SEX AND SCIENCE: HOW PROFESSOR GENDER PERPETUATES THE GENDER GAP*

SCOTT E. CARRELL
MARIANNE E. PAGE
JAMES E. WEST

Why aren't there more women in science? This paper begins to shed light on this question by exploiting data from the U.S. Air Force Academy, where students are randomly assigned to professors for a wide variety of mandatory standardized courses. We focus on the role of professor gender. Our results suggest that although professor gender has little impact on male students, it has a powerful effect on female students' performance in math and science classes, and high-performing female students' likelihood of taking future math and science courses, and graduating with a STEM degree. The estimates are largest for students whose SAT math scores are in the top 5% of the national distribution. The gender gap in course grades and STEM majors is eradicated when high-performing female students are assigned to female professors in mandatory introductory math and science coursework.

The inferior sex has got a new exterior. We got doctors, lawyers, politicians too. . . .

—Lennox and Stewart,
"Sisters Are Doin' It for Themselves" (1985)

I. INTRODUCTION

Why aren't there more women in science? During the past forty years, women have successfully entered many prestigious careers that were formerly dominated by men, and today graduate degrees in medicine, business, and law are almost equally divided between the sexes. In contrast, female college students are currently 37% less likely than males to obtain science and engineering B.A.s and females compose only 25% of the science, technology, engineering, and math (STEM) workforce (National Bureau of Economic Research 2005; National Science Foundation 2006).[1]

*Thanks go to USAFA personnel: J. Putnam, D. Stockburger, R. Schreiner, K. Carson, and P. Egleston for assistance in obtaining the data for this project, and to Deb West for data entry. Thanks also go to Charlie Brown, Caroline Hoxby, Deborah Niemeier, Kim Shauman, Douglas Staiger, Catherine Weinberger, and seminar participants at the NBER Higher Education Working Group, PPIC, SDSU, UC Davis, UC Irvine, UC Santa Barbara, UC Santa Cruz, and the University of Washington for their helpful comments and suggestions. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the USAF, the DoD, or the U.S. Government.

1. Among young workers in STEM careers, the fraction who are women is higher. For example, among STEM workers aged 30–39, 40% are female. This statistic, however, includes women in the biological sciences, who compose the

The Quarterly Journal of Economics, August 2010

What is the source of this discrepancy and why does it continue to exist when womens' expansion into other, traditionally male fields has been so much more rapid? This question has spurred hundreds of academic studies, widely publicized conferences, and government reports, but the exact manner in which cognitive and behavioral differences intertwine with social forces to produce differences in career outcomes remains a subject of spirited debate. Understanding how these possible mechanisms work is important: social scientists have shown that gender differences in entry into science careers explain a substantial portion of the gender pay differential among college graduates (Brown and Corcoran 1997; Weinberger 1999) and that the low representation of women in such careers may reduce aggregate productivity (Weinberger 1998).

What we do know is that through twelfth grade, the gender gap in math and science achievement tests is very small.[2] We also know that it has been declining over the past 20 years (Xie and Shauman 2003). The small differences that do exist in high school math and science achievement test scores are not predictive of men's higher likelihood of choosing a STEM career or major in college (Xie and Shauman 2003). *Conditional on proxies for ability*, the gender gap in the probability of completing a STEM degree is between 50% and 70% (Weinberger 2001). Nor are the nearly nonexistent differences in college preparatory math and science courses predictive of gender differences in college major (Xie and Shauman 2003; Goldin, Katz, and Kuziemko 2006). Because aptitude and preparedness of the two sexes upon entering college seem roughly equal, it seems that an important key to understanding the broader question of why men and women's representation in STEM careers is so different is understanding what happens to them during college.

This paper begins to shed light on this issue by exploiting data from the U.S. Air Force Academy (USAFA), where students are randomly assigned to professors for a wide variety of mandatory standardized courses. We focus on the role of professor gender.

majority of female STEM workers. Statistics from the National Science Foundation suggest that the gender gap in many STEM careers will continue to persist among young cohorts. For example, in 2002, women received only 21% of bachelor's degrees awarded in engineering, 27% in computer science, and 43% in physical science.

2. Some recent work by Ellison and Swanson (2009) and Pope and Sydnor (2010) suggests that there may be gender differences at the very upper tail of the ability distribution.

Why might professor gender affect female students' propensity to persist in STEM? Role model effects are frequently cited as potentially important factors affecting educational outcomes. Other factors might include gender differences in the academic expectations of teachers, differences in teaching styles, or differences in the extent to which teachers provide advice and encouragement. Experimental studies have documented that equally skilled men and women exhibit differences that might affect their career choices, including differences in self-perceptions of ability, preferences for taking on difficult tasks, levels of risk aversion, and expectations about future performance (Elliot and Harackiewicz 1994; Beyer and Bowden 1997; Eckel and Grossman 2008) but there is also a wide body of evidence suggesting that gender gaps in these characteristics are mutable (Spencer, Steele, and Quinn 1999). Teachers may be able to create an environment where this can occur.

Only a handful of studies have investigated the role of professor gender at the postsecondary level (Canes and Rosen 1995; Rothstein 1995; Neumark and Gardecki 1998; Bettinger and Long 2005; Hoffmann and Oreopoulos 2007), and all of these studies face identification challenges stemming from university students' ability to choose their courses and professors. Random placement of students into classrooms at the USAFA, together with mandatory math and science courses, allow us to investigate how professor gender influences student outcomes free of the self-selection and attrition problems that plague existing research. Because students are required to take specific math and science courses beyond the first year of study, we are also able to identify the long-term effects of professor gender. A further advantage of our data set is that course grades are not determined by an individual student's professor. Instead, all faculty members teaching the same course use an identical syllabus and give the same exams during a common testing period.[3] Our rich data, combined with the random assignment of students to professors in core math and science courses at the USAFA, allow us to overcome the self-selection and attrition problems that have limited the inferences that can be drawn from previous work in this area.

3. Although the students in Hoffman and Oreopoulos's data set are not randomly assigned and do not take mandatory STEM courses, their data set has one similarity to ours: course grades are determined by a general exam that is given to all students enrolled in the course, regardless of which professor they have taken the course from.

It is important to point out that if professor gender impacts female students, then these influences occur at a critical juncture in the life cycle. Decisions about choosing a STEM major are likely to have a substantial effect on future labor market opportunities. Furthermore, Xie and Shauman (2003) show that most women with a STEM bachelor's degree had initially planned on majoring in a non-STEM field. This suggests that the path toward a career in science is not primarily determined by the influence of social forces prior to college entry.

Our results suggest that although professor gender has only a limited impact on male students, it has a powerful effect on female students' performance in math and science classes, their likelihood of taking future math and science courses, and their likelihood of graduating with a STEM degree. The estimates are robust to the inclusion of controls for students' initial ability, and they are substantially largest for students with high SAT math scores. Indeed, among these students, the gender gap in course grades and college majors is eradicated when female students are assigned to introductory math and science professors who are female. The fact that we find the largest effects among high-ability women with a predisposition toward math and science is important because this group of women are, arguably, the set of women most suited for entering science and engineering careers. In contrast, the gender of professors teaching humanities courses has, at best, a limited impact on students' outcomes.

We also attempt to distinguish the role of professor gender itself from the role of other (unobservable) professor characteristics that are correlated with gender. We do this by estimating each professor's average "value-added" separately for male and female students. We find that some male professors are very effective at teaching female students—even more effective than they are at teaching male students. However, we find that the female introductory math and science professors continue to exert a positive influence on female students' long-run outcomes, even after controlling for professors' average value-added.

The remainder of the paper unfolds as follows: Section II describes our data set, and Section III discusses the statistical methods we will employ. In Section IV we present our main results. Section V investigates mechanisms, and Section VI concludes.

II. Data

Our data come from the USAFA, a fully accredited undergraduate institution of higher education with an approximate annual enrollment of 4,500 students. All students attending the USAFA receive 100% scholarship to cover their tuition, room, and board. Additionally, each student receives a monthly stipend of $845 to cover books, uniforms, computer, and other living expenses. All students are required to graduate within four years and typically serve minimum five-year commitments as commissioned officers in the United States Air Force following graduation.[4]

Despite the military setting, in many ways the USAFA is comparable to other selective postsecondary institutions in the United States. Similar to most selective universities and liberal arts colleges, USAFA faculty have earned their graduate degrees from a broad sample of high-quality programs in their respective fields. Approximately 40% of classroom instructors have terminal degrees, such as one might find at a university where introductory coursework is taught by graduate student teaching assistants. On the other hand, the number of students per section in any given course rarely exceeds twenty-five, and student interaction with faculty members in and outside of the classroom is encouraged. In this respect, students' learning experiences at USAFA more closely resemble those of students who attend small liberal arts colleges. Approximately thirty-two academic majors are offered at USAFA across the humanities, social sciences, basic sciences, and engineering.

Students at the USAFA are high achievers, with average math and verbal SAT scores in the eighty-eighth and eighty-fifth percentiles of the nationwide SAT distribution.[5] Students are drawn from each Congressional district in the United States by a highly competitive process, ensuring geographic diversity. Fourteen percent of applicants were admitted to the USAFA in 2007.[6] Approximately 17% of the students are female, 5% are black, 7% are Hispanic, and 6% are Asian. Seven percent of students at the

4. Special exceptions are given for religious missions, medical "set-backs," and other instances beyond the control of the individual.

5. See http://professionals.collegeboard.com/profdownload/sat_percentile_ranks_2008.pdf for SAT score distributions.

6. See the National Center for Education Statistics: http://nces.ed.gov/globallocator/.

USAFA have a parent who graduated from a service academy and 17% have a parent who previously served in the military.

Table I presents statistics for the USAFA and a set of comparison schools. We show the twenty-fifth and seventy-fifth percentiles of each school's verbal and SAT math scores, undergraduate enrollment, acceptance rates, and percent female for selected universities. SAT scores at the USAFA are comparable to the SAT scores of students at top-ranked public universities such as UCLA and UNC Chapel Hill, but, unlike students of these schools, only 17% of USAFA students are female. This characteristic makes the USAFA most comparable to selective universities that have strong traditions in science and technology, such as the Georgia Institute of Technology or Rensselaer Polytechnic Institute. Our results are thus most salient for women who enter college with a predisposition toward STEM. Although this group is not representative of all female college students, it is a group of particular relevance to the question under study. If professor gender has important effects among high-ability women who are already interested in science, and who have selected into an environment that is predominantly male, then the results have strong implications for the type of women who are most likely to choose to major in STEM out of high school. Put differently, our estimates probably speak most directly to retaining women with an interest in STEM, rather than the question of what causes women to enter STEM majors.

