STUDENT BIASES FOR MALE PROFESSORS IN STUDENT EVALUATIONS OF TEACHERS: WHAT CONSEQUENCES FOR FEMALE PROFESSORS?

Anne Boring, OFCE-Sciences Po and LEDa-DIAL

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 612413.

This paper argues that student evaluations of teachers (SETs) suffer from gender biases. Male students in particular tend to give higher overall satisfaction scores to male teachers, rewarding them for their perceived higher quality in course delivery style. These gender biases create different incentives for male and female teachers to change behaviors in order to improve their SET scores. As a consequence of student biases, female teachers tend to have to invest in more time-consuming dimensions of teaching if they want to increase their SET scores. These dimensions include the preparation, organization, and quality of course material. Male teachers, on the other hand, can invest in less time-consuming dimensions, such as leadership skills or discussing current events in class. Better teaching is not necessarily measured by SETs.

Keywords: Gender; Discrimination; Teaching effectiveness.

1. Introduction

Student evaluations of teachers (SETs) are the most common instrument that many universities throughout the world use to determine whether to promote or retain teachers. But for SETs to be a valid measure of teaching effectiveness, universities have to make three main assumptions. First, they must assume that students are unbiased. Second, they must assume that students know what a high quality teacher is. Third, they must assume that teachers do not act strategically to improve their SET scores.

This paper discusses the extent to which these three assumptions do not seem to hold, taking into account the existing literature and new evidence. This new evidence suggests that gender-based preferences by students bias all three assumptions regarding SET scores, thus preventing SETs from being a precise measure of teaching effectiveness. The paper more specifically analyzes the consequences of student biases in favor of male professors for female professors, using evidence from a French university (Boring 2014).

Understanding student gender-based biases is an important issue, because if male and female students do rate teachers differently according to gender, then male and female teachers are likely to face different incentives to improve their SET scores, and may thus have a lower likelihood of being retained. Also, if students valuate more time-consuming dimensions of teaching for one gender and less time-consuming dimensions of teaching for another gender, the teachers who need to invest effort in more time-consuming dimensions may thus have less time to invest in other tasks, such as research. Therefore, understanding gender biases in SET scores is essential to career management in academia.

2. Are students unbiased?

For SETs to be a valid measure of teacher productivity, universities must implicitly assume that students are objective evaluators. But there are good reasons to assume that students are in fact subjective evaluators of teacher productivity, since some evidence exists that the criteria on which

students judge their teachers are in part exogenous or unrelated to teachers' actual teaching qualities (e.g. DeWitte & Rogge 2011, McPherson 2006). If students are subjective evaluators applying stereotypes and expressing biases against some categories of teachers, the SET scores that these teachers receive are lower than their actual teaching ability.

2.1 Role model effect

Gender-based biases may have different effects on SET scores. First, a "role model" effect (e.g. Canes & Rosen 1995, Bettinger & Long 2005, Dee 2005, Hoffmann & Oreopoulos 2009, Carrell & West 2010) can partly explain how students evaluate their teachers. Assuming that students identify more closely with teachers of their own gender, male students may be more likely to rate male teachers higher, whereas female students may be more likely to rate female teachers higher.

Using data from first year undergraduate mandatory courses at a French university specialized in social sciences, Boring (2014) finds that male students tend to give much higher scores to male teachers in terms of overall satisfaction. Male students also tend to rate male teachers better than female students rate both female and male teachers. These results suggest that the role model effect, if it exists, could play for male students, but not necessarily for female students.

2.2 Stereotype effect

Second, a "stereotype effect" may influence SET scores. Assuming that students stereotype male teachers as more competent than female teachers, male students and female students are likely to both rate male teachers higher; if in fact students consider that female teachers are more competent than male teachers, then male and female students are likely to both rate female teachers higher.

In the fields of higher education and social psychology, this approach is called the *shifting standards theory* (Biernat, Manis & Nelson 1991, Biernat & Manis 1994). According to this theory, members of lower status groups tend to suffer from double standards when being evaluated: it is harder for the members of lower status groups to demonstrate their competence. On the other hand, members of higher status groups tend to be considered as competent: it is harder to prove incompetence for members of these higher status groups Basow, Phelan & Capotosto 2006, Foschi 2000). In the context of SETs, students may provide lower SET scores to female teachers, assuming that female teachers are part of a lower status group.

