



Educational evaluation schemes and gender gaps in student achievement[☆]



Torberg Falch^{a,b,*}, Linn Renée Naper^{c,d}

^a Department of Economics, Norwegian University of Science and Technology (NTNU), Dragvoll, N-7491 Trondheim, Norway

^b CESifo, Germany

^c EC Group AS, Beddingen 8, N-7014 Trondheim, Norway

^d Centre for Economic Research at NTNU, Norway

ARTICLE INFO

Article history:

Received 14 December 2011

Received in revised form 14 May 2013

Accepted 22 May 2013

JEL classification:

I21

J16

Keywords:

Educational evaluation schemes

Teacher grading

Gender gaps

Gender interactions

ABSTRACT

This paper investigates whether gender gaps in student achievement are related to evaluation schemes. We exploit different evaluations at the end of compulsory education in Norway in a difference-in-differences framework. Compared to the results at anonymously evaluated central exit exams, girls get significantly higher grades than boys when the same skills are assessed by their teacher. This gender grading gap in favor of the girls is found in both languages and mathematics. We find no evidence that the competitiveness of the environment can explain why boys do relatively better on the exam. We find some evidence that the gender grading gap is related to teacher characteristics, which indicates that the teacher–student interaction during coursework favors girls in the teacher grading.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Gender gaps in student test scores are observed throughout the world, most notably in favor of girls in languages (Machin & Pekkarinen, 2008), while the results in mathematics are more mixed (Guiso, Monte, Sapienza, & Zingales, 2008; Hyde, Lindberg, Linn, Ellis, & Williams, 2008). In addition, girls have recently improved their

position relative to boys (Hyde et al., 2008; Machin & McNally, 2005). Since literacy and numeracy skills are important determinants of success later in life, e.g., Murnane, Willett, and Levy (1995), Leuven, Oosterbeek, and Van Ophem (2004), and Heckman, Stixrud, and Urzua (2006), the gender achievement gaps might have important economic implications.

In this paper we analyze whether the observed gender gaps in student achievement are related to evaluation schemes by exploiting achievement scores for Norwegian students at the end of compulsory schooling. We find that girls get better grades than boys when assessed by their teacher compared to results at anonymously evaluated central exit exams. We investigate whether this gender grading gap in favor of girls is related to different competitiveness of the environment at the two evaluations and whether it is related to teacher characteristics.

Mechtenberg (2009) presents a game theoretical model in which different kinds of gender gaps are a result of teacher and student behavior in school. In equilibrium, the

[☆] Linn Renée Naper thanks the Ministry of Local Government and Regional Development for financial support. Comments from two anonymous referees, Hans Bonesrønning, Lars-Erik Borge, Julie Cullen, Astrid Kunze, and Sandra McNally are gratefully acknowledged. The authors bear the full responsibility for the analysis and the conclusions that are drawn.

* Corresponding author at: Department of Economics, Norwegian University of Science and Technology (NTNU), Dragvoll, N-7491 Trondheim, Norway. Tel.: +47 7359 6757.

E-mail addresses: Torberg.Falch@svt.ntnu.no (T. Falch), linnreeneaper@gmail.com (L.R. Naper).

gender gaps are similar to observed gender differences in school achievement, university enrollment, and wages. In her model, there are two subjects at school – mathematics and humanities – and students' beliefs about own abilities depend on teacher grading. The crucial assumption for the equilibrium is that girls do not fully trust bad grades in humanities and good grades in mathematics, while boys do not fully trust good grades in humanities. Teachers respond to these beliefs by easy grading of boys in humanities and of girls in mathematics, and hard grading of girls in humanities. Thus, the central theorem in [Mechtenberg \(2009\)](#) is the existence of a significant gender grading bias against girls in humanities and a smaller gender grading bias against boys in mathematics.

The observed gender gap in student achievement in favor of girls is often explained by increased share of female teachers. For example [Dee \(2005, 2007\)](#) and [Ammermueller and Dolton \(2006\)](#) find evidence that students profit from having a same-sex teacher. [Steel \(1997\)](#) discusses a phenomenon referred to as “stereotype threats” as an explanation of how demographic matches between students and teachers may influence educational outcomes. The idea is that students' academic self-confidence, and therefore their performance, is limited by possible and perceived stereotypes in the classroom. Another potential explanation, often referred to as “role-model” effects, is that the presence of a demographically similar teacher may raise students' academic motivation and expectations, and thus positively affects performance.

