



Gender biases in student evaluations of teaching[☆]



Anne Boring

Sciences Po, OFCE, 69, quai d'Orsay, 75007 Paris, France
 PSL, Université Paris-Dauphine, LEDa, DIAL UMR 225, Paris F-75016, France
 IRD, LEDa, DIAL UMR 225, Paris F-75010, France

ARTICLE INFO

Article history:

Received 16 April 2016
 Received in revised form 2 November 2016
 Accepted 7 November 2016
 Available online 12 November 2016

JEL classification:

H8
 I23
 J16

Keywords:

Student evaluations of teaching
 Gender biases
 Gender stereotypes
 Teaching effectiveness

ABSTRACT

This article uses data from a French university to analyze gender biases in student evaluations of teaching (SETs). The results of fixed effects and generalized ordered logit regression analyses show that male students express a bias in favor of male professors. Also, the different teaching dimensions that students value in male and female professors tend to match gender stereotypes. Men are perceived by both male and female students as being more knowledgeable and having stronger class leadership skills (which are stereotypically associated with males), despite the fact that students appear to learn as much from women as from men.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

To what extent do gender biases influence the way that evaluators assess individual competence? I study this research question in the context of a widely used mechanism to assess competence: student evaluations of teaching (SETs). I find evidence that students discriminate in online evaluations of professors at a French university specialized in social sciences. Female professors receive lower

SET scores, despite evidence that female professors are as efficient instructors as their male colleagues.

The database that I use offers a unique opportunity to test for gender biases in SETs. The university requires first year undergraduate students to take six mandatory courses, so students do not select their courses when they register. Students' assignment to male and female professors is random. The administration makes students' online ratings of professors mandatory. As all students across all sections of a discipline take the same final exam, it is possible to compare student learning at the end of the semester. The database's properties therefore enable me to conduct the analysis in the context of a natural experiment. The database includes 20,197 observations of individual SET scores over five academic years, as well as student, professor, and course characteristics.

First, I study whether a match between student and professor gender has an impact on a professor's overall satisfaction score. Gender biases appear to exist: male students give significantly higher overall satisfaction scores to male professors than to female professors. Male students also rate male professors significantly higher than how female students rate both female and male professors. Male students are more likely to give *excellent* overall satisfaction scores to male professors. For instance, a male professor being rated by a male student is approximately 11 percentage points more likely

[☆] I would like to thank Stéphane Auzanneau for his help in collecting the different pieces of data, as well as Françoise Mélonio whose interest and support in this research project were essential. I would also like to thank Abdullah Al-Bahrani, Lee Badgett, Léopold Biarreau, Jen Brown, David Card, Sarah Cattani, Quoc-Anh Do, Manon Garrouste, Daniel Hamermesh, Benoît Kloeckner, Estelle Koussoubé, Ilyana Kuziemko, Cristina Lopez-Mayan, Ronald Oaxaca, Kellie Ottoboni, Hélène Périvier, Arnaud Philippe, Anna Raute, Georg Schaur, Ricarda Schmidl, Sarah Smith, Philip B. Stark, Myra Strober, Camille Terrier, Maxime Tô, Etienne Wasmer and Ulf Zölitz, as well as seminar participants at LEDa-DIAL, LIEPP, OFCE Sciences Po, the University Paris Dauphine, Southern Methodist University, Université de Franche-Comté, UT-Arlington, ENS-Lyon, and conference participants at AFSE, CTREE, EDGE, EEA, IAAE, IAFFE, JMA and RESUP for stimulating discussions and valuable comments and suggestions. This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 612413, for the EGERA (Effective Gender Equality in Research and the Academia) European project.

to be rated as *excellent* compared to how he would be rated by a female student. As a result, a male professor's expected *excellent* overall satisfaction score is approximately 20% higher than a female professor's expected *excellent* overall satisfaction score. I also find that students perform equally well on final exams whether their professor was a man or a woman, suggesting no difference in actual teaching effectiveness. Thus, the results suggest that differences in teaching skills are not driving gender differences in evaluations.

Second, I measure gender biases over different teaching dimensions related to course content and curriculum, learning assignments, course delivery style, and the perceived knowledge of the professor. I find that male and female students tend to give more favorable ratings to male professors on teaching dimensions that are associated with male stereotypes (of authoritativeness and knowledgability), such as class leadership skills and the professor's ability to contribute to students' intellectual development. I find that, on average, students rate female professors similarly to male professors for teaching skills that are more closely associated with female stereotypes (of being warm and nurturing), such as preparation and organization of classes, quality of instructional materials, clarity of the assessment criteria, usefulness of feedback on assignments, and ability to encourage group work.