II.A. The Data Set

Our data set includes 9,015 students who compose the USAFA graduating classes of 2001 through 2008. Data for each student's high school (pretreatment) characteristics and his or her achievement while at the USAFA were provided by USAFA Institutional Research and Assessment and were stripped of individual identifiers by the USAFA Institutional Review Board. Student-level pretreatment data include whether students were recruited as athletes, whether they attended military preparatory schools, and measures of their academic, athletic, and leadership aptitude. Academic aptitude is measured through SAT verbal and SAT math scores and an academic composite computed by the USAFA admissions office, which is a weighted average of an individual's high school GPA, class rank, and the quality of the high school attended. The measure of pretreatment athletic aptitude is a score

TABLE I
COMPARISON SCHOOLS

	Percent female	SAT verbal 25th	SAT verbal 75th	SAT math 25th	SAT math 75th	2007 undergraduate enrollment	Percent admitted
Kettering University	14.9	510	630	600	690	2,178	23.0
Air Force Academy	18.6	590	670	620	700	4,461	14.0
Rose–Hulman Institute of Technology	20.6	560	680	630	710	1,936	69.7
Rensselaer Polytechnic Institute	26.6	600	690	650	730	5,146	49.4
Georgia Tech	28.6	590	690	650	730	17,936	28.0
California Institute of Technology	30.6	700	780	770	800	913	16.9
Virginia Tech	41.6	530	630	570	670	23,041	67.1
Case Western Reserve University	42.3	580	690	620	720	4,207	74.7
UCLA	44.7	570	680	610	720	25,928	25.8
University of Illinois at Urbana Champaign	46.9	550	670	640	740	31,472	71.0
University of Michigan	50.3	590	690	630	730	25,555	50.3
UC San Diego	52.6	540	660	600	700	22,048	45.6
University of Virginia	55.8	590	700	610	720	15,078	35.2
UNC Chapel Hill	58.7	590	690	610	700	17,628	34.1

Note. Data originally from National Center for Education Statistics (2007–2008).

on a fitness test required by all applicants prior to entrance.[7] The measure of pretreatment leadership aptitude is a leadership composite computed by the USAFA admissions office, which is a weighted average of high school and community activities (e.g., student council offices, Eagle Scout participation, captaining a sports team).

Table II provides summary statistics, and Figure I plots the distribution of pretreatment academic variables by gender. As in nationally representative samples, the upper tail of the math score distribution is somewhat thicker for male than it is for female students. Because our estimation strategy is based on random assignment and includes pretreatment characteristics as controls, small differences in distributions will not affect our analysis.

Our academic performance measures consist of final grades in core courses for each individual student by course and section–semester–year. Students at USAFA are required to take a core set of approximately thirty courses in mathematics, basic sciences, social sciences, humanities, and engineering, but we focus only on mandatory introductory and follow-on courses in mathematics, physics, chemistry, engineering, history, and English.[8] A distinct advantage of our data set is that all students are required to take a follow-on related curriculum. Grades are determined on an A, A−, B+, B, ... , C−, D, F scale, where an A is worth 4 grade points, an A− is 3.7 grade points, a B+ is 3.3 grade points, etc. The sample grade point average in core STEM coursework is 2.72 among females and 2.85 among males. The grade point average in core humanities courses is 2.81 among females and 2.73 among males. We standardize these course grades to have a mean of zero and a variance of one within each course, semester, and year.

We also examine students' decisions to enroll in optional follow-on math and science classes, whether they graduate with a bachelor's degree, and their choice of academic major. In our sample, female students are less likely than males to take higher-level elective math courses (34% of females versus 51% of males) and

7. Barron, Ewing, and Waddell (2000) find a positive correlation between athletic participation and educational attainment and Carrell, Fullerton, and West (2009) find a positive correlation between fitness scores and academic achievement.

8. Course descriptions for Math 130, 141, 142, Physics 110, 221, Chemistry 141, 142, History 101, 202, English 111, 211, and the required engineering courses (aeronautical, astronautical, electrical, mechanical, civil, and thermodynamics) can be found at http://www.usafa.edu/df/dfr/curriculum/CHB.pdf. Additionally, Carrell and West (2010, Table I) provides a list of the required STEM courses at USAFA.

TABLE II
SUMMARY STATISTICS

	Female students		Male students	
Student-level variables	Observations	Mean (std. dev.)	Observations	Mean (std. dev.)
Total course hours	1,504	25.71	7,511	25.56
		(5.89)		(6.13)
Math and science core course grades	7,547	−0.09	36,739	0.02
(normalized by course by semester)		(1.00)		(1.00)
English and history core course grades	5,349	0.08	27,274	−0.02
(normalized by course by semester)		(0.99)		(1.00)
Withdraw in first year	1,504	0.06	7,511	0.07
		(0.23)		(0.25)
Withdraw in first or second year	1,504	0.14	7,511	0.15
		(0.35)		(0.36)
Take higher-level math elective	1,504	0.35	7,511	0.51
		(0.48)		(0.50)
Take higher-level humanities elective	1,504	0.25	7,511	0.22
		(0.43)		(0.42)
Graduate	1,504	0.84	7,511	0.81
		(0.37)		(0.39)
Graduate with a math, science, or engineering degree	1,504	0.41	7,511	0.46
		(0.49)		(0.50)
Graduate with a math, science, or engineering degree	1,504	0.25	7,511	0.41
(excludes biological sciences)		(0.43)		(0.49)
Graduate with a humanities degree	1,504	0.10	7,511	0.07
		(0.30)		(0.26)
Proportion female professors	1,492	0.23	7,430	0.23
(introductory math & science)		(0.27)		(0.28)

TABLE II
(CONTINUED)

Student-level variables	Female students		Male students	
	Observations	Mean (std. dev.)	Observations	Mean (std. dev.)
Proportion female professors (introductory humanities)	1,489	0.16	7,437	0.15
		(0.28)		(0.27)
SAT verbal	1,504	637.65	7,511	630.05
		(67.08)		(64.41)
SAT math	1,504	650.21	7,511	666.40
		(59.72)		(61.24)
Academic composite	1,504	13.11	7,510	12.62
		(1.97)		(2.17)
Algebra/trigonometry placement score	1,496	59.89	7,461	62.79
		(19.13)		(19.39)
Leadership composite	1,503	17.65	7,503	17.23
		(1.92)		(1.83)
Fitness score	1,502	4.67	7,510	4.86
		(0.92)		(0.94)
Black	1,504	0.07	7,511	0.05
		(0.25)		(0.21)
Hispanic	1,504	0.08	7,511	0.07
		(0.27)		(0.25)
Asian	1,504	0.07	7,511	0.04
		(0.26)		(0.20)
Recruited athlete	1,504	0.31	7,511	0.26
		(0.46)		(0.44)
Attended preparatory school	1,504	0.16	7,511	0.21
		(0.36)		(0.41)

TABLE II
(CONTINUED)

	Math, physics, and chemistry introductory courses			
	Female professors		Male professors	
	Observations	Mean (std. dev.)	Observations	Mean (std. dev.)
Professor-level variables				
Number of sections per instructor	47	6.09	202	4.61
		(4.29)		(3.36)
Instructor is a lecturer	47	0.57	200	0.42
		(0.50)		(0.49)
Instructor is an assistant professor	47	0.30	200	0.37
		(0.46)		(0.48)
Instructor is an associate/full professor	47	0.13	202	0.22
		(0.34)		(0.42)
Instructor has a terminal degree	47	0.28	199	0.43
		(0.45)		(0.50)
Instructor's teaching experience	47	3.17	199	4.81
		(3.16)		(6.05)
Class-level variables				
Class size	286	19.18	935	18.97
		(3.10)		(3.97)
Average number of female students	286	3.31	935	3.26
		(1.81)		(1.99)
Average class SAT verbal	286	625.16	935	625.78
		(22.55)		(27.04)
Average class SAT math	286	653.42	935	651.26
		(28.69)		(32.60)
Average class academic composite	286	12.47	935	12.40
		(0.89)		(1.02)
Average class algebra/trig score	286	58.03	935	56.58
		(11.97)		(12.24)

TABLE II
(CONTINUED)

	English and history introductory courses			
	Female professors		Male professors	
	Observations	Mean (std. dev.)	Observations	Mean (std. dev.)
Professor-level variables				
Number of sections per instructor	24	6.92	88	8.93
		(5.77)		(7.42)
Instructor is a lecturer	24	0.54	88	0.52
		(0.51)		(0.50)
Instructor is an assistant professor	24	0.42	88	0.33
		(0.50)		(0.47)
Instructor is an associate/full professor	24	0.04	88	0.15
		(0.20)		(0.36)
Instructor has a terminal degree	24	0.17	88	0.32
		(0.38)		(0.47)
Instructor's teaching experience	24	3.35	88	4.42
		(3.31)		(5.04)
Class-level variables				
Class size	166	15.14	786	16.10
		(4.86)		(3.89)
Average number of female students	166	2.58	786	2.58
		(1.83)		(1.74)
Average class SAT verbal	166	623.12	786	627.88
		(28.18)		(27.89)
Average class SAT math	166	659.01	786	662.25
		(28.34)		(27.21)
Average class academic composite	166	12.75	786	12.64
		(0.94)		(0.96)
Average class algebra/trig score	166	61.67	786	61.92
		(8.57)		(8.03)

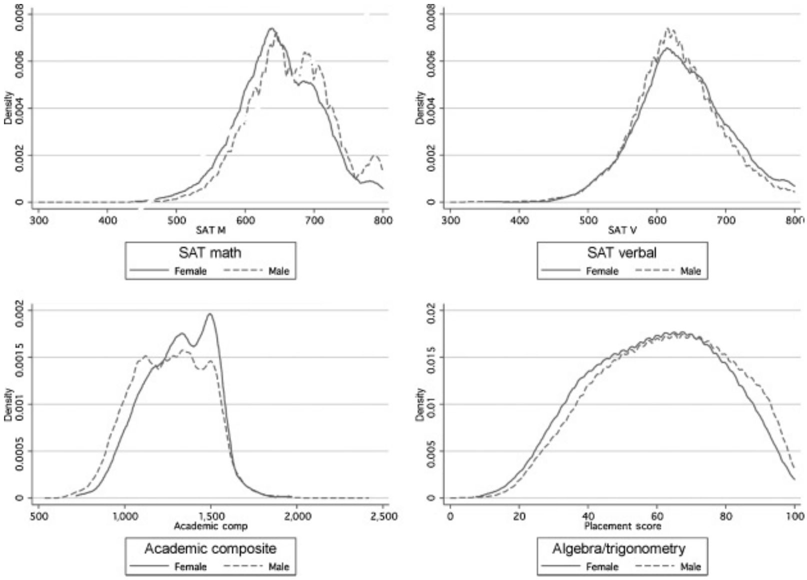

FIGURE I

Distribution of Academic Pretreatment Measures by Gender

Figures represent the distribution of pre-Academy characteristics by student gender for the USAFA graduating classes of 2001–2008.

less likely to major in STEM (24% versus 41% but are more likely to graduate (84% versus 81%).[9]

Individual professor-level data were obtained from USAFA historical archives and the USAFA Center for Education Excellence and were matched to the student achievement data for each course taught, by section–semester–year.[10] We have information on each professor's academic rank, gender, education level (M.A. or Ph.D.), and years of teaching experience at USAFA. During the period we study, 249 different faculty members taught introductory mathematics, chemistry, or physics courses. Nineteen percent (47 of 249) of these faculty were female and taught 23% (286 of 1,221) of the introductory math and science course sections. One hundred twelve different faculty members taught humanities courses,

9. Figures for STEM majors exclude the biological sciences, which require less mathematics and have historically higher rates of female participation. When biological sciences are included the gender difference is smaller (40% versus 45%).