Also, the literature on higher education and social psychology suggests that male students, in particular, tend to rate according to gender stereotypes (Basow, Phelan & Capotosto 2006), and attribute different standards to male and female teachers. Boring (2014) finds that there is some econometric evidence that male students do apply different standards when evaluating male and female teachers. More specifically, it is harder for female teachers to obtain high SET scores from male students.

3. Do students know what high quality teaching is?

Through the use of SETs, universities delegate the task of monitoring teachers to students, since students can directly observe their teachers in the classroom. Yet, students are not necessarily able to evaluate what a high quality teacher is. The criteria according to which students find a teacher to be effective are not necessarily linked to a teacher's ability.

3.1 The role of teacher expressiveness in SETs

The way that students perceive teacher expressiveness tends to separate the teachers whom they consider as being "effective" from "ineffective" (Radmacher & Martin 2001). Teachers who want to increase their SET scores may choose to work on their extroversion and enthusiasm in class. However,

an experiment by Arbuckle and Williams (2003) suggests that students spontaneously rate (young) male teachers higher than female teachers controlling for a *same* level of teacher enthusiasm. If students express a positive bias for male teachers on extroversion (and hence on teacher expressiveness), then male teachers have a comparative advantage in this dimension of teaching.

While improving enthusiasm is not a time-consuming task, students' gendered expectations may create incentives for female teachers to invest in more time-consuming dimensions of teaching, such as course preparation or more detailed feedback on homework assignments (Sprague & Massoni 2005). Time-consuming teaching tasks mechanically reduce time available for other tasks, and may not even be all that rewarding for female teachers if students persevere in rating female teachers more severely.

In her analysis of undergraduate evaluations in a French university, Boring (2014) finds that the main reason why male students overrate male teachers is that male students perceive male teachers as being better in terms of delivery style and the course's link to current issues. While female teachers receive higher scores on the course content in itself (a more time-consuming task), their comparative advantage over male teachers is not as large as the comparative advantage that male teachers have in terms of delivery style.

3.2 Teaching methods

In a sense, it may be unfair to ask students to judge teaching effectiveness, considering that teachers themselves often disagree on the required qualities of an effective professor. Yet, most would agree that a desirable quality would be a teacher's ability to develop students' long-run knowledge. Carrell and West (2010), however, find that SET scores tend to be higher for teachers whose students perform well on contemporaneous course achievement, but not necessarily on follow-on achievement. If female teachers tend to focus more on learning techniques that aim for higher follow-on achievement, this may partly explain why they receive lower SET scores (Boring 2014).

4. Professors' strategic behavior according to gender

With SETs, teachers have two goals they need to achieve. First, they must be high quality teachers. Second, they must obtain high SET scores. As illustrated in the previous example, these two objectives are not always compatible. Some teachers may also decide to adopt strategies to obtain a high measure of teacher effectiveness, without necessarily improving their teaching.

4.1 The "Dr. Fox effect"

Teachers with little knowledge on a topic can earn high SET scores, with students judging teachers on the basis of their personalities rather than educational content. This effect has been called the "Dr. Fox effect", since the experiment by Naftulin, Ware and Donnelly (1973).

The results by Boring (2014) suggest that male teachers may be more likely to be Dr. Foxes, compared to female teachers. Indeed, male teachers benefit from student biases regarding class leadership skills. Because this teaching dimension tends to be a strong determinant of student overall satisfaction, male teachers who focus on leadership skills with a minimum of course preparation can obtain high SET overall satisfaction scores. However, it is harder for a female teacher to be a Dr. Fox, since students tend to be biased against female teachers on leadership skills and course delivery style in general.

4.2 Grade inflation

Another piece of evidence suggests that male and female teachers may not face similar incentives to improve SET scores involves students' perception of course grading. The economics literature has focused on the incentives for teachers to inflate students' grades in order to purchase higher SET

scores (e.g. Ewing 2012, Isely & Singh 2005, Krautmann & Saunder 1999, McPherson 2006). The social psychology literature suggests that students who receive poor grades tend to be harsher in their evaluations towards female teachers (referring to them as more incompetent), than towards male teachers who attribute equally bad grades (Sinclair & Kunda 2000). Said differently, female teachers who give more negative feedback to students receive poorer evaluations than male teachers who give equally negative feedback. Female teachers thus have higher incentives to give better grades and more positive feedback to students. If they do not give higher grades, they may receive lower SET scores.

