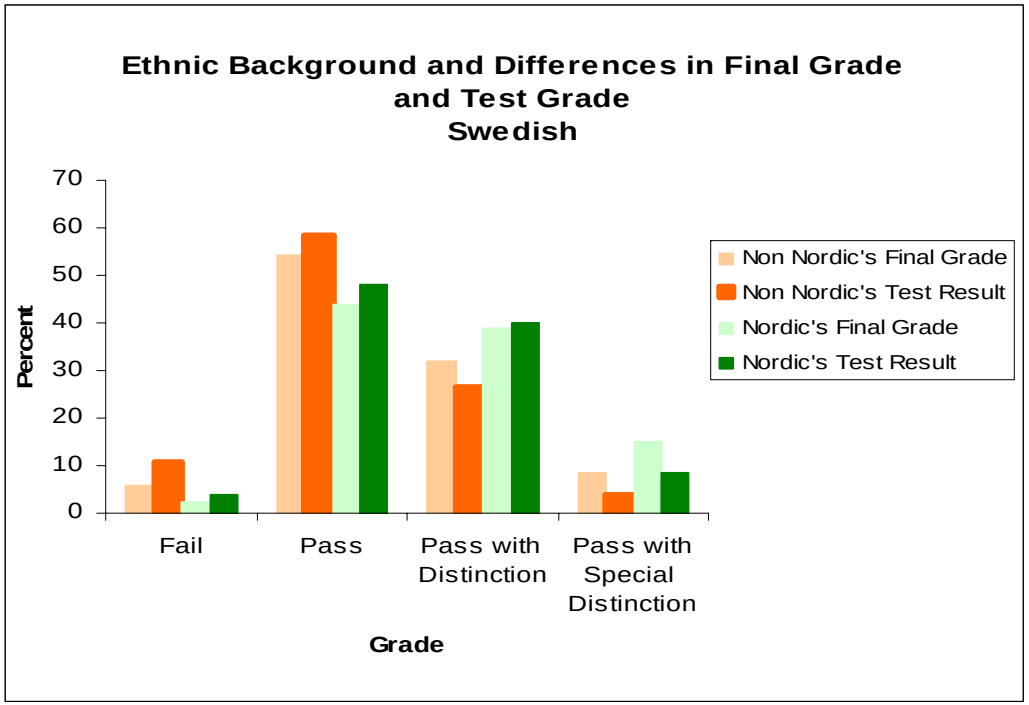

Ethnic Background and Differences in Final Grade and Test Grade
Swedish
Percent
70
60
50
40
30
20
10
0
Fail
Pass
Pass with Distinction
Pass with Special Distinction
Grade
Non Nordic's Final Grade
Non Nordic's Test Result
Nordic's Final Grade
Nordic's Test Result