10. We were only able to obtain the professor observable data for the mathematics, chemistry, physics, English, and history departments. Hence, we focus our analysis on these courses.

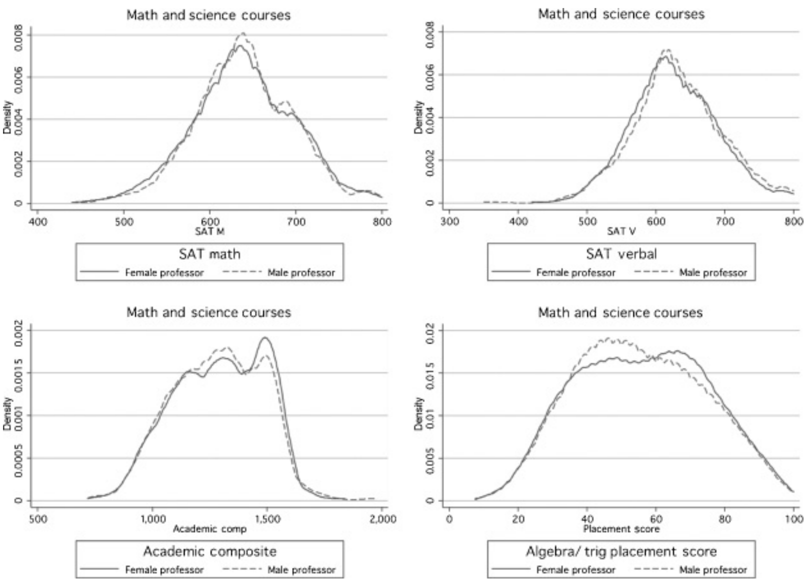

FIGURE II
Math and Science Courses: Distribution of Female Student Pretreatment Characteristics by Professor Gender

Figures represent the distribution of pre-Academy characteristics for female students by professor gender for the USAFA graduating classes of 2001–2008.

and 21% of them were female. Figure II shows the distribution of female student pretreatment characteristics by professor gender.

II.B. Student Assignment to Courses and Professors

Prior to the beginning of the freshman year, students take placement exams in mathematics, chemistry, and select foreign languages, and the scores on these exams are used to place students in the appropriate beginning core courses (i.e., remedial math, Calculus I, Calculus II, etc.). Conditional on course placement, the USAFA Registrar randomly assigns students to core course sections.[11] Thus, throughout their four years of study,

11. The USAFA Registrar employs a stratified random assignment algorithm to place students into sections within each course and semester. The algorithm first assigns all female students evenly throughout all offered sections, then places male recruited athletes, and then assigns all remaining students. Within each group (i.e., female, male athlete, and all remaining males), assignments are random. The one exception is introductory chemistry, where the 92 lowest-ability freshman students each year are ability grouped into four separate sections and are taught by the most experienced professors. Our results are not sensitive to the exclusion of these sections.

students have no ability to choose their required core course professors. Because faculty members teaching the same course use an identical syllabus and give the same exams during a common testing period, grades in core courses are a consistent measure of relative achievement across all students.[12] These institutional characteristics ensure that there is no self-selection of students into (or out of) courses or toward certain professors.

Table II indicates that the types of students assigned to female faculty are nearly indistinguishable from those assigned to male faculty. In math and science courses, the average class size for female faculty is 19.2 compared to 19.0 for males. In addition, male and female professors have a similar numbers of female students per section, and similar average scores on SAT verbal, SAT math, academic composite, and algebra/trigonometry tests.

To formally test whether course assignment is random with respect to faculty gender, we regressed faculty gender on individual student characteristics. The results of this analysis are shown in Table III. Panel A shows results for math and science courses, and Panel B shows results for humanities courses. Across all subgroups we see that the correlation between faculty gender and student characteristics is generally small and statistically nonsignificant. For each specification, we calculated the joint significance of all individual covariates and found these to be nonsignificant in fifteen of the sixteen estimates. Additionally, in Carrell and West (2010), we showed that student assignment to core courses at USAFA is random with respect to peer characteristics and faculty academic rank, experience, and terminal degree status. In that paper, we used resampling methods to construct 10,000 sections drawn from the relevant course and semester and found that the distribution of academic ability by assigned section is indistinguishable from the distribution observed in the resampled sections. Results from these analyses indicate that the algorithm that assigns students to course sections is consistent with random assignment.

12. The one exception is that in some core courses at USAFA, 5% to 10% of the overall course grade is earned by professor-/section-specific quizzes and/or class participation. Among the introductory courses we examine in this study, grades in calculus were not based on any professor-specific assignments between 2000 and 2007. Introductory physics professors were allowed to establish 5% of the course grade and introductory chemistry professors were allowed to establish 4% of the course grade. The introductory course effects we find do not vary significantly across math, chemistry, and physics courses; hence, we believe that the subjective portion of course grades has very little influence on our estimates.

TABLE III
RANDOMNESS CHECK REGRESSIONS OF FACULTY GENDER ON STUDENT CHARACTERISTICS

	All students		SAT math ≦ 660 (median)		SAT math > 660 (median)		SAT math > 700 (75th pctile)	
Specification	Male & female (1)	Female (2)	Male & female (3)	Female (4)	Male & female (5)	Female (6)	Male & female (7)	Female (8)
Panel A. Math and science courses								
Female student	0.003	NA	0.005	NA	−0.001	NA	0.022	NA
	(0.008)		(0.008)		(0.012)		(0.023)	
SAT verbal	−0.005	−0.019	0.002	−0.003	−0.01	−0.046**	−0.019	−0.038
	(0.006)	(0.014)	(0.008)	(0.018)	(0.008)	(0.020)	(0.011)	(0.026)
SAT math	−0.001	−0.008	−0.003	−0.026	−0.009	0.059	−0.041	−0.038
	(0.009)	(0.016)	(0.014)	(0.030)	(0.016)	(0.042)	(0.030)	(0.090)
Academic composite	0.231	0.321	0.512	0.743	−0.256	−0.514	−0.253	−1.921*
	(0.262)	(0.450)	(0.356)	(0.579)	(0.303)	(0.648)	(0.413)	(1.055)
Algebra/trig placement	0.068	0.083	0.06	0.061	0.07	0.103	0.063	−0.016
	(0.064)	(0.074)	(0.063)	(0.073)	(0.075)	(0.102)	(0.087)	(0.175)
Observations	23,056	3,963	13,861	2,721	9,195	1,242	4,046	489
P-value: Joint significance of all individual covariates	.626	.210	.714	.676	.419	.135	.684	.021

TABLE III
(CONTINUED)

Specification	All students: Male & female (1)	All students: Female (2)	SAT math ≦ 660 (median): Male & female (3)	SAT math ≦ 660 (median): Female (4)	SAT math > 660 (median): Male & female (5)	SAT math > 660 (median): Female (6)	SAT math > 700 (75th pctile): Male & female (7)	SAT math > 700 (75th pctile): Female (8)
Panel B. Humanities courses								
Female student	0.011	NA	0.019*	NA	−0.002	NA	0.002	NA
	(0.009)		(0.010)		(0.014)		(0.021)	
SAT verbal	−0.008	−0.051**	−0.016	−0.044**	−0.002	−0.057**	−0.02	−0.007
	(0.009)	(0.019)	(0.011)	(0.021)	(0.009)	(0.025)	(0.014)	(0.031)
SAT math	0.007	−0.003	−0.004	0.008	0.003	0	−0.02	−0.032
	(0.007)	(0.018)	(0.013)	(0.023)	(0.012)	(0.036)	(0.019)	(0.073)
Academic composite	0.372	0.710*	0.292	0.85	0.525	0.678	0.613	0.55
	(0.289)	(0.388)	(0.319)	(0.532)	(0.390)	(0.889)	(0.535)	(1.155)
Algebra/trig placement	0.007	0.081	0.02	0.037	−0.008	0.158	0.041	−0.016
	(0.024)	(0.068)	(0.031)	(0.071)	(0.029)	(0.103)	(0.048)	(0.171)
Observations	15,044	2,438	8,071	1,560	6,973	878	3,396	380
p-value: Joint significance of all individual covariates	.362	.145	.116	.245	.731	.441	.797	.223

Notes. Each specification represents results for a regression where the dependent variable is an indicator variable for female faculty. The SAT verbal, SAT math, academic composite, and algebra/trig placement variables were divided by 100 prior to running the regression. For brevity, coefficients for indicators for black, Hispanic, Asian, recruited athlete, and attended a preparatory school are not shown. Standard errors are clustered at the professor level.

*Significant at the .10 level. **Significant at the .05 level. ***Significant at the .01 level.

III. Statistical Methods

We begin by estimating the following linear regression model:

$$(1)\qquad Y_{icjst} = \phi_1 + \beta_1 F_i + \beta_2 F_j + \beta_3 F_i F_j + \phi_2 X_{icst} + \phi_3 P_j + \gamma_{ct} + \epsilon_{icjst},$$

where Y_{icjst} is the outcome measure for student i in course c with professor j in section s in semester–year t, F_i is an indicator for whether student i is female, and F_j is an indicator for whether professor j is female. The β coefficients are the primary coefficients of interest in our study. β_1 represents the difference in mean performance between female and male students. β_2 is the value added by having a female professor, and β_3 indicates the extent to which having a female professor differentially affects female versus male students. Because students are randomly assigned, estimates of the β coefficients are unbiased.

The vector X_{icst} includes the following student characteristics: SAT math and SAT verbal test scores, academic and leadership composites, algebra/trigonometry placement test score, fitness score, race, whether the student was recruited as an athlete, and whether he or she attended a military preparatory school. We also include cohort dummies. P_j is a vector of professor characteristics including indicators of the professor's academic rank, teaching experience, and terminal degree status. γ_{ct} are course by semester–year fixed effects, which control for unobserved mean differences in academic achievement or grading standards across courses and time. The inclusion of these fixed effects ensures that the model identifies professor quality using only the within-course by semester–year variation in student achievement. We also include course and time-of-day fixed effects. ϵ_{icjst} is the error term. Standard error estimates are clustered by professor.

