View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

NBER WORKING PAPER SERIES

A PROFESSOR LIKE ME: THE INFLUENCE OF INSTRUCTOR GENDER ON COLLEGE ACHIEVEMENT

Florian Hoffman
Philip Oreopoulos

Working Paper 13182
http://www.nber.org/papers/w13182

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2007

We are extremely grateful to Marianne Lynham, Matthew Hendrickson, Josie Lalonde, Rick Hayward, and Karel Swift for helping us obtain and work with the data used in this project, and to Rita Ahmadyar, Nicole Corbett, and Paul Ternamian for outstanding research assistance. We also thank seminar participants at UC Davis and the University of Toronto. Oreopoulos thanks the Social Sciences and Research Council and the Canadian Labour Market and Skills Researcher Network for financial support. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by Florian Hoffman and Philip Oreopoulos. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

A Professor Like Me: The Influence of Instructor Gender on College Achievement
Florian Hoffman and Philip Oreopoulos
NBER Working Paper No. 13182
June 2007
JEL No. I2

ABSTRACT

Many wonder whether teacher gender plays an important role in higher education by influencing student achievement and subject interest. The data used in this paper helps identify average effects from male and female college students assigned to male or female teachers. In contrast to previous work at the primary and secondary school level, our focus on large first-year undergraduate classes isolates gender interaction effects due to students reacting to instructors rather than instructors reacting to students. In addition, by focusing on college, we examine the extent to which gender interactions may exist at later ages. We find that assignment to a same-sex instructor boosts relative grade performance and the likelihood of completing a course, but the magnitudes of these effects are small. A same-sex instructor increases average grade performance by at most 5 percent of its standard deviation and decreases the likelihood of dropping a course by 1.2 percentage points. The effects are similar when conditioning on initial ability (high school achievement), and ethnic background (mother tongue not English), but smaller when conditioning on mathematics and science courses. The effects of same-sex instructors on upper-year course selection are insignificant.

Florian Hoffman
Department of Economics
University of Toronto
140 St. George St. Suite 707
Toronto, Ontario M5S3G7
Canada
flhoffmann@gmx.de

Philip Oreopoulos
Department of Economics
University of Toronto
140 St. George St. Suite 707
Toronto, ON M5S 3G7
Canada
and NBER
oreo@economics.utoronto.ca

I. Introduction

Education outcomes often differ by gender, and many of these differences seem to increase with age. The National Center for Education Statistics documents these trends for the United States [e.g. Freemen, 2004]. In early years, boys and girls appear to be on similar footing, performing about the same on tests of general knowledge, reading and mathematics. But by fourth grade, girls perform substantially better than boys at reading and slightly worse at mathematics. Gender differences by subject persist into high school and college, and occur for all reporting OECD countries. More men than women complete bachelor degrees in math, physical and computer sciences, business, and engineering. More men than women also graduate with masters and doctoral degrees in these subjects and complete programs in law, dentistry, and medicine, although these differences have narrowed over time.

While men tend to take more courses and perform better at subjects related to higher-paying occupations, women consistently outshine men in terms of overall educational attainment. Women are less likely to repeat a grade, drop out of high school, and more likely to enroll in college, finish college, and complete an advanced degree. Women currently receive 57 percent of all bachelor's degrees, and about 56 percent of all graduate degrees, reflecting steady increases since the early 1970s.

Role model effects are frequently considered key for explaining gender differences in education.[1] There is rich evidence within the psychology literature that girls and boys respond differently to mothers and fathers [e.g. Brown, 1990, Brown et al.,

[1] For a review of explanations of gender differences in education and specialization, see Jacobs (1996), and DiPrete and Buchmann (2005).

1986], and pick different celebrities and athletes to emulate. Male and female teachers are also potential role models. Students spend large portions of weekdays interacting with them. Perhaps not coincidently, females still constitute the majority of teachers in elementary and secondary schools during the period when girls repeat grades less than boys and form views about going to college. Conversely, male teachers, especially in college, dominate fields in mathematics, engineering, and sciences while male students enroll in these subjects more.

A few recent papers have used datasets with multiple student-teacher matches in elementary school to compare differences in student performance with differences in teacher gender for the same student. [Dee, 2006, 2007, Holmlund and Sund (2005)]. These studies improve on earlier ones by controlling for unobservable student traits that are common across the classroom, but they are not able to distinguish between role model effects – from students reacting to teachers depending on teacher gender – or teacher bias effects – from teachers reacting to students depending on student gender.

Our study is the first to estimate the impact of having a same-sex instructor on classroom performance in college using both within student and within instructor variation. Since we focus on large first-year undergraduate classes where teachers do not grade students' exams and students do not typically receive differential treatment from teachers, we can more confidently equate gender interaction effects with role-model effects. In addition, by focussing on college we examine the extent to which gender role model effects exist at later ages. Many social scientists wonder whether role model effects function mostly at young ages, and whether encounters at later ages can have any significant impact on social-economic success. Lastly, our paper speaks directly to the debate about increasing female representation in male-dominated fields. There have been many widely publicized efforts by the government, companies, and schools to increase female representation in math and science. This paper estimates the impact of male and female undergraduates' exposure to same sex teachers and whether such exposure can affect student achievement and subject interest.

Our results suggest teacher gender plays little or no role in student achievement and field of study choice. While we find some evidence that female students perform relatively better in terms of grade performance and are less likely to drop a course when encountered with a female instructor instead of male instructor, the magnitude of these effects are small. The evidence also holds when we consider subgroups across different subjects (mathematics and science), different pre-college ability (high school achievement), and different ethnicities (mother tongue not English).