Both stereotype threats and role-model effects are “passive” teacher effects in that they are not related to intentional behavior of teachers. Thus, passive teacher effects cannot explain systematic differences in performance across evaluation schemes as far as they test the same skills.

The hypothesis in [Lavy \(2008\)](#) is that schools and teachers are sources of stereotypes that harm girls. The hypothesis is tested by exploiting that the matriculation exam in the academic track at Israeli high schools consists of both a state exam, which is anonymously graded, and an internal school exam. Contrary to the hypothesis, [Lavy \(2008\)](#) finds that the bias on the non-blind test is in favor of girls in all subjects.

Compared to the exam system in Israel, the potential for discrimination is higher in countries where teacher grading is based on more than a single test. In a review of the literature on gender differences in economic experiments, [Croson and Gneezy \(2009\)](#) argue that women's behavior is more context-dependent than men's behavior. If the way people treat others depends on their gender, the teacher–student interaction in coursework might induce statistical discrimination. The findings of [Emanuelsson and Fischbein \(1986\)](#), [Stobart, Elwood, and Quinlan \(1992\)](#), [Lindahl \(2007a\)](#), and [Bonesrønning \(2008\)](#) indicate, however, that placing greater weight on coursework elements in the evaluation improves the relative performance of girls. [Machin and McNally \(2005\)](#) present similar evidence. They show that when the importance of coursework in the examination system in the UK increased in 1988, the girls started to outperform the boys in the assessments.

In the Norwegian case, teacher set grades are based on written tests throughout the school year, and all students conduct a written central exit examination which evaluates the same skills and are graded anonymously. The students are randomly selected to an exit examination in either mathematics, English, or Norwegian language. All grades matter for admission to upper secondary schools and they are in this respect high-stake tests. We find that girls obtain better scores than boys in teacher grading relative to the central exit exam in all subjects in the empirical period 2002–2005. Thus, our results are not in accordance with [Mechtenberg's \(2009\)](#) central theorem. The gender grading gaps estimated are, however, similar to those found by [Lavy \(2008\)](#), [Bonesrønning \(2008\)](#), and [Lindahl \(2007a\)](#).

We investigate whether the finding in [Gneezy, Niederle, and Rustichini \(2003\)](#) that males perform relatively better in competitive environments can explain the estimated gender grading gaps. We exploit the variation across counties in the extent to which grades matter for admission to upper secondary schools. We also exploit the fact that one cohort conducted a separate low-stakes test. The results indicate that the competitiveness of the environment cannot explain the gaps. In addition, the results for the low-stakes test indicate that the gaps are not simply related to the anonymous vs. non-anonymous dimension. However, we find some evidence that the gender of the teacher and teacher experience matter for the gender grading gaps.

The next section offers a more detailed description of the Norwegian educational system and student evaluation schemes. Section 3 presents the data. Section 4 includes the main results on the gender grading gap in teacher assessments, while Section 5 investigates some possible explanations of the observed gender gap. Section 6 discusses the results and concludes.

2. Institutional setting

Norway has 10 years of compulsory schooling (from the year children turn six to the year they turn 16). None repeat grades, which implies that every student graduates on-time after 10 years. Multi-purpose municipalities are responsible for the schools and assign students to schools according to neighborhood rules. In 2005, 1164 public schools provided education at the lower secondary level (8–10th grade).

At the end of lower secondary education, students are evaluated both non-anonymously by their teachers (grades given in all curricula-based subjects) and anonymously in central exit exams. Each student takes one central written exit exam of 5 h, which take place at the end of the final year. The Norwegian Directorate for Education and Training prepares the written central exams, while local authorities are responsible for a random assignment of examination subjects to schools and individual students. The Directorate determines the share of students in each examination subject. The schools and the teachers have no influence in the assignment of examination subject. The students, as well as the schools, are informed about their exam subject on the same day all

Download English Version:

<https://daneshyari.com/en/article/6840915>

Download Persian Version:

<https://daneshyari.com/article/6840915>

[Daneshyari.com](https://daneshyari.com)