We implement a slightly modified version of equation (1) to estimate the effect of professor gender in initial courses on performance in follow-on related courses:

$$(2)\qquad Y_{ic's't'} = \phi_1 + \beta_1 F_i + (\beta_2 + \beta_3 F_i)\frac{\sum_{j|i} F_{jt}}{n_{it}} + \phi_2 X_{icst} + \gamma_{c's't'} + \epsilon_{ic's't'},$$

where $Y_{ic'ks't'}$ is performance in the follow-on course c' in section s' and semester–year t'. $\sum_{j|i} F_{jt'}/n_{it'}$ is the proportion of introductory course faculty j who were female for student i at time t'. Including this variable allows us to measure the average impact of having more female professors in introductory math and science courses. We have also estimated regressions in which we include separate variables indicating each introductory course professor's gender. In principle, this specification should allow us to separately identify the effects of introductory math versus chemistry versus physics professors, but in practice the estimated coefficients on the separate indicator variables are too noisy to identify differential effects. The proportion of female professors teaching the students' introductory courses efficiently summarizes the interesting variation. To adjust for any possible professor, section, or year effects in the follow-on course, we include a section by course by semester–year fixed effect, $\gamma_{c's't'}$. As in equation (1), we are primarily interested in the βs, which measure the average differences across male and female students, the effect of having more female professors in the introductory STEM courses, and the differential effect across male and female students of being assigned more female professors in introductory courses. Because students are rerandomized into the mandatory follow-on course sections, estimates of the β coefficients are again unbiased.

To estimate the effect of professor gender on longer-term outcomes, such as choosing to take higher-level math or graduating with a technical degree, we estimate a variation of equation (2),

$$D_{it'} = \phi_1 + \beta_1 F_i + (\beta_2 + \beta_3 F_i)\frac{\sum_{j|i} F_{jt}}{n_{it}} + \phi_2 X_{it} + \epsilon_{it'}, \tag{3}$$

where $D_{it'}$ is a dummy variable that indicates whether student i at time t' chose to take a higher-level math course or chose a STEM major. As before, the β coefficients are the coefficients of interest.

IV. Estimated Effects of Introductory Course Professor Gender in Science and Math Classes

IV.A. Estimated Effects on Course Performance in the Professor's Own Course

Figure III provides unconditional mean estimates by student and professor gender. The pattern of estimates shown in the figure are quantitatively and qualitatively similar to those produced by equation (1), which include all of the covariates discussed in the previous section and are shown in Table IV. The first two columns of Table IV show the estimated effects for all students, whereas

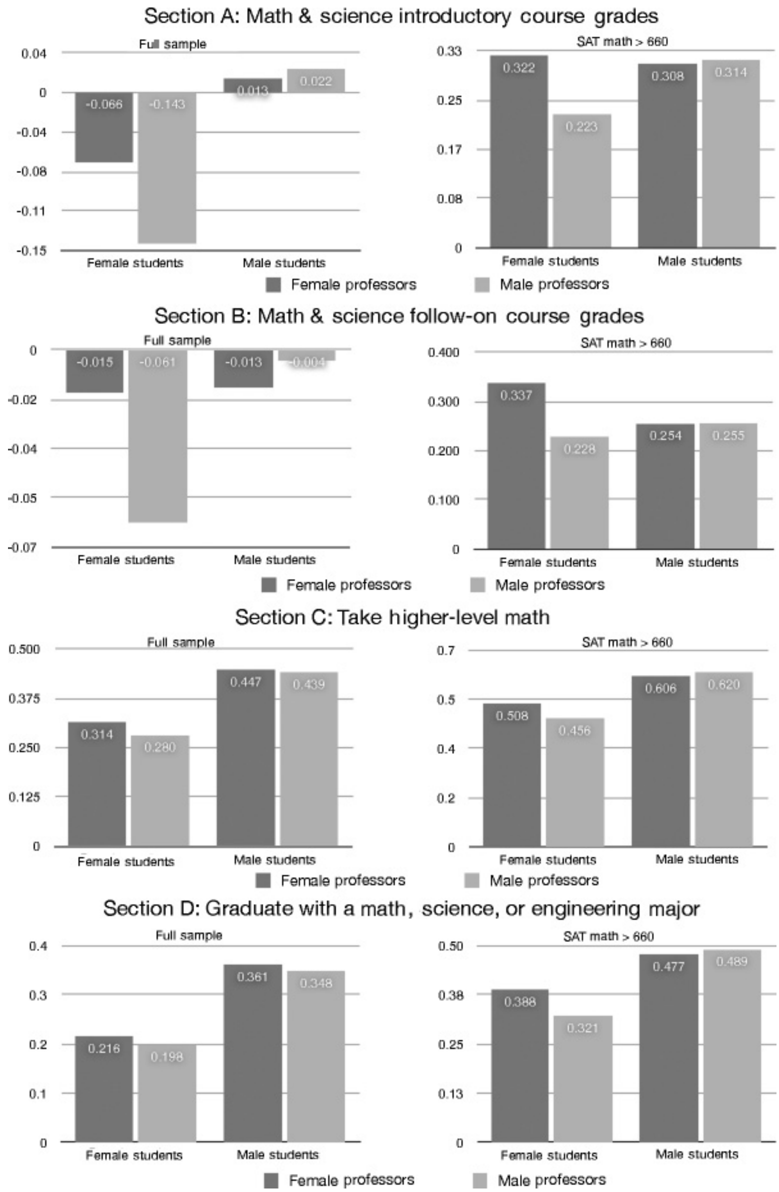

FIGURE III
Unconditional Mean Performance by Student and Professor Gender
Data for the USAFA graduating classes of 2001–2008.

TABLE IV
MATH AND SCIENCE INTRODUCTORY COURSE PROFESSOR GENDER EFFECTS ON INITIAL COURSE PERFORMANCE

	All students		SAT math ≦ 660 (median)		SAT math > 660 (median)		SAT math > 700 (75th pctile)	
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female professor	−0.050*	−0.043**	−0.050*	−0.051**	−0.052	−0.055	−0.028	−0.029
	(0.028)	(0.020)	(0.029)	(0.022)	(0.033)	(0.047)	(0.036)	(0.057)
Female student	−0.149***	NA	−0.147***	NA	−0.153***	NA	−0.162***	NA
	(0.021)		(0.026)		(0.032)		(0.043)	
Female student × female	0.097**	0.139***	0.086*	0.156***	0.115	0.080	0.172**	0.170**
professor	(0.044)	(0.034)	(0.046)	(0.036)	(0.074)	(0.058)	(0.079)	(0.068)
Individual fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	22,956	23,127	13,778	13,889	9,178	9,238	4,043	4,077
Dependent variable mean/	−0.122		−0.291		0.247		0.420	
std dev (female students)	(1.018)		(1.014)		(0.925)		(0.891)	
Dependent variable mean/	0.026		−0.186		0.321		0.502	
std dev (male students)	(0.994)		(0.984)		(0.929)		(0.846)	

Notes. The dependent variable in all specifications is the normalized grade in the course. Robust standard errors in parentheses are clustered by instructor. Control variables: Course by semester fixed effects, graduation class fixed effects, and course time of day fixed fixed effects. Individual-level SAT verbal, SAT math, academic composite, leadership composite, fitness score, algebra/trig placement score and indicator variables for students who are black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school. Introductory course professor-level academic rank dummies, teaching experience, and terminal degree status dummy.

*Significant at the .10 level. **Significant at the .05 level. ***Significant at the .01 level.

the remaining columns focus on subsets of students with different math skills. We include detailed student-level control variables in column (1); column (2) replaces the control variables with individual-student fixed effects.

For the full sample, our estimates on the female-faculty dummy variable indicate that when male students are taught by female professors they end up with somewhat lower course grades than when they are taught by males.[13] The coefficient on the female-professor dummy is between −0.05 (column (1)) and −0.06 (column (2)), which suggests that female professors lower male students' course grades by about 5% to 6% of a standard deviation. The magnitude of the teacher gender effects is swamped, however, by the estimated coefficient on the female student dummy (column (1), row (2)), which indicates that women, on average, score 15% of a standard deviation lower than men *whose math skills were comparable upon entry into the USAFA* when assigned male professors. Given that we are controlling for initial skills, this is a dramatic discrepancy, which can only be documented because of the randomized nature of our study. In most university settings, the possibility of differential selection into courses would make it impossible to detect this phenomenon.

The third row of Table IV displays the estimated coefficient on the female student × female professor interaction. Focusing first on column (1), we see that the estimate is of substantial magnitude (10% of a standard deviation) and positive, indicating that female students' performance in math and science courses improves substantially when the course is taught by a female professor. In fact, taken together with the estimates in rows (1) and (2), the estimated coefficient on the interaction term suggests that having a female professor reduces the gender gap in course grades by approximately two-thirds. This finding reflects both the fact that male students do worse when they have a female professor, and the fact that female students do significantly better. The absolute gain to women from having a female professor is 5% of a standard deviation (−0.050 + 0.097).

The estimates shown in column (1) are based on regressions that control for observable proxies of ability and provide information about the relative gains to men and women from having a male versus female professor in first-year math and science

13. The estimated effect is not statistically significant across all of the subsamples indicated in columns (3)–(6) or across all of the specifications that we use in our robustness analyses.

classes. The next column replaces the student control variables with a student fixed effect. In this regression, the coefficient on the interaction term indicates how much better female students do when they have female professors, compared to their own performance in other mandatory first-year math and science courses. When the estimated coefficients on the female-professor dummy and interaction term are added together $(-0.043 + 0.139)$, the resulting estimate indicates that, conditional on proxies of own ability, female students' performance improves by nearly 10% of a standard deviation.

Columns (3)–(8) focus on subgroups of women defined according to their observed math skills at the time they entered college. Columns (3) and (4) show the regression estimates for students whose SAT math scores were below 660, columns (5) and (6) show the regression estimates for students whose math SATs were above 660, and columns (7) and (8) show the same results for students who scored above 700. These scores correspond to the median and seventy-fifth percentile of the distribution at USAFA, and to the ninetieth and ninety-fifth percentiles of the national SAT math distribution. Because we control for initial SAT math scores and math placement test scores in our regressions, this is unlikely to reflect men's higher likelihood of scoring at the very top of the distribution prior to college. Rather, it suggests that either (1) there are gender differences in math/science ability that are not captured by the initial controls, or (2) something about the college experience has a particularly detrimental effect on the math and science performance of highly skilled women.