II. Background

Teachers may respond differently depending on the gender of a student, or students may respond differently depending on the gender of a teacher. In the first case, teachers discriminate, and exhibit bias with respect to how they engage or evaluate boys and girls in the classroom. The way teachers behave interacting with boys or girls may depend on whether teachers themselves are male or female. These effects may be conscious or unconscious. In the second case, students may see teachers more as role models if they are of the same sex, and exhibit greater intellectual engagement, conduct, and interest. Students may also react to teachers when they fear being viewed through negative ‘stereotype threats’ – for example, when female students are reminded about a belief they are not supposed to be good at math when being taught by a male teacher. In one study [Spencer, Steele, and Quinn (1999)], for example, women underperformed men on a math test when told that the test produces gender differences but did not when told the opposite. Another possibility is that male and female students respond differently to male and female teaching styles. If girls and boys respond differently to teacher behavior rather than teacher gender per se, relative differences in academic achievement could still arise.

At the primary and secondary school level, a number of recent studies have estimated effects from being taught by a same-sex teacher, without attempting to disentangle why such effects exist. Results have been mixed. Nixon and Robinson (1999) regressed education attainment on the proportion of female faculty in an individual’s high school, using the National Longitudinal Survey of Youth (NLSY). With linear controls for family background, they concluded that raising the percentage of high school female faculty increases high school and college completion among girls, but decreases these outcomes among boys. On the other hand, Ehrenberg, Goldhaber, and Brewer (1995) adopted a better identification strategy by regressing individual test score gains between Grade 8 and Grade 10 on Grade 10 teacher gender and race characteristics using the National Educational Longitudinal Study (NELS). Their analysis suggested these characteristics have no affect on test scores, but do affect teachers’ subjective evaluations of students.

Dee (2007) also used the NELS but for Grade 8 students with two recorded subject outcomes. His study was the first to use a ‘matched pairs’ approach to estimate the effects of same-sex teachers on grade performance and teacher evaluations. Using the NELS, Dee examined whether test scores and student evaluations for boys and girls systematically differed between classes depending on teacher gender differences. Test scores were lower for boys assigned to female teachers while no difference occurred for girls. Dee argued that his data suggested female math teachers may have been assigned to lower-achieving classes, and therefore excluded the sample with math teachers in his baseline results. Test scores were about 4 percent of a standard deviation higher for girls assigned to female teachers and 4 percent of a standard deviation lower for boys. Female teachers were also more likely to believe boys are disruptive and don't do homework.

Holmlund and Sund (2005) adopted a similar approach using a large dataset of secondary students in Sweden. In contrast to Dee's results, they found no significant effects on grade performance. Carrington and Tymms (2005, 2007) used multiple classroom data for Grade 6 students from the Performance Indicators in Primary Schools (PIPS). They found no significant gender interactions for subject test score performance and subjective attitudes towards math, reading, and science. Lahelma (2000) interviews 13 and 14 year-olds about what they think about the importance of teacher gender. Although students often commented on the lack of male teachers in their schools, the issue of gender did not figure prominently in their observations about the quality of teaching that they valued. Students emphasized teachers who were engaging, friendly, sensitive, impartial, and able to maintain discipline, regardless of gender. Finally, in related work, Lavy and Schlosser (2007) used idiosyncratic variation in the proportion of girls in a class and conclude more girls in a class lowers disruption, improves student-teacher relationships and lessens teachers' fatigue.

Few studies have examined gender interactions at the college level. Canes and Rosen (1995) used year-to-year variation in the proportion of female faculty in a department and found no correlation with year-to-year variation in the proportion of females majoring in related subjects. On the other hand, Neumark and Gardecki (1998) found female graduate students in faculties with more women and with female advisors do better on the job, and Rothstein (1995) found that the probability a female college student obtains an advanced degree is positively associated with the percentage of faculty at her undergrad institution who are female. As with the earlier secondary school studies, many of these results are prone to possible omitted variables bias and apply only to limited cases. Bettinger and Long (2005) improved on this earlier work by using within course and student variation. They examined the impact of same-sex instructors on the choice of major and course credits and find small positive effects for females. Their data, however, did not allow them to explore interaction effects on more immediate classroom outcomes, such as course dropout and grade.

The existing research on the role of gender in higher education has been significantly hampered by lack of appropriate data. Most of the earlier studies are limited to small samples and prone to possible omitted variables bias. The data used in this paper provides better identification of student-teacher gender interactions in college, specifically at the classroom level. We use both within student and within class variation to estimate average counterfactual outcomes from male and female students assigned to male or female teachers. Our focus on large first-year undergraduate classes where teachers do not grade students' exams and students do not typically receive differential treatment from teachers, allows us to more confidently isolate gender interaction effects due to role-model effects rather than discrimination effects. In addition, by focussing on college we examine the extent to which gender interactions exist at later ages. Many social scientists wonder whether role model effects function mostly at young ages, and whether encounters at later ages can have any significant impact on social-economic success. Lastly, our paper speaks directly to the debate about increasing female representation in male-dominated fields.

III. Data and Statistical Methodology

Our study uses detailed student and instructor administrative data from the University of Toronto's Arts and Science Faculty. The data cover the Fall and Winter school year periods between 1996 and 2005. We focus on the 34,352 students that entered into full-time undergraduate programs from Ontario high schools, and were 17 to 20 years old on September 1 in the year of entry. We also focus on the 88 largest first year courses with at least 50 students in a section. This sample includes 85 percent of all first-year classes. Focussing on large courses minimizes the possibility that results depend on small and anomalous classes, and helps speed statistical computation.