The results are consistent with role congruity theory (Eagly and Karau, 2002): students may expect women to behave according to female gender stereotypes and men according to male gender stereotypes, while also evaluating overall teaching competence according to the characteristics of the stereotypical male professor (Kierstead et al., 1988; Basow et al., 2006; MacNell et al., 2014). These double standards are consistent with findings from studies conducted in experimental settings, in which the researchers were able to control for teaching styles (Arbuckle and Williams, 2003, and MacNell et al., 2014).¹

The fact that gender stereotypes may be driving students' ratings is consistent with statistical discrimination theory (Arrow, 1973; Phelps, 1972). According to this theory, evaluators may rely on stereotypes when assessing competence in contexts in which they lack information on actual productivity (Altonji and Blank, 1999). This theory suggests that when biased individuals are exposed to more information, they rely less on stereotypes, and they discriminate less. Here, I find that despite being exposed to male and female professors during entire semesters, students continue to discriminate in SET scores. A possible explanation could be that students are unable to assess actual teaching effectiveness, even after an entire semester.

As universities use SETs to decide on promotions and contract renewals, these results imply that promotion and hiring in universities may be biased (possibly unintentionally) against women. The gender biases that I find may therefore have negative consequences for female professors. These biases may also be harmful to female students, given the main results from the literature that discusses the impact of a role model effect on student performance (see Bettinger and Long, 2005; Dee, 2005; Hoffmann and Oreopoulos, 2009; Carrell et al., 2010). For instance, in the stereotypically male STEM fields, Carrell et al. (2010) find that female students perform better and are more likely to continue taking math and science courses when their introductory level professor was a woman. Given these results, gender biases in SETs may have a negative impact on female students' choices and success at the higher education level if competent female professors are offered fewer courses because of low SET scores.

This article is organized as follows. Section 2 explains the context of the natural experiment. Section 3 describes the data. Sections 4 and 5 examine the impact of student and professor gender on overall satisfaction scores, and on the different dimensions of teaching, respectively. Section 6 concludes.

2. The organization of courses and the SET system

The database presents a unique opportunity to test for the impact of gender biases in SETs. For several reasons linked to the organization of the first year mandatory undergraduate courses and how the SET scores are collected, the context satisfies the requirements of a natural experiment.

2.1. The "triplet" system

The first important feature of the database is that there is no selection bias of courses by students. First year undergraduates follow six mandatory courses: introduction to microeconomics, political institutions, and history during the fall semester; and introduction to macroeconomics, political science, and sociology during the spring semester. Students follow each course for 4 h a week: 2 h in a large lecture format (all taught by male professors) and 2 h in a small class section format called "seminars" (approximately 20 students per seminar). For each main lecture, there are between 43 and 49 seminars per year. The database includes students' individual evaluations of professors in the seminar classes of each of the six mandatory first year courses, for five academic years in a row (2008–2009 to 2012–2013). The data for the sociology and political science courses are for three academic years; these two courses were introduced as mandatory first year undergraduate courses in the 2010–2011 academic year.

The triplet system eliminates selection biases from students choosing based on a professor's gender. Students register for a fixed combination of three seminar professors (called a "triplet") for the fall semester mandatory seminars. All students of the same triplet therefore share the same combination of seminar professors. The administration creates the triplets such that each combination of three seminars offers similar advantages in terms of scheduling. For the fall semester courses, students register before the beginning of classes, and they are not allowed to change triplets after registration. After the fall semester, the administration requires students to stay together in the same triplet for the three spring semester seminars. The administration randomly assigns new seminar professors to each triplet for the spring semester courses.

This random assignment of new professors is convenient for the analysis of the spring semester courses. The triplet system imposes strong constraints on registration for the fall semester courses, which largely eliminate professor selection by students for the fall semester courses as well. To test for the absence of professor selection bias in the fall semester, I apply permutation tests.² I assume that if students were able to select professors, the share of male students would be related to the share of male professors in the teaching team. More specifically, male students with a preference for male professors would tend to register in triplets taught by more male professors in the fall semester. Overall, the correlation between the number of male students in a triplet and the number of male professors teaching the triplet is small and not statistically significant (Table 1). The sign of the correlation is inconsistent across years: sometimes positive, sometimes negative. For four out of five academic years, the correlation is not statistically significant.

In the 2012–13 academic year, however, the results of the permutation test suggests that there are significantly more male students

¹ Experimental settings suggest that students rate male and female professors differently even on objective criteria. In a reanalysis of the data from MacNell et al. (2014), Boring et al. (2016) find that students rate a female instructor as less prompt in grading assignments even in an experimental context in which the male and the female instructor graded assignments at exactly the same time.

² See Boring et al. (2016) for more information on the method.

Download English Version:

<https://daneshyari.com/en/article/5101824>

Download Persian Version:

<https://daneshyari.com/article/5101824>

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