The most striking pattern in Table IV is that *as female students' initial math skills increase, the relative importance of professor gender also increases*. In fact, at the top of the distribution (column (7)), having a female professor completely closes the gender gap $(-0.162 + 0.172)$. Notably, at higher skill levels, the evidence that professor gender matters to male students also weakens. We speculate that something about the classroom environment created by female math and science professors has a powerful effect on the performance of women with very strong math skills—with virtually no expense incurred by their comparable male peers. This result is particularly relevant because men and women with high math ability are precisely those needed in the STEM labor market.[14]

14. The improvements in initial course grade are unlikely to result from female instructors engaging in preferential treatment. In the math courses that we

Our estimates are robust to changes in specification that allow the correlation between student characteristics and course grades to vary with student gender. They are also insensitive to the inclusion of interactions between the professor-gender dummy and professor characteristics, and to the inclusion of interactions between the student-gender dummy and the professor-level control variables. The results will be discussed further in Section VI.

We have also extended our analyses to include a full set of professor gender indicators, one for each of the three introductory math and science courses, plus interactions between these indicators and the student gender dummy. The magnitudes of the effects are larger for mathematics but not significantly different from those for chemistry and physics. We also examined and found no evidence for spillover effects across the introductory courses. For example, students' introductory math course grades are affected by the gender of their math professor but not by the gender of their introductory physics or chemistry professor. Similarly, introductory chemistry and physics grades are only affected by the gender of the chemistry or physics professor and not the genders of the professor teaching the other introductory math/science subjects. Results from this analysis are available in Table 1 in the Online Appendix.

IV.B. Longer-Term Effects of Professor Gender

Our main finding is that female students perform substantially better in their math and science courses when they are taught by a woman. Because we are interested in understanding why the prevalence of women in science careers is lower than that of men, our next task is to examine whether these effects persist to longer-term outcomes; course performance itself is only interesting to the extent that it affects pathways into STEM careers. Table V provides the results from estimating the effect of professor gender, measured by the proportion of introductory courses taught by female faculty, on longer-term outcomes. We look at four outcomes: whether the student withdraws from the USAFA, the

study, all exams are graded by a team of faculty and these grades form the basis of the course grade. In all courses, the final grade-cut lines are not determined by the faculty member. To formally test this, we obtained the percentage of points earned in the course for a two-thirds subset of our data. We found nearly identical results when using these continuous data rather than the categorical data. For example, the magnitude of the female student×female professor interaction variable for the highest-ability students (Table IV, column (7)) is 0.172 for the categorical data and 0.192 for the continuous data.

TABLE V
MATH AND SCIENCE INTRODUCTORY COURSE PROFESSOR GENDER EFFECTS ON LONGER-TERM OUTCOMES

Specification	Follow-on STEM course performance (1)	Withdraw in first 2 years (2)	Take higher-level math (3)	Graduate with STEM degree[a] (4)	(5)
	Panel A. All students				
Proportion of professors female	−0.048*	0.008	0.001	0.022	0.010
(introductory courses)	(0.027)	(0.015)	(0.019)	(0.019)	(0.019)
Female student	−0.046**	−0.000	−0.140***	−0.032*	−0.136***
	(0.022)	(0.013)	(0.017)	(0.017)	(0.016)
Female student × proportion of professors female	0.032	−0.049	0.078*	0.030	0.032
	(0.062)	(0.036)	(0.045)	(0.047)	(0.046)
Observations	58,929	8,851	8,851	8,851	8,851
Dependent variable mean/std dev	−0.021	0.140	0.350	0.412	0.247
(female students)	(0.976)	(0.347)	(0.477)	(0.492)	(0.431)
Dependent variable mean/std dev	0.004	0.150	0.508	0.461	0.407
(male students)	(1.002)	(0.358)	(0.500)	(0.499)	(0.491)
	Panel B. SAT math $\leqq$ 660 (median)				
Proportion of professors female	−0.001	0.024	0.050*	0.053*	0.064**
(introductory courses)	(0.041)	(0.024)	(0.028)	(0.029)	(0.027)
Female student	−0.034	−0.005	−0.118***	−0.010	−0.099***
	(0.030)	(0.019)	(0.022)	(0.023)	(0.021)
Female student × proportion of professors female	−0.070	−0.025	0.019	−0.071	−0.086
	(0.089)	(0.053)	(0.063)	(0.065)	(0.060)
Observations	31,517	4,673	4,673	4,673	4,673

TABLE V
(CONTINUED)

Specification	Follow-on STEM course performance (1)	Withdraw in first 2 years (2)	Take higher-level math (3)	Graduate with STEM degree[a] (4)	Graduate with STEM degree[a] (5)
Dependent variable mean/std dev	−0.228	0.159	0.241	0.314	0.161
(female students)	(0.948)	(0.366)	(0.428)	(0.464)	(0.368)
Dependent variable mean/std dev	−0.246	0.169	0.350	0.335	0.281
(male students)	(0.975)	(0.375)	(0.477)	(0.472)	(0.450)
Panel C. SAT math > 660 (median)					
Proportion of professors female	−0.080**	0.002	−0.030	0.003	−0.028
(introductory courses)	(0.033)	(0.019)	(0.025)	(0.026)	(0.026)
Female student	−0.065*	0.006	−0.169***	−0.057**	−0.179***
	(0.032)	(0.019)	(0.026)	(0.027)	(0.027)
Female student × proportion of professors female	0.157**	−0.080	0.136**	0.140**	0.155**
	(0.080)	(0.050)	(0.066)	(0.070)	(0.070)
Observations	27,414	4,178	4,178	4,178	4,178
Dependent variable mean/std dev	0.315	0.109	0.526	0.569	0.384
(female students)	(0.925)	(0.312)	(0.500)	(0.496)	(0.487)
Dependent variable mean/std dev	0.268	0.131	0.670	0.589	0.535
(male students)	(0.961)	(0.338)	(0.470)	(0.492)	(0.499)

TABLE V
(CONTINUED)

Specification	Follow-on STEM course performance (1)	Withdraw in first 2 years (2)	Take higher-level math (3)	Graduate with STEM degree[a] (4)	(5)
	Panel D. SAT math > 700 (75th pctile)				
Proportion of professors female	−0.104***	−0.010	−0.018	0.036	0.021
(introductory courses)	(0.041)	(0.025)	(0.033)	(0.036)	(0.037)
Female student	−0.104**	0.029	−0.235***	−0.071*	−0.265***
	(0.045)	(0.029)	(0.037)	(0.041)	(0.042)
Female student × proportion of professors female	0.228**	−0.096	0.193**	0.110	0.258***
	(0.102)	(0.069)	(0.090)	(0.099)	(0.101)
Observations	13,110	2,040	2,040	2,040	2,040
Dependent variable mean/std dev	0.462	0.116	0.564	0.610	0.398
(female students)	(0.879)	(0.321)	(0.497)	(0.489)	(0.490)
Dependent variable mean/std dev	0.429	0.118	0.7498	0.648	0.600
(male students)	(0.920)	(0.323)	(0.434)	(0.478)	(0.490)

Notes. Robust standard errors in parentheses are clustered by student in Specification 2. Control variables: Graduation class fixed effects. Individual-level SAT verbal, SAT math, academic composite, leadership composite, fitness score, algebra/trig placement score and indicator variables for students who are black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school. Introductory course proportion of professors who are associate or full professors, mean teaching experience, and proportion with a terminal degree. For Specification 2 we also include course by semester by section fixed effects.

[a]Specification 5 excludes biological sciences.

*Significant at the .10 level. **Significant at the .05 level. ***Significant at the .01 level.

student's performance in all required follow-on STEM coursework, whether the student chooses to take higher-level math courses beyond those that are required for graduation with a non-STEM degree, and whether she graduates with a STEM degree.[15] All four of these outcomes are correlated with future career choices. Beginning with the top panel, column (2) shows that, conditional on entering math skills, women and men are equally likely to withdraw from the USAFA. However, female students perform significantly worse in follow-on STEM coursework, are less likely to take higher-level math courses, and are less likely to graduate with a STEM degree than male students. It is also clear that gender differences in college major are much larger when we exclude biological sciences (column (5) versus (4)), which typically require less math, and have higher rates of female participation.[16]

The estimated effect of professor gender on these long-term outcomes varies across the sub-samples, with the biggest effects, by far, accruing to women with high entering math ability. Across the full sample, there is no statistically significant evidence that having a higher proportion of female professors affects a woman's likelihood of withdrawing, her performance in follow-on coursework, her probability of taking higher-level math courses, or her probability of graduating with a STEM major. Similar results are shown in Panel B, where we focus on the subgroup of women whose math SAT scores were below the median. However, as the sample narrows to include increasingly high-skilled women (as approximated by their SAT math scores), the estimated effects of professor gender become much larger and statistically significant. Among the top quartile of female students, and for each long-term outcome, higher proportions of female professors in introductory math and science courses are associated with reductions in the gender gap. In fact, the estimates suggest that increasing the fraction of female professors from 0% to 100% would completely eliminate the gender gap in math and science majors. For example, column (5) of Panel C indicates that among the highest-ability women, those whose introductory math and science professors are exclusively female are twenty-six percentage points more likely to major in STEM than those who are exclusively assigned to male

15. The attrition results we present in Table V show attrition after the second year; however, results are qualitatively similar for one-year and four-year attrition. See Carrell and West (2010, Table I) for a list of the required follow-on coursework.

16. We find qualitatively similar results when we also exclude environmental engineering, a field with a relatively higher rate of female participation.

faculty. For this high ability group, the male/female gap in the probability of completing a STEM major is 27 percentage points.

At the same time, there is no evidence that having a female professor affects a female student's likelihood of dropping out, regardless of her ability level. This suggests that whatever it is about female professors that affects women in their first-year math and science courses, it is not something that changes retention rates but rather something that changes their preferences for math and science. This interpretation is consistent with Zafar (2009), who finds evidence at Northwestern University that the gender gap in academic major is "due to differences in beliefs about enjoying coursework and differences in preferences." Hence, our findings suggest that female professors may be changing female student's beliefs and preferences toward STEM coursework and careers. We have also estimated regressions in which we include three separate dummy variables indicating each introductory course professor's gender. This allows us to investigate the possibility that our estimated long-run effects are driven by professor gender in a particular course.[17] We find little evidence that our long-run estimates are driven by professor gender in a particular subject or that professor gender in the same previous subject is more important than professor gender in "cross" subjects.[18]

Our findings are robust to changes in model specification that exclude individual controls or that increase model flexibility by including interactions between individual characteristics and student gender. They are not generated by a few outliers: when we estimate teacher value-added for each professor and plot the effects by professor and student gender we find that among female professors over two-thirds of the value-added shrinkage estimates are positive for their high-ability female students.[19]

IV.C. Estimated Effects of Professor Gender in English and History Classes

Next, we consider the role of professor gender in humanities courses. Table VI shows the estimated effects of professor

17. The results from this analysis can be found in Table 1, Panel B, in the Online Appendix.

18. We find one exception. Among women with SAT math scores greater than 700, we find that the effects of professor gender on graduating with a STEM degree and taking higher-level math are significantly greater for calculus professors compared to chemistry or physics professors.