We have enrolment data that include gender, date of birth, mother tongue, citizenship, entering program of study, and high-school grades. We also have data for registration status at the start of each Fall and Winter term, the number of credits students are enrolled in, financial status with the university, cumulative and current Grade Point Average (GPA), program of study, and graduation status. Our course data contain information on courses enrolled in and credits received for each year and each course. The data distinguish between course enrolment status on September 1, November 1, January 1, March 1, and the most current status. An advantage of this file is that it allows us to match to courses that students enrolled in before their first day of class, regardless of whether they completed the course or not. The course data also includes section information and final grade received, and is matched to instructors. We also use a number of objective and subjective teacher quality measures such as instructor rank and average evaluation score.[2]

We first estimate gender interactions for male and female students separately. Our initial empirical model takes on the following specification:

(1) $$y_{ikt} = \beta * f\_instructor_{ikt} + \delta_i + \delta_k + \delta_t + u_{ikt}$$

where y_{ikt} is a classroom or subject-specific outcome for student i taking course k in school year t, $f\_instructor_{ikt}$ is an indicator variable for whether a teacher have the is female, δ_i, δ_k, and δ_t are fixed effects for student, course, and year respectively, and u_{ikt} is the error term. β measures the average effect from assignment to a female versus male instructor, and captures both a gender interaction effect and an instructor quality effect (if males and females teach differently). The difference between the β coefficient for the female sample compared to male sample is the relative gender difference predicted from assignment to a female versus male instructor.

[2] See also Hoffmann and Oreopoulos (2006) for more description of related data.

To explore the importance of unobserved student and teacher characteristics, we replace student fixed effects with individual controls. We also explore the sensitivity of these estimates when including female indicators instead of fixed effects, and time-of-day controls. Remaining potential selection biases are mitigated by focusing on large classes with multiple sections where the final instructor allocation is not indicated in course calendars, and by focusing on first year students that have limited flexibility in choosing courses. We also explore (and find similar) results from using courses with only one instructor per year. This further removes students' ability to target particular courses.

Our data also allow for classroom fixed effects using the following specification:

$$(2) \qquad y_{ic} = \delta * f\_student_{ic} * f\_instructor_{ic} + \theta_{kg} + \delta_i + \delta_c + u_{ic}$$

where y_{ic} is a classroom or subject-specific outcome for student i in classroom c, $f\_student_{ic}$ is an indicator variable for whether a student is female, δ_i, and δ_c are fixed effects for student and course respectively, and θ_{kg} are course by gender fixed effects. These last controls allow gender differences in performance that are not attributable to teacher differences to vary across subjects courses. These are necessary to account for the possibility that the courses in which males and females tend to diverge are also the courses in which instructors tend to be more likely male or more likely female. The coefficient δ reflects the average outcome gain for females, relative to males, from assignment to a female versus male instructor or, conversely, the average outcome loss for males, relative to females, from assignment to a female versus male instructor.

Focusing on first-year students helps minimize gender-based course selection for two reasons. First, first year students cannot easily identify instructors, and especially gender of instructors, prior to enrollment. Course calendars at the University of Toronto usually do not indicate the instructor teaching the class, and when they do, only first initials are included. Second, first year students are inexperienced about teacher allocation mechanisms of the university and cannot rely either on their own or peer groups' past experience. We also restrict our sample to full year and first semester courses. Dropping courses taken in the second semester further minimizes opportunities for selecting courses by instructor. Students are matched to classes chosen before the first week of school. For purposes of comparison, we also include in the appendix separate and pooled results using second year classes. The possibility of selecting classes based on instructor is greater in second year, but the variety of courses and instructors teaching them is greater.

For our main sample, we tested for evidence of gender-specific selection by regressing the fraction of female students in a section on whether an instructor was female, conditioning on course or course-by-year fixed effects. There was no significant relationship.[3] The proportion of females in a class was consistently uncorrelated with the gender of the instructor under all specifications we tried. In addition, we estimated equation (2) with a student's high school grade as the outcome variable, and without student fixed effects. As expected in the absence of gender specific sorting, we found no

[3] Details of these results are available on request. The coefficient from regressing the fraction of female students in a first year classroom on whether an instructor was female, with course and year fixed effects is 0.004, with a standard error of 0.006. Results were similar when using course by year fixed effects or adding instructor and student background characteristics as controls.

relative differences in high school grades between males and females within classrooms.[4]

We use four student outcome variables at the student by course level: Whether students dropped the course ("Dropped Course"), the grade received for students that completed the course ("Grade"), the number of additional courses students take in the same subject in all subsequent years ("Subject Course, Subsequent Years"), and the number of subsequent credits received in the same subject in all subsequent years, ("Subject Credits, Subsequent Years"). To receive a credit requires both taking a course and passing it. Other than the binary variable, "Dropped Course", all variables are normalized for each course to have mean zero and standard deviation one.

Table 1 presents summary descriptive statistics for the sample of entering first year full time students between 1996 and 2004. The main dataset has one observation per student-class. Each student takes 4.2 half and full-year classes, on average. After restricting the sample to large full year and first semester classes, and dropping classes co-taught by male and female instructors, the average number of classes per student in our sample is 2.6. Sixty percent of first-year students are female. Fourteen percent of them take courses in math (usually calculus) compared to 17 percent of males. Sixteen percent of females take courses in chemistry and physics, compared to 15 percent of males. Notably, substantially fewer females compared to males take courses in business, economics, and computer science, but more take courses in psychology and sociology. Twenty-three percent of first-year instructors are female (24 percent, on average, per course). There are 1,450 classes within 88 courses over this 9 year period, with 16.8 classes on average per course, and 2.4 classes on average per course in each year. The

[4] The coefficient from regressing high school grade average (in a student's last year) on the interaction between being a female student and facing a female instructor, with female student, course-by-female-student, and classroom fixed effects is 0.03 percent, with a standard error of 0.16.

table indicates that course dropout and performance does not differ noticeably by gender across first year courses. Second year statistics are presented for comparison. By the second year, female students are slightly less likely to drop courses, have higher average grades and acquire less course credits than their male fellow students.