Figure 5

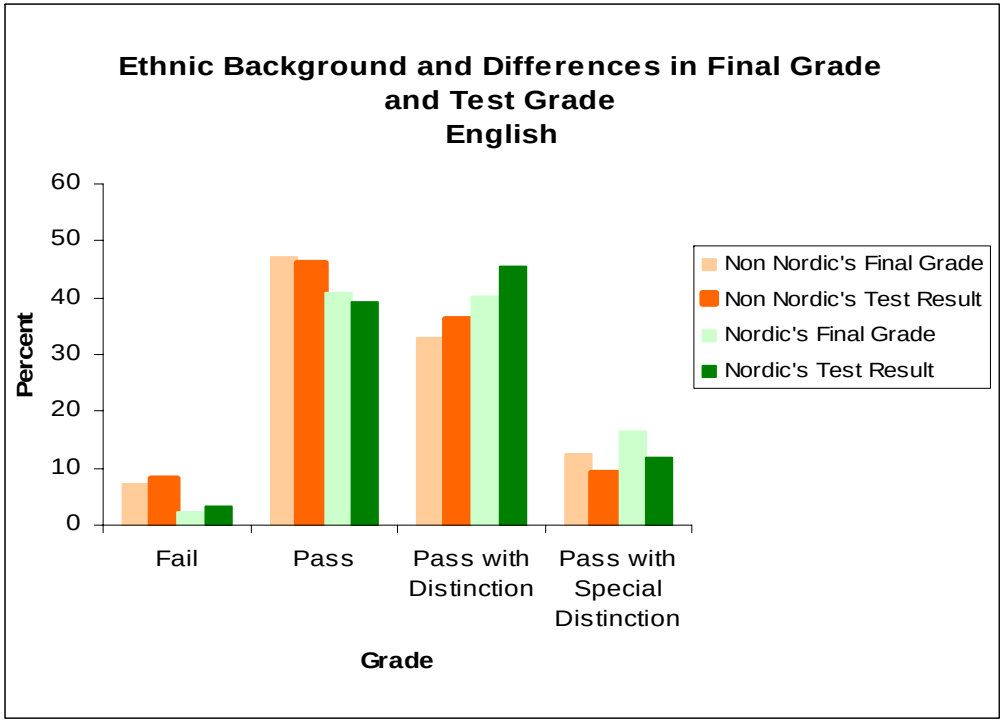
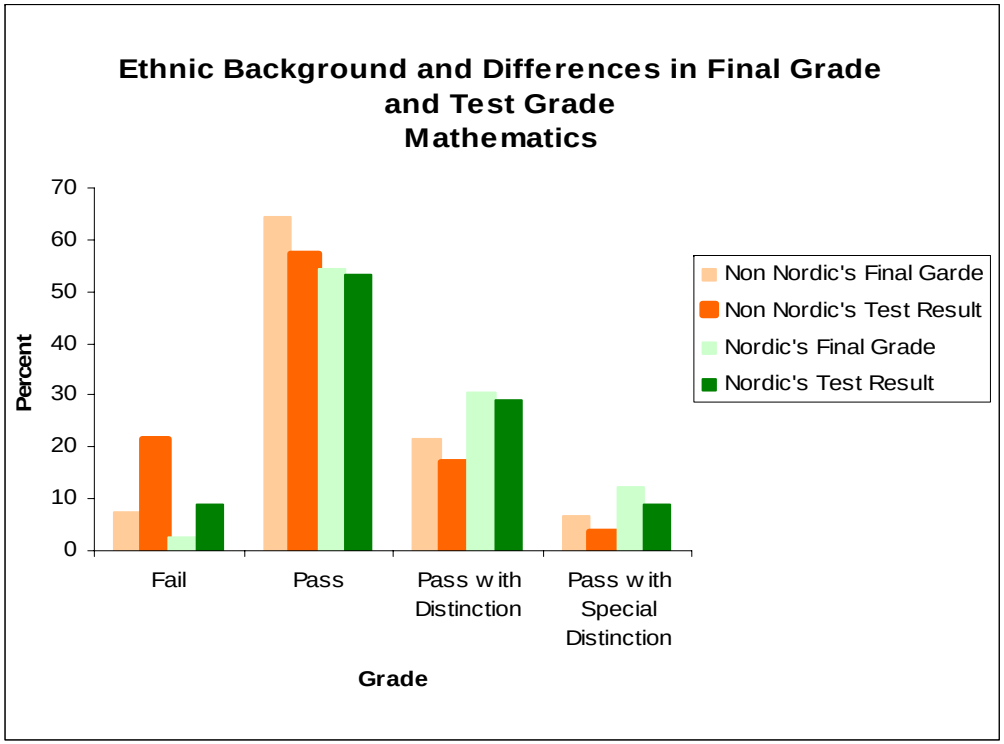

Ethnic Background and Differences in Final Grade and Test Grade
English
60
50
40
30
20
10
0
Percent
Fail
Pass
Pass with Distinction
Pass with Special Distinction
Grade
Non Nordic's Final Grade
Non Nordic's Test Result
Nordic's Final Grade
Nordic's Test Result

Figure 6


Ethnic Background and Differences in Final Grade and Test Grade Mathematics
Percent
70
60
50
40
30
20
10
0
Fail
Pass
Pass with Distinction
Pass with Special Distinction
Grade
Non Nordic's Final Garde
Non Nordic's Test Result
Nordic's Final Grade
Nordic's Test Result