19. See Section V for details of how we calculated the value-added estimates. Figure IV shows plots of the value-added shrinkage estimates by student and professor gender.

TABLE VI
ENGLISH AND HISTORY INTRODUCTORY COURSE PROFESSOR GENDER EFFECTS

	Initial course performance (1)	Follow-on course performance (2)	Take higher-level humanities (3)	Graduate with humanities degree (4)	Take higher-level math (5)	Graduate with STEM degree[a] (6)
Panel A. All students						
Proportion of professors female	−0.113*	−0.008	−0.016	−0.002	0.008	−0.007
(introductory courses)	(0.064)	(0.038)	(0.018)	(0.012)	(0.019)	(0.020)
Female student	−0.018	0.037	0.020	0.019**	−0.110***	−0.123***
	(0.036)	(0.025)	(0.014)	(0.009)	(0.015)	(0.015)
Female student × proportion	0.028	−0.098	0.019	−0.009	−0.087*	−0.043
of professors female	(0.074)	(0.078)	(0.042)	(0.027)	(0.045)	(0.045)
Observations	15,044	13,661	8,720	8,720	8,720	8,720
Panel B. SAT math $\leqq$ 660 (median)						
Proportion of professors female	−0.107	−0.014	−0.041	0.004	0.028	−0.076
(introductory courses)	(0.065)	(0.050)	(0.028)	(0.018)	(0.027)	(0.055)
Female student	−0.064	−0.018	0.026	0.019	−0.094***	−0.109***
	(0.040)	(0.032)	(0.019)	(0.012)	(0.019)	(0.018)
Female student × proportion	−0.009	−0.027	0.015	−0.0004	−0.122**	0.076
of professors female	(0.008)	(0.102)	(0.059)	(0.038)	(0.058)	(0.055)
Observations	8,071	7,244	4,619	4,619	4,619	4,619

TABLE VI
(CONTINUED)

	Initial course performance (1)	Follow-on course performance (2)	Take higher-level humanities (3)	Graduate with humanities degree (4)	Take higher-level math (5)	Graduate with STEM degree[a] (6)
Panel C. SAT math > 660 (median)						
Proportion of professors female	−0.115*	−0.028	0.001	−0.009	−0.016	−0.024
(introductory courses)	(0.068)	(0.054)	(0.023)	(0.015)	(0.028)	(0.030)
Female student	0.050	0.077*	0.011	0.019	−0.135***	−0.145***
	(0.045)	(0.041)	(0.019)	(0.013)	(0.023)	(0.025)
Female student × proportion	0.086	−0.102	0.047	−0.024	−0.048	−0.005
of professors female	(0.083)	(0.122)	(0.058)	(0.038)	(0.070)	(0.075)
Observations	6,973	6,417	4,101	4,101	4,101	4,101
Panel D. SAT math > 700 (75th pctile)						
Proportion of professors female	−0.101)	−0.087	0.005	−0.007	−0.017	−0.026
(introductory courses)	(0.073)	(0.073)	(0.030)	(0.019)	(0.037)	(0.042)
Female student	0.021	0.110*	0.007	−0.0003	−0.187***	−0.209***
	(0.052)	(0.061)	(0.027)	(0.017)	(0.034)	(0.038)
Female student × proportion	0.083	−0.137	0.101	0.048	−0.026	0.038
of professors female	(0.104)	(0.177)	(0.079)	(0.048)	(0.097)	(0.109)
Observations	3,396	3,155	1,997	1,997	1,997	1,997

Notes. For Specification (1) standard errors are clustered by professor. For Specification (2) standard errors are clustered by student. Control variables: Graduation class fixed effects. Individual-level SAT verbal, SAT math, academic composite, leadership composite, fitness score, algebra/trig placement score, and indicator variables for students who are black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school. For Specifications (1) and (2) we control for the academic rank, teaching experience, and terminal degree status of the professor. For Specifications (3)–(5) we control for the introductory course proportion of professors who are associate or full professors, mean teaching experience, and proportion with a terminal degree. For Specification (1) we include a course by semester fixed effect. For Specification (2) we include course by semester by section fixed effects.

[a]Excludes biological sciences.

*Significant at the .10 level. **Significant at the .05 level. ***Significant at the .01 level.

gender when we estimate equation (1) for introductory English and history courses. The estimates are strikingly different. There is no observable gender gap in course performance, and there is no evidence that female students' course grades are improved when they have a female professor. As in Tables IV and V, we find weak evidence that both men and women have lower humanities grades when the course is taught by a female professor, but most of the coefficient estimates on the female professor dummy are barely significant at the 10% level.[20] Specifications 3–6 carry forward our analyses for longer-term outcomes. We look at the effect of professor gender in initial humanities courses on later course selection and choice of major. All of the estimated female professor coefficients are small, and none is statistically significant. This indicates that the gender of professors in initial humanities courses has no effect on male students' longer-term choices. Similarly, most of the estimated coefficients on the interaction term are small, and only one is statistically different from zero, suggesting that female students' long-run choices are also unrelated to the sex of the professors who teach their humanities courses.

These results stand in direct contrast to our estimated professor effects in math and science, where it appears that female students with strong math skills are powerfully affected by the gender of their introductory course professors. These results also indicate the effects we find are not likely driven by the general (military) culture of the institution we study. In the next section, we explore mechanisms that might be behind this effect.

IV.D. Contemporaneous Effects of Professor Gender in Follow-On Courses

We have seen evidence that female students' paths into math and science careers are influenced significantly by the gender of the professors who teach their *introductory* math and science courses. Next, we examine how the gender of professors in more advanced follow-on math and science courses affects

20. We have also estimated individual student fixed effects model analogous to the specification that is employed in columns (2), (4), (6), and (8) of Table IV. The results from this specification suggest that when male students are taught by women in introductory humanities courses, their grades are about 20% of a standard deviation lower. Because we observe this effect only for male students with one professor of each gender (19% of sample), any sort of grade discrimination on the part of professors is not driving the effect. Rather, the result is consistent with a story of effort/response on the part of male students who have this very specific treatment. Among female students, course performance seems to be unrelated to professor gender. Results are available upon request from the authors.

contemporaneous student STEM outcomes.[21] Results in Table VII show negligible effects of professor gender in mandatory follow-on math and science courses on (contemporaneous) course grades, whether the student takes higher-level math, and whether the student graduates with a degree in STEM. We find that none of the estimated interaction terms is statistically different from zero, most are small in magnitude, and a few are in the opposite direction from our earlier estimates. Because these courses are taken later in students' educational paths, the effect of professor gender may be different due either to a mechanical effect (i.e., academic majors may already be chosen) or to the fact that preferences and self-perceptions of student ability may already be formed at this juncture. Nevertheless, these results suggest that the classroom environment has its strongest influence on female students early in the college career.

V. Mechanisms: Is It All about Professor Gender?

Table IV suggests that female students' initial math and science grades are substantially higher when they are taught by female professors. The estimated effects are particularly large among female students in the upper quartile of the SAT math distribution. In this section, we investigate whether gender differences in student performance are driven by professor gender *per se*, or might be driven by some other professor characteristic that is correlated with professor gender. For example, male and female students may respond in different ways to younger versus older professors, or they may have different responses to alternative teaching styles that are correlated with, but not exclusive to, professor gender.

To investigate possible mechanisms further, we conduct three additional analyses. First, we interact all of our professor-level variables with the professor and student gender dummies to see whether the importance of particular professor characteristics varies with student and/or professor gender. The results of these regressions, which are shown in Table VIII, indicate that it is not differences in observables, or differences in student-gender specific responsiveness to those observables, that are driving our results.

21. Specifically, we examine how the gender of the professor teaching mandatory second-semester courses in calculus, chemistry, and physics affects course grades.

TABLE VII
CONTEMPORANEOUS EFFECTS OF PROFESSOR GENDER IN FOLLOW-ON COURSES

Outcome	Course grade	Take higher-level math	Graduate with STEM degree[a]
	Panel A. All students		
Female professor	0.016	0.003	0.009
	(0.028)	(0.008)	(0.009)
Female student	−0.043*	−0.124***	−0.128***
	(0.023)	(0.011)	(0.009)
Female student × female professor	0.019	−0.002	−0.012
	(0.039)	(0.021)	(0.018)
Observations	19,315	19,315	19,315
	Panel B. SAT math ≦ 660 (median)		
Female professor	0.015	−0.001	0.010
	(0.032)	(0.012)	(0.014)
Female student	−0.031	−0.121***	−0.121***
	(0.027)	(0.014)	(0.012)
Female student × female professor	−0.015	0.003	−0.003
	(0.044)	(0.029)	(0.022)
Observations	11,211	11,211	11,211
	Panel C. SAT math > 660 (median)		
Female professor	0.009	0.013	0.012
	(0.030)	(0.012)	(0.013)
Female student	−0.059	−0.122***	−0.134***
	(0.038)	(0.018)	(0.017)
Female student × female professor	0.064	−0.028	−0.058
	(0.068)	(0.048)	(0.042)
Observations	8,104	8,104	8,104
	Panel D. SAT math > 700 (75th pctile)		
Female professor	0.010	0.006	−0.015
	(0.040)	(0.017)	(0.021)
Female student	−0.108**	−0.197***	−0.224***
	(0.052)	(0.032)	(0.034)
Female student × female professor	0.071	0.016	−0.020
	(0.090)	(0.055)	(0.056)
Observations	3,602	3,602	3,602

Notes. The dependent variable in all specifications is the normalized grade in the course. Robust standard errors in parentheses are clustered by instructor. Control variables: Contemporanous course by semester fixed effects, graduation class fixed effects, and course time of day fixed effects. Introductory course by semester by section fixed effects. Individual-level SAT verbal, SAT math, academic composite, leadership composite, fitness score, algebra/trig placement score, and indicator variables for students who are black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school. Professor-level academic rank dummies, teaching experience, and terminal degree status dummy.

[a]Excludes biological sciences.

*Significant at the .10 level. **Significant at the .05 level. ***Significant at the .01 level.