IV. Results

Table 2 presents estimates of equation (2) separately for male and female students. In the first two columns we regress student achievement on whether an instructor is female, controlling for course and school year. For females, we estimate no significant difference in the likelihood of dropping a class based on whether the instructor is male or female. Males, on the other hand, are about 1.8 percentage points more likely to drop a course when beginning a course with a female instructor. The difference between the female and male student effects is the predicted relative effect between gender groups from facing a female instead of male instructor. The second set of columns shows results from including student controls for students' last year of high school average grade, program of study, and age, and the third set of columns shows results from including student fixed effects across courses. Neither of these alternative specifications alters the point estimates by very much.

Without conditioning for student background, males perform slightly better, on average, with a male instructor. The estimated relative gain to male students from assignment to a male instructor is about 5 percent of a standard deviation, without student controls. This translates into a 0.6 percentage point increase in expected grade (out of 100 percent). When student controls or fixed effects are added, the estimated effect falls further, and we cannot reject that the estimated effect is zero. The relative effect falls and becomes statistically insignificant when student fixed effects are added, in part because the estimated effect from females with a female instructor is slightly negative.

Table 3 presents results after pooling males and females in the same regression. Column 1 shows the coefficient estimates of the female-student-female-instructor interaction, including course fixed effects and student background controls. These results are the same ones listed in column 6 of Table 1. They show the expected change in average achievement for females relative to males from assignment to a female instructor. This can also be interpreted as the expected relative loss in average male achievement from assignment to a female instructor. The coefficients in column 2 are the same ones listed in column 9 of Table 2 from including student fixed effects instead of student controls.

Pooling males and females together allows for classroom fixed effects. With classroom fixed effects and student controls in column 3, females are about 1 percentage point less likely than males in the same class to drop a course in a class with a female instructor. Conversely, males are 1 percentage point less likely than females to drop a class if the instructor is male. The 95 percent confidence region for these effects, however, includes zero. With classroom fixed effects and student controls, the difference between female and male average grade performance is 3.8 percent of a standard deviation higher (0.4 percentage points) with a female instructor. With both classroom and student fixed effects, the estimated effect is zero. Turning to subject interest, relative differences in male and female likelihood of taking related courses in subsequent years, and passing these courses, appear generally unaffected by whether a female or male teaches a first-year class.

Table 4 presents the results for sub-populations by mother tongue, type of course, and initial ability. For comparison, the first row replicates baseline findings in Table 3. The point estimates provide imprecise evidence that the estimated same-sex instructor effects on grade performance and course completion are larger for native English speaking students than for non-English speaking students, and smaller for math and science instructors than for social science instructors. In general, all the point estimates are small and mostly insignificant. We do not find evidence that the effects depend on students' initial ability (using high school entry grade as a proxy).[5]

V. Sample Selection for Grade Outcomes

Estimation of gender-interaction effects in college on grades is possible only for the sample of students that write the final exam. Table 2 suggests that the propensity to drop a course is affected by gender interactions as well. This creates a sample selection problem, formally described by the following set of equations:

$$(3) \qquad Grade_{ic} = \delta^{grade} * f\_student_{ic} * f\_instructor_{ic} + \theta_{kc}^{grade} + \delta_{i}^{grade} + \delta_{c}^{grade} + u_{kc}^{grade}$$

[5] We repeat the analysis for a sample of second year students and for a pooled sample of first and second year students.

(4)

$$Dropped_{ic} = \delta^{dropped} * f\_student_{ic} * f\_instructor_{ic} + \theta_{kc}^{dropped} + \delta_{i}^{dropped} + \delta_{c}^{dropped} + u_{kc}^{dropped}$$

(5) $$Grade_{kc} = 1[Dropout_{kc} \geq 0] * Grade_{kc}^{*}.$$

Equations (3) and (4) replicate equation (2) for “Grade” and “Dropped Course” as outcome variable, while equation (5) accounts for the potential selection bias. OLS-estimates of the parameter of interest, δ^{grade}, is biased if $\delta^{dropped}$ is different from zero. Our earlier analysis indicates that female students are indeed less likely to drop a course, relative to male students, when the class is taught by a female teacher (and vice versa).

Correcting for sample selection is difficult in our case since any variable affecting dropout behavior arguably also affects potential grades. Without exclusion restrictions, identification in a standard Heckman-selection model is solely based on the non-linearity of the correction term. Instead of relying on this source of variation we estimate upper bounds of δ^{grade} using a procedure similar to the ones described by Krueger and Whitmore (2002) and Lee (2005).

In general, OLS-estimates are downward biased if relatively more students stay to complete a course when the instructor is of the same sex, and if these marginal students are from the left tail of the grade distribution. We can therefore estimate an upper bound of δ^{grade} when applying OLS to a sample without the ($\delta^{dropped}$*100)-percent worst female students (relative to males) from female-taught classes.

We therefore apply the following procedure: In the first step we estimate dropout equations following the same specifications as in table 2. This provides us with an estimate of $\delta^{dropped}$, the female-male student difference in dropout behavior when taught by a female teacher. We then calculate the ($\delta^{dropped}$*100) percentile of the female-grade distribution for every class taught by a female teacher and drop all female students with a final grade lower than this percentile. Since we are focusing on selection due to the relative difference from having a female versus a male instructor between female and male students we do not need to trim marginal male students. In the second step we use this restricted sample to estimate the same equation as in the first step, but with final grade replacing the dropout variable.