5. Conclusion: what consequences for the academic environment?

These gender biases have a negative impact both on teachers and on students. Female teachers may be less likely to receive promotions or to be offered new courses. They may also find themselves with less time for tasks such as research activities, since they need to spend more time preparing courses due to student biases.

These gender biases may also have a negative impact on students. Assuming that fewer female teachers are kept on the teaching team due to poorer evaluation scores, this may have strong consequences on female students. Indeed, Carrell, Page and West (2010) show that female students tend to perform better and to take more courses in scientific and math courses when they had female teachers in introductory courses in college.

Finally, Boring (2014) argues that there is no difference between male teachers and female teachers if one measures effectiveness in terms of how well students perform on the final exam. Thus, SET scores could reflect the pleasure that students (especially male students) feel when attending classes, more than actual teaching quality.

References

Arbuckle, J. & Williams, B. D. (2003). Students' perceptions of expressiveness: Age and gender effects on teacher evaluations. In: Sex Roles. 49, pp. 507-516.

Basow, S. A., Phelan, J. E., & Capotosto, Laura (2006). Gender patterns in college students' choices of their best and worst professors. In: Psychology of Women Quarterly. 30, pp. 25-35.

Bettinger, E. P. & Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. In: American Economic Review. 95, pp. 152-157.

Biernat, M. & Manis, M. (1994). Shifting standards and stereotype-based judgments. In: Journal of personality and social psychology. 66, pp. 5-20.

Biernat, M., Manis, M. & Nelson, T. E. (1991). Stereotypes and standards of judgment. In: Journal of Personality and Social Psychology. 60, pp. 485-499.

Boring, A. (2014). Gender biases in student evaluations of teachers. Mimeo.

Canes, B. J. & Rosen, H. S. (1994). Following in her footsteps-faculty gender composition and women's choices of college majors. In: Industrial and Labor Relations Review. 48, pp. 486-504.

Carrell, S. E. & West, J. E. (2010). Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors. In: Journal of Political Economy. 118, pp. 409-432.

Carrell, Scott E., Page, Marianne E. & West, James E. Sex and science: How professor gender perpetuates the gender gap. In: The Quarterly Journal of Economics. 125, pp. 1101-1144.

De Witte, K. & Rogge, N. (2011). Accounting for exogenous influences in performance evaluations of teachers. In: Economics of Education Review. 30, pp. 641-653.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter?. In: American Economic Review. 95, pp. 158-165.

Ewing, A. M. (2012). Estimating the impact of relative expected grade on student evaluations of teachers. In: Economics of Education Review. 31, pp. 141-154.

Foschi, M. (2000). Double standards for competence: Theory and research. In: Annual Review of Sociology. 26, pp. 21-42.

Hoffmann, F. & Oreopoulos, P. (2009). A professor like me the influence of instructor gender on college achievement. In: Journal of Human Resources. 44, pp. 479-494.

Isely, P. & Singh, H. (2005). Do higher grades lead to favorable student evaluations?. In: The Journal of Economic Education. 36, pp. 29-42.

Krautmann, A. & Sander, W. (1999). Grades and student evaluations of teachers. In: Economics of Education Review. 18, pp. 59-63.
McPherson, M. A. (2006). Determinants of how students evaluate teachers. In: The Journal of Economic Education. 37, pp. 3-20.
Naftulin, D. H., Ware Jr, J. E. & Donnelly, F. A. (1973). The Doctor Fox lecture: A paradigm of educational seduction. In: Academic Medicine. 48, pp. 630-635.
Radmacher, S. A. & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. In: The Journal of Psychology. 135, pp. 259-268.
Sinclair, L. & Kunda, Z. (2000). Motivated stereotyping of women: She's fine if she praised me but incompetent if she criticized me. In: Personality and social psychology bulletin. 26, pp. 1329-1342.
Sprague, J. & Massoni, K. (2005). Student evaluations and gendered expectations: What we can't count can hurt us. In: Sex Roles. 53, pp. 779-793.

Author

Anne BORING, received her PhD in economics from the University Paris Dauphine in 2012, studying the political economy of trade. Since 2011, her love of teaching has brought her to serve as Assistant Dean for Undergraduate Studies in Economics and Mathematics at Sciences Po – Paris. This position has led her to investigate a new field of research: gender biases in higher education. She is now a post-doc researcher at Sciences Po, working on the EGERA project (Effective Gender Equality in Research and the Academia), which is co-financed by the European Commission and for which Sciences Po is the coordinator.
e-mail: anne.boring@sciencespo.fr