TABLE VIII
MATH AND SCIENCE INTRODUCTORY COURSE PROFESSOR GENDER EFFECTS WITH ADDITIONAL CONTROLS

Outcome	Introductory course performance (1)	Follow-on STEM course performance (2)	Withdraw in first 2 years (3)	Take higher-level math (4)	Graduate with STEM degree[a] (5)	Graduate with STEM degree[a] (6)
Panel A. All students						
Female student × proportion of professors female	0.098**	0.038	−0.051	0.076*	0.036	0.047
	(0.040)	(0.064)	(0.037)	(0.046)	(0.048)	(0.047)
Observations	22,956	58,929	8,851	8,851	8,851	8,851
Dependent variable mean/std dev	−0.122	−0.021	0.140	0.350	0.412	0.247
(female students)	(1.018)	(0.976)	(0.347)	(0.477)	(0.492)	(0.431)
Dependent variable mean/std dev	0.026	0.004	0.150	0.508	0.461	0.407
(male students)	(0.994)	(1.002)	(0.358)	(0.500)	(0.499)	(0.491)
Panel B. SAT math $\leqq$ 660 (median)						
Female student × proportion of professors female	0.080*	−0.072	−0.022	0.015	−0.069	−0.066
	(0.044)	(0.090)	(0.055)	(0.064)	(0.066)	(0.061)
Observations	13,778	31,517	4,673	4,673	4,673	4,673
Dependent variable mean/std dev	−0.291	−0.228	0.159	0.241	0.314	0.161
(female students)	(1.014)	(0.948)	(0.366)	(0.428)	(0.464)	(0.368)
Dependent variable mean/std dev	−0.186	−0.246	0.169	0.350	0.335	0.281
(male students)	(0.984)	(0.975)	(0.375)	(0.477)	(0.472)	(0.450)

TABLE VIII
(CONTINUED)

Outcome	Introductory course performance (1)	Follow-on STEM course performance (2)	Withdraw in first 2 years (3)	Take higher-level math (4)	Graduate with STEM degree[a] (5)	Graduate with STEM degree[a] (6)
Panel C. SAT math > 660 (median)						
Female student × proportion of professors female	0.133*	0.204**	−0.082	0.138**	0.140*	0.164**
	(0.070)	(0.081)	(0.051)	(0.069)	(0.072)	(0.073)
Observations	9,178	27,414	4,178	4,178	4,178	4,178
Dependent variable mean/std dev	0.247	0.315	0.109	0.526	0.569	0.384
(female students)	(0.925)	(0.925)	(0.312)	(0.500)	(0.496)	(0.487)
Dependent variable mean/std dev	0.321	0.268	0.131	0.670	0.589	0.535
(male students)	(0.929)	(0.961)	(0.338)	(0.470)	(0.492)	(0.499)
Panel D. SAT math > 700 (75th pctile)						
Female student × proportion of professors female	0.211**	0.274**	−0.094	0.183*	0.073	0.239**
	(0.086)	(0.110)	(0.074)	(0.097)	(0.106)	(0.108)
Observations	4,043	13,110	2,040	2,040	2,040	2,040
Dependent variable mean/std dev	0.420	0.462	0.116	0.564	0.610	0.398
(female students)	(0.891)	(0.879)	(0.321)	(0.497)	(0.489)	(0.490)
Dependent variable mean/std dev	0.502	0.429	0.118	0.7498	0.648	0.600
(male students)	(0.846)	(0.920)	(0.323)	(0.434)	(0.478)	(0.490)

Notes. Robust standard errors in parentheses are clustered by professor in Specification (1) and student in Specification (2). For Specification (1) the female faculty and value-added variables correspond to the professor who taught the initial course. Control variables: Individual-level SAT verbal, SAT math, academic composite, leadership composite, fitness score, algebra/trig placement score and indicator variables for students who are black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school. Introductory course proportion of professors who are associate or full professors, mean teaching experience, and proportion with a terminal degree. Graduation class by gender fixed effects. Interactions between the professor-level variables and student gender. Interactions between the professor-level variables and professor gender. Interactions of student gender and all individual-level controls. For Specification (1) we include course-by-semester fixed effects. For Specification (2) we include course-by-semester-by-section fixed effects.

[a]Excludes biological sciences.

*Significant at the .10 level. **Significant at the .05 level. ***Significant at the .01 level.

Second, we examined the role of voluntary interaction between students and professors outside of formal classroom instruction. To do so, the Mathematics Department at USAFA collected office hour data for each student by professor during the fall of 2008. These data showed that female students were no more likely to attend office hours with female than with male professors.[22] Although the data were from a single course in a single semester, the results suggest that the mechanisms that are driving our estimated effects are not likely driven by gender differences in willingness to approach professors for additional instruction.

Finally, we examine the role of unobservables through a professor "value-added" analysis. This is implemented through a two-step process: first, for each professor and course, we estimate a student gender-specific random effect, which summarizes the professor's average value-added separately for female and for male students.[23] This provides us with estimates of each professor's "value-added" for both female and male students. Figure IV shows the distribution of the gender-specific estimated value-added, $\hat{\xi}$. As expected, the distribution of the female-student–female-teacher effects (middle column) is to the right of the distribution of female-student–male-teacher effects. These results reconfirm our previous finding that, on the average, female students perform better when their math and science courses are taught by female faculty, but also make clear that many male professors are very effective at teaching female students. In other words, student performance in the introductory course is correlated with professor gender, but not exclusively.

22. Female students were much more likely to attend office hours than male students across all professors.

23. We estimate a Bayesian shrinkage estimate for each professor's value-added by student gender in a random effects framework as in Rabe-Hesketh and Skrondal (2008). The shrinkage estimates take into account the variance (signal to noise) and the number of observations for each professor. Because we have random assignment, both random effects and fixed effects models will produce consistent estimates, but random effects models are efficient. To eliminate classroom-specific common shocks, we estimated professor *j*'s value-added in section *s* using professor *j*'s students not in section *s* (i.e., we use sections other than the student's own section). The value-added estimates are based on regressions that control for all variables in equation (1), except for professor gender. In addition, we include interactions between student gender and professor academic rank, experience, and terminal degree status and interactions between student gender and individual-level covariates. The raw correlation between the within-professor male and female student value-added is 0.19. For recent work estimating teacher value-added models see Rivkin, Hanushek, and Kain (2005), Kane, Rockoff, and Staiger (2008), Kane and Staiger (2008), Hoffmann and Oreopoulos (2009), and Carrell and West (2010).

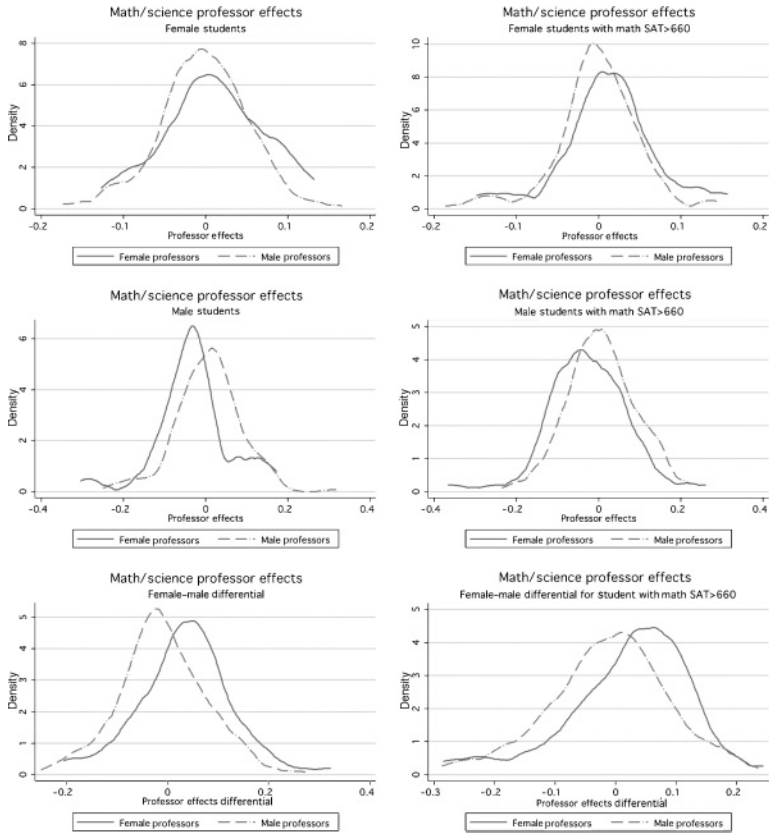

FIGURE IV
Distribution of Professor Value-Added by Student and Professor Gender

Figures represent the distribution of professor value-added estimates (Bayes shrinkage) by student and professor gender in introductory math and science courses for the USAFA graduating classes of 2001–2008.

Our next step is to reestimate the follow-on equations, (2) and (3), while including the average of the estimated professor value-added, $\hat{\xi}$, as explanatory variables:

$$(4) \qquad Y_{ic's't'} = \phi_1 + \beta_1 F_i + \phi_2 X_{icst} + (\beta_2 + \beta_3 F_i)\frac{\sum_{j|i} F_{jt}}{n_{it}} + \beta_4 F_i \frac{\sum_{j|i} \hat{\xi}_{fj}}{n_{it}} + \beta_5 F_i \frac{\sum_{j|i} \hat{\xi}_{mj}}{n_{it}} + \beta_6 M_i \frac{\sum_{j|i} \hat{\xi}_{fj}}{n_{it}} + \beta_7 M_i \frac{\sum_{j|i} \hat{\xi}_{mj}}{n_{it}} + \gamma_{c's't'} + \epsilon_{ic's't'}.$$

M_i is an indicator variable of whether student i is male. This equation allows us to investigate whether students' outcomes are affected by professors who have high "male/female value-added," conditional on professor gender. In other words, we can estimate the impact of professor "quality" separately from the impact of professor gender itself. We present results for this analysis in Table IX. Column (1) shows that both the professor gender and professor "value-added" variables are strong predictors of student performance in the introductory STEM courses. However, results in columns (2)–(4) show that although professor gender continues to exert a positive effect on female student outcomes, the introductory course professor value-added has no predictive power for the longer-term outcomes. As in Carrell and West (2010), we find no persistence of introductory course value-added into follow-on course performance at USAFA. Thus, it appears that the influence of female professors on their female students' future math and science performance operates largely through factors other than value-added in the introductory course grades.

VI. Conclusions

Why aren't there more women in science careers? If we want to know the answer to this question, we need to make sense of what happens to women in college. College is a critical juncture in the life cycle, and in spite of the fact that men and women enter college with similar levels of math preparation, substantially fewer women leave college with a science or engineering degree. This, in turn, closes the door to many careers in science and technology.

The goal of this paper is to shed light on how women's paths toward science are affected by the college environment, focusing on the role of professor gender. Unlike previous research on this topic, we are blessed with experimental conditions that ensure that our estimates are uncontaminated by self-selection and attrition bias. This is possible because the USAFA randomly assigns students to professors over a wide variety of mandatory standardized courses. A further advantage of studying this campus is that course grades are not determined by an individual student's professor.