The first set of columns in Table 5 presents these results. The upper bound effect on relative grade performance by gender is about 5 to 7 percent of a standard deviation. Thus, if same-sex instructors increase course completion for students at the bottom of the class, accounting for this selection leads to a small, but no longer insignificant gender interaction effect on grades. Expected grades may increase by up to 0.6 to 0.8 percentage points from being matched to a same-sex instructor.

In the second set of columns in Table 5, we repeat the same selection analysis, but from estimating the first-stage regression for each course separately. This yields course-specific estimates of ($\delta^{dropped}$*100), which are then used to trim the female-taught grade distributions within the same course. Since every student is allowed to take every course only once, a specification including individual fixed effects is not identified in this case. Table 5 reveals that the upper bound effect on grade performance is similar: assignment to a same sex instructor, leaving out students that finished the course because of same-sex assignment, increases relative grade performance by about 5 percent of a standard deviation (0.6 percentage points). These results suggest that, under conservative estimates that account for course completion effects, assignment to a same-sex instructor improves expected grade performance, but not by an amount that would substantially impact a student's GPA.

VI. Conclusion

In this paper, we address the importance of gender interactions between teachers and students at the college level to explain educational performance and subject interest. Our detailed administrative dataset from a large public university provides a rare opportunity to predict how classroom outcomes between males and females systematically differ depending on whether an instructor assigned to the class is male or female. Using within class variation for students taking multiple courses, we find that students react only marginally to an instructor's gender. Students taught by a same-sex instructor are about one percentage point less likely to drop a course (a 10 percent change from the mean). Relative grade performance is about 1 to 5 percent of a standard deviation better for students with a same-sex instructor. The small effects appear driven more by males performing worse when assigned to a female instructor, with females performing about the same. We also find no important influence from same-sex instructors on taking or passing subsequent courses in related subjects.

Our grade score estimates are somewhat smaller than the 5 to 10 percent standard deviation effects reported by Dee (2007) at the primary school level (using similar methodology), but not by much. Two possibilities may explain the difference. First, same-sex instructors may matter more at earlier ages, when development of cognitive and non-cognitive ability occurs more rapidly. Second, reactions from students over the gender of a teacher may matter less than reactions from teachers over the gender of a student. As mentioned earlier, college instructors do not typically interact on a one-on-one basis with students in large first year classes and do not typically grade tests, so there is less chance of discrimination. Another result that matches some of Dee's findings is that the interaction effect seems to stem more from male students performing worse with female instructors, while female performance appears unaffected.

We interpret these findings to suggest instructor gender plays only a minor role in determining college student achievement. Most of our baseline estimates imply that expected changes to performance and subject interest are small from same-sex instructor assignment, and many of these estimates are statistically insignificant. The small effects we do detect appear more due to social science courses than math or physical science courses, and do not appear to differ by initial student ability. It should be noted, however, that all the estimates in this paper relate to cases where one instructor is replaced at the margin for another who differs by gender. We cannot explore potential non-linear effects from more dramatic changes in the proportion of male or female faculty in a department or institution with this methodology.

The results are consistent with our earlier research [Hoffmann and Oreopoulos, 2006], which finds that observable instructor characteristics, such as rank, experience, and salary, do not explain differences in student performance. Subjective instructor quality, however, does predict these differences, although overall instructor effects are small. Hard-to-measure instructor qualities may matter more in predicting achievement, even for instructors that exhibit the same age, salary, rank, and gender.

References

Bettinger, Eric and Long, Bridget T. “Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students.” American Economic Review, 2005, pp. 152-157.

Canes, Brandice and Harvey Rosen. "Following in Her Footsteps? Faculty Gender Composition and Women's Choices of College Majors." Industrial & Labor Relations Review, April 1995, 48(3), pp. 486-504.

Carrington, B. , Tymms, P. & Merrell, C. (2007, in press) ‘Role models, school improvement and the “gender gap” – Do men bring out the best in boys and women the best in girls?’, British Educational Research Journal.

Carrington, B. , Tymms, P. & Merrell, C. (2005), “Role models, school improvement and the ‘gender gap’ Do men bring out the best in boys and women the best in girls?”, mimeo.

DiPrete, Thomas A. and Claudia Buchmann. “Gender-specific trends in the value of education and the emerging gender gap in college completion,” working paper, September 2005. Dee, T. S. (2004). “Teachers, race and student achievement in a randomized experiment.” The Review of Economics & Statistics, 86, 1, 195210.

Dee, Thomas S. “A Teacher Like Me: Does Race, Ethnicity or Gender Matter?” American Economic Review, 2005, 95(2), pp. 158-165.

Dee, Thomas S. “Teachers and the Gender Gaps in Student Achievement,” Journal of Human Resources, forthcoming.

Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). “Do teachers' race, gender, and ethnicity matter? Evidence from the National Education Longitudinal Study of 1988.” Industrial and Labor Relations Review, 48, 547561.

Freeman, C.E. (2004). Trends in Educational Equity of Girls & Women: 2004 (NCES 2005–016). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Hoffmann, Florian and Oreopoulos, Philip. “Professor Qualities and Student Achievement.” NBER Working Paper, No. 12596, 2006.

Holmlund, Helena and Sund, Krister. “Is the Gender Gap in School Performance Affected by the Sex of the Teacher?” Swedish Institute for Social Research, Stockholm University Working Paper No. 5, 2005.

Jacobs, Jerry A. “Gender inequality and higher education.“ Annual Review of Sociology, Vol. 22, 1996, pp. 153-85.

Reference: Krueger, A. & Whitmore, D. (2002). Would smaller classes help close the black-white achievement gap? In J. Chubb and T. Loveless (Eds.), Bridging the achievement gap. Washington, DC: Brookings Institute Press.

Lahelma, E. (2000). “Lack of male teachers: A problem for students or teachers?” Pedagogy, Culture and Society, 8, 2,17385.

Lavy, Victor. “Do Gender Stereotypes Reduce Girls’ Human Capital Outcomes? Evidence From a Natural Experiment.” NBER Working Paper No. 10678, 2004.

Lavy, Victor, and Analia Schlosser (2007). “Mechanisms and Impacts of Gender Peer Effects at School,” mimeo.

Lee, David (2005). “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects,” NBER Working Paper #11721.

Neumark, David and Rosella Gardecki. "Women Helping Women? Role Model and Mentoring Effects on Female Ph.D. Students in Economics." Journal of Human Resources, 1998, 33(1), pp. 220-46.

Nixon, Lucia A. and Michael D. Robinson. “The educational attainment of young women: Role model effects of female high school faculty,” Demography 36(2): May 1999, 185-194.

Rothstein, Donna S (1995). “Do Female Faculty Influence Female Students’ Educational and Labor Market Attainments?” Industrial and Labor Relations Review, vol. 48(3), pp. 515-530.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women’s math performance. Journal of Experimental Social Psychology, 35, 4-28.

TABLE 1 - DESCRIPTIVE STATISTICS

PANEL A: First Year Students

Variable	Full Sample			Female			Male		
	Mean	S.D.	*Sample Size*	Mean	S.D.	*Sample Size*	Mean	S.D.	*Sample Size*
Highschool Grade	85.2	5.8	*34,061*	85.5	5.6	*20,714*	84.6	6.0	*13,347*
Dropped Course	0.112	0.316	*98,861*	0.110	0.313	*58,592*	0.115	0.319	*40,269*
Grade	68.9	14.0	*87,775*	68.8	13.3	*52,121*	68.9	15.0	*35,654*
Subject Courses, Subsequent Years	1.443	2.916	*98,861*	1.370	2.814	*58,592*	1.550	3.055	*40,269*
Subject Credits, Subsequent Years	0.725	1.462	*98,861*	0.689	1.412	*58,592*	0.778	1.530	*40,269*
Female Teacher	0.246	0.431	*574*	0.248	0.432	*569*	0.241	0.428	*568*

PANEL B: Second Year Students

Variable	Full Sample			Female			Male		
	Mean	S.D.	*Sample Size*	Mean	S.D.	*Sample Size*	Mean	S.D.	*Sample Size*
Highschool Grade	85.5	5.7	*24,734*	85.8	5.5	*15,027*	85.1	5.9	*9,707*
Dropped Course	0.119	0.324	*56,744*	0.115	0.319	*33,751*	0.126	0.332	*22,993*
Grade	70.4	12.6	*49,966*	70.6	12.0	*29,873*	70.1	13.4	*20,093*
Subject Courses, Subsequent Years	2.371	3.128	*56,744*	2.376	3.059	*33,751*	2.364	3.225	*22,993*
Subject Credits, Subsequent Years	1.199	1.569	*56,744*	1.203	1.535	*33,751*	1.193	1.617	*22,993*
Female Teacher	0.24	0.43	*577*	0.24	0.43	*574*	0.24	0.43	*575*

TABLE 2 - ESTIMATED EFFECT OF FEMALE INSTRUCTOR ASSIGNMENT, BY GENDER

										Observations	
	Female	Male	Diff	Female	Male	Difference	Female	Male	Difference	*Female*	*Male*
Dropped Course	0.002 [0.008]	0.018 [0.008]**	-0.016 [0.008]*	0.001 [0.008]	0.015 [0.008]*	-0.014 [0.008]*	-0.007 [0.006]	0.01 [0.007]	-0.017 [0.008]**	*58,562*	*40,249*
Grade (with mean 0, stand. dev. 1)	-0.03 [0.027]	-0.076 [0.030]**	0.047 [0.028]*	-0.009 [0.025]	-0.035 [0.024]	0.026 [0.024]	-0.016 [0.016]	-0.002 [0.016]	-0.014 [0.018]	*52,121*	*35,654*
Subject Courses, Subsequent Years (with mean 0, stand. dev. 1)	-0.01 [0.018]	-0.046 [0.020]**	0.036 [0.022]*	-0.008 [0.018]	-0.041 [0.020]**	0.033 [0.022]	-0.019 [0.018]	-0.039 [0.019]**	0.019 [0.023]	*58,562*	*40,249*
Subject Credits, Subsequent Years (with mean 0, stand. dev. 1)	-0.009 [0.018]	-0.045 [0.020]**	0.036 [0.022]*	-0.006 [0.018]	-0.04 [0.020]**	0.033 [0.022]	-0.019 [0.018]	-0.038 [0.019]	0.019 [0.023]	*58,562*	*40,249*
Course FE		Yes			Yes			Yes			
Student FE		No			No			Yes			
Student Controls		No			Yes			No			

Notes: Each cell reports the coefficient of the student-teacher gender interaction from a separate linear probability regression. Regressions without individual FE include fixed effects for academic year. Student controls are: gender, highschool grade average and fixed effects age. One, two, and three astricies indicate statistical significance at the 10, 5, and 1 percent levels respectively.