The nature of our data allows us to document a number of interesting patterns. First, we find that *compared to men with the same entering math ability*, female students perform substantially less well in their introductory math and science courses. To our knowledge, this is the first study that has been able to document

TABLE IX
MATH AND SCIENCE INTRODUCTORY COURSE PROFESSOR VALUE-ADDED EFFECTS BY STUDENT GENDER

	Panel A. All students				Panel B. SAT math $\leqq$ 660 (median)			
	Introductory course performance	Follow-on STEM course performance	Take higher-level math	Graduate with STEM degree	Introductory course performance	Follow-on STEM course performance	Take higher-level math	Graduate with STEM degree
Specification	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
			Female student coefficients					
Female student × proportion of professors female	0.079**	0.029	0.084*	0.041	0.067	−0.060	0.027	−0.082
	(0.038)	(0.063)	(0.046)	(0.046)	(0.043)	(0.089)	(0.063)	(0.060)
Mean initial course professor value-added for female students	0.040**	−0.013	−0.045**	−0.013	0.028	−0.021	−0.047*	−0.015
	(0.019)	(0.025)	(0.020)	(0.020)	(0.022)	(0.034)	(0.026)	(0.025)
Mean initial course professor value-added for male students	0.044***	0.009	0.014	0.024	0.043***	0.032	0.019	0.016
	(0.015)	(0.25)	(0.018)	(0.018)	(0.016)	(0.033)	(0.023)	(0.022)
			Male student coefficients					
Mean proportion of professors female	−0.030**	−0.049*	0.003	0.012	−0.029	−0.006	0.048*	0.063**
	(0.014)	(0.027)	(0.019)	(0.019)	(0.019)	(0.041)	(0.028)	(0.027)
Mean initial course professor value-added for female students	0.013*	0.010	0.004	−0.006	0.011	0.025	0.002	−0.002
	(0.007)	(0.012)	(0.008)	(0.008)	(0.010)	(0.018)	(0.013)	(0.011)
Mean initial course professor value-added for male students	0.080***	−0.016	0.0004	0.001	0.078***	−0.020	−0.006	−0.002
	(0.008)	(0.011)	(0.008)	(0.008)	(0.010)	(0.017)	(0.012)	(0.011)
Observations	22,342	58,493	8,770	8,770	13,433	31,490	4,668	4,668

TABLE IX
(CONTINUED)

	Panel C. SAT math > 660 (median)				Panel D. SAT math > 700 (75th pctile)			
Specification	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	Female student coefficients							
Female student × proportion	0.106	0.137*	0.144**	0.179**	0.157**	0.215**	0.216**	0.294***
of professors female	(0.069)	(0.081)	(0.068)	(0.072)	(0.076)	(0.106)	(0.093)	(0.103)
Mean initial course professor value-	0.062**	−0.017	−0.045	−0.010	0.034	0.020	−0.071*	−0.008
added for female students	(0.029)	(0.033)	(0.030)	(0.032)	(0.035)	(0.047)	(0.041)	(0.045)
Mean initial course professor value-	0.052*	−0.014	0.011	0.038	0.049	−0.023	0.019	0.076
added for male students	(0.029)	(0.033)	(0.028)	(0.029)	(0.040)	(0.046)	(0.039)	(0.043)
	Male student coefficients							
Mean proportion	−0.032	−0.077**	−0.028	−0.026	−0.008	−0.099**	−0.018	0.024
of professors female	(0.021)	(0.033)	(0.025)	(0.027)	(0.028)	(0.041)	(0.033)	(0.037)
Mean initial course professor value-	0.015*	0.001	0.003	−0.013	−0.014	−0.013	0.011	−0.015
added for female students	(0.011)	(0.015)	(0.012)	(0.012)	(0.016)	(0.019)	(0.015)	(0.017)
Mean initial course professor value-	0.081***	−0.010	0.010	0.006	0.090***	0.008	0.009	0.010
added for male students	(0.012)	(0.014)	(0.011)	(0.012)	(0.016)	(0.018)	(0.014)	(0.016)
Observations	8,909	27,003	4,102	4,102	3,912	12,806	1,984	1,984

Notes. Specification (1). The female faculty and value-added variables correspond to the professor who taught the initial course. Standard errors are clustered by professor. Control variables: Course by semester fixed effects, graduation class fixed effects, and course time of day fixed effects. Individual-level SAT verbal, SAT math, academic composite, leadership composite, fitness score, algebra/trig placement score, and indicator variables for students who are black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school. Introductory course professor-level academic rank, teaching experience, and terminal degree status. Specifications (2)–(4). Standard errors are clustered by student. Control variables: Graduation class fixed effects. Individual-level SAT verbal, SAT math, academic composite, leadership composite, fitness score, algebra/trig placement score and indicator variables for students who are black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school. Introductory course proportion of professors who are associate or full professors, mean teaching experience, and proportion with a terminal degree. For Specification (2) we also include course by semester by section fixed effects.

*Significant at the .10 level. **Significant at the .05 level. ***Significant at the .01 level.

this fact—it is only knowable because of the mandatory nature of introductory math and science courses at the USAFA. We document a gender gap in most other dimensions of STEM success, as well. Second, we find that the gender gap is mitigated considerably when female students have female professors.[24] Conversely, professor gender seems to be irrelevant in the humanities. Third, we find that the effect of female professors on female students is largest among students with high math ability. In particular, we find that among students in the upper quartile of the SAT math distribution, being assigned to a female professor eliminates the gender gap in introductory course grades and science majors. We also find that professor gender has minimal effects on male students' outcomes.

This research raises a number of interesting questions about why professor gender is important, particularly among students whose math skills are at the top of the ability distribution. Do female professors serve as role models? Do they teach in ways that female students find more accessible? Are they more encouraging of their female students? We have begun to investigate these questions by looking at the distribution of each professor's gender-specific average value-added. We find that professor value-added is correlated with professor gender, but is not exclusive to it. Additionally, professor gender continues to be a positive predictor of long-term STEM success even when controlling for professor value-added. In future research, we hope to investigate whether there are observable characteristics of male and female teachers that can help explain this phenomenon. Although this is not possible with our current data, it would provide invaluable information to policy makers who seek to improve women's representation in science.

UC Davis and National Bureau of Economic Research
UC Davis and National Bureau of Economic Research
U.S. Air Force Academy

References

Barron, John M., Bradley T. Ewing, and Glen R. Waddell, "The Effects of High School Participation on Education and Labor Market Outcomes," *Review of Economics and Statistics*, 82 (2000), 409–421.

24. Note that the impact of female professors may reflect the high quality of faculty at the USAFA, and that substituting lower-quality female professors for high-quality male professors is not a policy that would be recommended by the authors.

Bettinger, Eric, and Bridget Terry Long, "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students," *American Economic Review*, 95 (2005), 152–157.

Beyer, Sylvia, and Edward M. Bowden, "Gender Differences in Self-Perceptions: Convergent Evidence from Three Measures of Accuracy and Bias," *Personality and Social Psychology Bulletin*, 23 (1997), 157–172.

Brown, Charles, and Mary Corcoran, "Sex-Based Differences in School Content and the Male–Female Wage Gap," *Journal of Labor Economics*, 15 (1997), 431–465.

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

Carrell, Scott E., Richard L. Fullerton, and James E. West, "Does Your Cohort Matter? Estimating Peer Effects in College Achievement," *Journal of Labor Economics*, 27 (2009), 439–464.

Carrell, Scott E., and James E. West, "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," *Journal of Political Economy,* 118 (2010), 409–432.

Eckel, Catherine C., and Philip J. Grossman, "The Difference in the Economic Decisions of Men and Women: Experimental Evidence," in *Handbook of Experimental Economics Results*, Vol. 1, Charles Plott and Vernon Smith, eds. (Amsterdam: Elsevier, 2008).

Elliot, Andrew J., and Judith M. Harackiewicz, "Goal Setting, Achievement Orientation, and Intrinsic Motivation: A Mediational Analysis," *Journal of Personality and Social Psychology*, 66 (1994), 968–980.

Ellison, Glenn, and Ashley Swanson, "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions," NBER Working Paper No. 15238, 2009.

Goldin, Claudia, Lawrence Katz, and Ilyana Kuziemko, "The Homecoming of American College Women: The Reversal of the College Gender Gap," *Journal of Economic Perspectives*, 20 (2006), 133–156.

Hoffmann, Florian, and Philip Oreopoulos, "A Professor Like Me: The Influence of Instructor Gender on College Achievement," NBER Working Paper No. 13182, 2007.

——, "Professor Qualities and Student Achievement," *Review of Economics and Statistics*, 91 (2009), 83–92.

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger, "What Does Certification Tell Us about Teacher Effectiveness? Evidence from New York City," *Economics of Education Review*, 27 (2008), 615–631.

Kane, Thomas J., and Douglas O. Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," NBER Working Paper No. 14607, 2008.

Lennox, Annie, and Dave Stewart, "Sisters Are Doin' It for Themselves," in *Eurythmics: Be Yourself Tonight*, Dave Stewart, ed. (RCA Records, 1985).

National Bureau of Economic Research, "Diversifying the Science and Engineering Workforce: Women, Underrepresented Minorities, and Their Science and Engineering Careers" (http://www.nber.org/sewp/events/2005.01.14/Agenda-1-14-05-WEB.htm, 2005).

National Science Foundation, "Science and Engineering Degrees: 1966–2004," Manuscript NSF 07-307, National Science Foundation, Division of Science Resources Statistics, 2006.

Neumark, David, and Rosella Gardecki, "Women Helping Women? Role Model and Mentoring Effects on Female Ph.D. Students in Economics," *Journal of Human Resources*, 33 (1998), 220–246.

Pope, Devin G., and Justin R. Sydnor, "A New Perspective on Stereotypical Gender Differences in Test Scores," *Journal of Economic Perspectives*, 24 (2010), 95–108.

Rabe-Hesketh, Sophia, and Anders Skrondal, *Multilevel and Longitudinal Modeling Using Stata*, 2nd ed. (College Station, TX: Stata Press, 2008).

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain, "Teachers, Schools and Academic Achievement," *Econometrica*, 73 (2005), 417–458.

Rothstein, Donna S., “Do Female Faculty Influence Female Students Educational and Labor Market Attainments?” *Industrial and Labor Relations Review*, 48 (1995), 515–530.

Spencer, Steven J., Claude M. Steele, and Diane M. Quinn, “Stereotype Threat and Womens Math Performance,” *Journal of Experimental Social Psychology*, 35 (1999), 4–28.

Weinberger, Catherine J., “Race and Gender Wage Gaps in the Market for Recent College Graduates,” *Industrial Relations*, 37 (1998), 67–84.

——, “Mathematical College Majors and the Gender Gap in Wages,” *Industrial Relations*, 38 (1999), 407–413.

——, “Is Teaching More Girls More Math the Key to Higher Wages?” in *Squaring Up: Policy Strategies to Raise Women’s Incomes in the U.S.*, Mary C. King, ed. (Ann Arbor: The University of Michigan Press, 2001).

Xie, Yu, and Kimberlee A. Shauman, *Women in Science: Career Processes and Outcomes* (Cambridge, MA: Harvard University Press, 2003).

Zafar, Basit, “College Major Choice and the Gender Gap,” Federal Reserve Bank of New York, Working Paper 364, 2009.