TABLE 3 - ESTIMATED EFFECTS ON STUDENT PERFORMANCE FROM SAME-SEX INSTRUCTOR ASSIGNMENT

	(1)	(2)	(3)	(4)	*Sample Size*
Dropped Course	-0.015 [0.008]*	-0.017 [0.008]**	-0.01 [0.008]	-0.011 [0.007]	*98,811*
Grade (with mean 0, stand. dev. 1)	0.023 [0.024]	-0.014 [0.018]	0.038 [0.023]*	0.001 [0.017]	*87,775*
Subject Courses, Subsequent Years (with mean 0, stand. dev. 1)	0.034 [0.022]	0.019 [0.023]	0.034 [0.022]	0.02 [0.023]	*98,811*
Subject Credits, Subsequent Years (with mean 0, stand. dev. 1)	0.035 [0.022]	0.019 [0.023]	0.034 [0.022]	0.019 [0.023]	*98,811*
Course FE	Yes	Yes	No	No	
Student FE	No	Yes	No	Yes	
Classroom FE	No	No	Yes	Yes	
Student Controls	Yes	No	Yes	No	

Notes: Each cell reports the coefficient of the student-teacher gender interaction from a separate linear probability regression. All regressions include course-by-gender fixed effects. Student controls are: highschool grade average and fixed effects for academic year, age, mother tongue and program enrolled. One, two, and three astricies indicate statistical significance at the 10, 5, and 1 percent levels respectively.

TABLE 4 - ESTIMATED EFFECTS ON STUDENT PERFORMANCE FROM SAME-SEX INSTRUCTOR ASSIGNMENT
BY BACKGROUND CHARACTERISTICS

	GRADE (normalized)					**SUBJECT COURSES, SUBSEQUENT YEARS (normalized)**				
	(1)	(2)	(3)	(4)	*Sample Size*	(5)	(6)	(7)	(8)	*Sample Size*
Full Sample	0.023 [0.024]	-0.014 [0.018]	0.038 [0.023]*	0.001 [0.017]	*87,775*	0.034 [0.022]	0.019 [0.023]	0.034 [0.022]	0.02 [0.023]	*98,811*
Mother tongue: English	0.029 [0.027]	-0.011 [0.022]	0.046 [0.027]*	0.025 [0.033]	*59,375*	0.057 [0.027]**	0.027 [0.030]	0.065 [0.028]**	-0.026 [0.062]	*66,119*
Mother tongue: Other	0.016 [0.042]	-0.021 [0.025]	0.012 [0.046]	-0.007 [0.026]	*28,400*	-0.003 [0.041]	0.004 [0.039]	0.003 [0.044]	0.022 [0.042]	*32,692*
Major: Mathematics/Science	-0.004 [0.031]	0.014 [0.030]	0.008 [0.031]	0.014 [0.030]	*23,916*	0.009 [0.029]	0.018 [0.035]	0.021 [0.030]	0.029 [0.038]	*26,528*
Major: Other	0.045 [0.035]	0.013 [0.029]	0.07 [0.035]**	0.074 [0.039]*	*63,859*	0.077 [0.034]**	0.056 [0.036]	0.065 [0.033]**	-0.011 [0.054]	*72,283*
Below Highschool-Grade Median	0.031 [0.039]	-0.015 [0.029]	0.047 [0.036]	-0.008 [0.028]	*43,750*	0.038 [0.032]	0.026 [0.033]	0.034 [0.034]	0.027 [0.035]	*50,121*
Above Highschool-Grade Median	0.014 [0.029]	-0.014 [0.020]	0.023 [0.031]	-0.026 [0.029]	*44,025*	0.036 [0.032]	0.022 [0.035]	0.036 [0.032]	0.035 [0.041]	*48,690*
Course FE	Yes	Yes	No	No		Yes	Yes	No	No	
Student FE	No	Yes	No	Yes		No	Yes	No	Yes	
Classroom FE	No	No	Yes	Yes		No	No	Yes	Yes	
Student Controls	Yes	No	Yes	No		Yes	No	Yes	No	

Notes: Each cell reports the coefficient of the student-teacher gender interaction from a separate linear probability regression. All regressions include course-by-gender fixed effects. Student controls are: highschool grade average and fixed effects for academic year, age, mother tongue and program enrolled.One, two, and three astricies indicate statistical significance at the 10, 5, and 1 percent levels respectively.

TABLE 5 - EFFECTS ON GRADE PERFORMANCE FROM SAME-SEX INSTRUCTOR ASSIGNMENT WITH CORRECTION FOR SAMPLE-SELECTION

	TRUNCATION: OVERALL FEMALE DROPOUT BEHAVIOUR			**TRUNCATION: COURSE-SPECIFIC DROPOUT BEHAVIOUR**		
	(1)	(2)	(3)	(4)	(5)	(6)
Uncorrected gender interaction	0.023	-0.014	0.038	0.023	NA	0.038
	[0.024]	[0.018]	[0.023]*	[0.024]		[0.023]*
Sample Size		*87,775*			*87,775*	
Corrected gender interaction (Upper Bound)	0.068	0.068	0.046	0.048	NA	0.047
	[0.023]***	[0.023]***	[0.024]*	[0.024]**		[0.024]**
Sample Size	*87,641*	*87,641*	*87,714*	*87,687*	-	*87,694*
Course FE	Yes	Yes	No	Yes	Yes	No
Student FE	No	Yes	No	No	Yes	No
Classroom FE	No	No	Yes	No	No	Yes
Student Controls	Yes	No	Yes	Yes	No	Yes

Notes: The table shows uncorrected and sample-selection corrected estimates for the gender interaction when grade is used as outcome variable. We first estimate the gender-interaction in dropout-regressions (not shown in table). The estimate provides us with the x-percentage difference of the propensity to drop the course between female and male students when taught by a female teacher. We calculate x-percentage quintiles of the female grade distribution in female taught classes and drop all female students with grades below this quintile. Our upper-bound estimates come from regressions on the restricted sample. The first three rows show estimates when we trim the overall female grade distribution in female-taught classes. The last three rows repeat the analysis when we trim course-specific distributions instead. In this case, the specification with individual fixed effects is not identified. Each cell reports the coefficient of the student-teacher gender interaction from a separate linear probability regression. All regressions include course-by-gender fixed effects. Student controls are: highschool grade average and fixed effects for academic year, age, mother tongue and program enrolled. One, two, and three astricies indicate statistical significance at the 10, 5, and 1 percent levels respectively.

TABLE A1 - ESTIMATED EFFECTS ON STUDENT PERFORMANCE FROM SAME-SEX INSTRUCTOR ASSIGNMENT BY BACKGROUND CHARACTERISTICS, SECOND YEAR STUDENTS

	GRADE				SUBJECT COURSES, SUBSEQUENT YEARS			
	(1)	(2)	(3)	*Sample Size*	(4)	(5)	(6)	*Sample Size*
Full Sample	0.007 [0.032]	0.054 [0.026]**	-0.013 [0.032]	*49,966*	0.048 [0.032]	-0.011 [0.033]	0.022 [0.033]	*56,688*
Mother tongue: English	-0.001 [0.039]	0.071 [0.030]**	-0.02 [0.040]	*33,867*	0.062 [0.038]	0.004 [0.040]	0.039 [0.039]	*37,759*
Mother tongue: Other	0.019 [0.058]	0.009 [0.050]	0.027 [0.066]	*16,099*	0.009 [0.050]	-0.039 [0.050]	0.002 [0.049]	*18,929*
Major: Mathematics/Science	0 [0.130]	0.145 [0.083]*	-0.042 [0.145]	*7,202*	0.164 [0.097]*	-0.024 [0.137]	0.098 [0.107]	*8,221*
Major: Other	0.009 [0.033]	0.048 [0.028]*	-0.011 [0.032]	*42,764*	0.045 [0.034]	0 [0.034]	0.017 [0.034]	*48,467*
Below Highschool-Grade Median	-0.035 [0.048]	0.035 [0.036]	-0.056 [0.046]	*25,336*	0.049 [0.041]	0.006 [0.040]	0.012 [0.042]	*29,287*
Above Highschool-Grade Median	0.056 [0.039]	0.08 [0.034]**	0.041 [0.042]	*24,630*	0.039 [0.051]	-0.027 [0.049]	0.024 [0.055]	*27,401*
Course FE	Yes	Yes	No		Yes	Yes	No	
Student FE	No	Yes	No		No	Yes	No	
Classroom FE	No	No	Yes		No	No	Yes	
Student Controls	Yes	No	Yes		Yes	No	Yes	

Notes: Each cell reports the coefficient of the student-teacher gender interaction from a separate linear probability regression. All regressions include course-by-gender fixed effects. Student controls are: highschool grade average and fixed effects for academic year, age, mother tongue and program enrolled. One, two, and three astricies indicate statistical significance at the 10, 5, and 1 percent levels respectively.

TABLE A2 - ESTIMATED EFFECTS ON STUDENT PERFORMANCE FROM SAME-SEX INSTRUCTOR ASSIGNMENT BY BACKGROUND CHARACTERISTICS, FIRST AND SECOND YEAR STUDENTS

	GRADE				**SUBJECT COURSES, SUBSEQUENT YEARS**			
	(1)	(2)	(3)	*Sample Size*	(4)	(5)	(6)	*Sample Size*
Full Sample	0.016 [0.019]	-0.003 [0.015]	0.022 [0.019]	*137,741*	0.035 [0.019]*	0.029 [0.019]	0.03 [0.019]	*155,555*
Mother tongue: English	0.017 [0.022]	-0.001 [0.018]	0.022 [0.023]	*93,242*	0.056 [0.023]**	0.047 [0.025]*	0.057 [0.023]**	*103,897*
Mother tongue: Other	0.018 [0.035]	-0.012 [0.024]	0.021 [0.037]	*44,499*	-0.002 [0.034]	-0.004 [0.032]	0.004 [0.035]	*51,658*
Major: Mathematics/Science	-0.006 [0.030]	0.005 [0.028]	0.003 [0.030]	*31,118*	0.014 [0.029]	-0.005 [0.037]	0.026 [0.029]	*34,749*
Major: Other	0.024 [0.024]	0.024 [0.021]	0.03 [0.025]	*106,623*	0.061 [0.024]**	0.044 [0.025]*	0.04 [0.024]	*120,806*
Below Highschool-Grade Median	0.011 [0.031]	-0.009 [0.024]	0.019 [0.030]	*68,901*	0.027 [0.026]	0.031 [0.025]	0.019 [0.027]	*79,208*
Above Highschool-Grade Median	0.02 [0.024]	0.002 [0.017]	0.025 [0.024]	*68,840*	0.044 [0.027]	0.032 [0.029]	0.043 [0.027]	*76,347*
Course FE	Yes	Yes	No		Yes	Yes	No	
Student FE	No	Yes	No		No	Yes	No	
Classroom FE	No	No	Yes		No	No	Yes	
Student Controls	Yes	No	Yes		Yes	No	Yes	

Notes: Each cell reports the coefficient of the student-teacher gender interaction from a separate linear probability regression. All regressions include course-by-gender fixed effects. Student controls are: highschool grade average and fixed effects for academic year, age, mother tongue and program enrolled. One, two, and three astricies indicate statistical significance at the 10, 5, and 1 percent levels respectively.