ISSN 2385-2755
DiSSE Working papers
[online]

SAPIENZA
UNIVERSITÀ DI ROMA

WORKING PAPERS SERIES
DIPARTIMENTO DI SCIENZE SOCIALI ED ECONOMICHE

Closing the Gender Divide in STEM The impact of female college professors on female student performance

Francesco D’Angelo Giuseppe De Arcangelis
Joanna Kopinska

N. 5/2025

SAPIENZA - UNIVERSITY OF ROME
P.le Aldo Moro n.5 – 00185 Roma T(+39) 0649910563
CF80209930587 – P.IVA 02133771002

Closing the Gender Divide in STEM
The impact of female college professors on female student performance*

Francesco D'Angelo
Giuseppe De Arcangelis
Joanna A. Kopinska

April 10, 2025

Abstract

This study examines the role model effects of female professors on female students' academic performance. We leverage student-professor matched data from STEM Bachelor's programs at Sapienza University, the largest university in Europe, covering the period from 2017 to 2022. We exploit a quasi-experimental design based on the quasi-random assignment of students to professors of different genders due to the section composition of large classes based only on last name initials. Our findings show that the presence of female role models significantly benefits female students, both by narrowing the performance gap with male peers at the exam level and by improving long-term academic outcomes, such as meeting progression benchmarks and increasing persistence in the program. Importantly, we find no evidence of negative effects on male students' performance. From a policy perspective, fostering the presence of female instructors in first-year courses may enhance female students' performance, while also enhancing a less gendered perception of STEM fields among male students.

Keywords: Gender gap, STEM, Higher education, Quasi-experimental analysis

***JEL* codes:** I23, J16, J24, C26, I21

*Contact information: Francesco D'Angelo, Sapienza University of Rome, Department of Economics and Law, fr.dangelo@uniroma1.it; Giuseppe De Arcangelis, Sapienza University of Rome, Department of Economics and Social Sciences, giuseppe.dearcangelis@uniroma1.it; Joanna Kopinska, Sapienza University of Rome, Department of Economics and Social Sciences, joanna.kopinska@uniroma1.it.

We would like to thank Bruno Sciarretta and the Statistics Office of Sapienza for enabling this analysis and offering constant support. We also thank Edoardo Di Porto, Salvatore Lattanzio, Anna Maria Mayda, and Ashley Wong for their useful suggestions and feedback received.

1 Introduction

Although the role of women in STEM (Science, Technology, Engineering, and Mathematics) occupations has increased since 2016, the gender gap remains pronounced. Women represent only 28.2% of the global STEM workforce, whereas they make up approximately half of the workforce in non-STEM sectors, according to the Gender Gap Report 2024 published by the World Economic Forum (Kali Pal et al., 2024). At the highest levels in STEM, only 10% of the leadership positions are held by women, compared to one-quarter in non-STEM fields.

In high-income countries, the situation is somewhat more favorable. In Europe, for instance, Eurostat reports that in 2023, out of 71.8 million employees in science and technology aged 24-65, 37.7 million (over 50%) were women.[1] However, women remain underrepresented in top positions. For example, they make up 41% of scientists and engineers, a proportion that has stagnated for more than a decade.

The gender divide is even larger when considering enrollment in certain STEM degrees. In 2022, female enrollment in ICT and engineering programs in Europe was 20.2% and 27.3%, respectively, while in Italy, the figures were 15.4% and 28.5%. These differences in the chosen field of study are particularly relevant since they are able to explain a significant share – more than 50% – of the gender pay gaps in early career years, as documented by Bovini et al. (2024) and Arellano-Bover et al. (2024), and this dynamic is likely to persist given the increasing demand of ICT skills in the labor market.

In addition to direct policies – as promoted, for example, in Europe (Commission et al., 2022) – other interventions can also play a crucial role in reducing gender stereotypes and encouraging women to pursue careers in STEM.

Exposure to female role models, whether in early life or throughout secondary and tertiary education, has been shown to significantly influence career choices. As reported in Riise et al. (2022), in Norway, encountering female physicians during childhood increased the probability of pursuing STEM and medicine programs by 4%. Similarly, a randomized control trial in French high schools demonstrated that a one-hour talk by a female scientist significantly boosted the

[1] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Human_resources_in_science_and_technology#Women_in_science_and_technology

probability of 12-year-old girls choosing STEM programs in college by 3.4% (Breda et al., 2023).

The impact of teacher gender has also been a focus of research, in particular for primary and secondary tracks. de Gendre et al. (2023) conducted a meta-analysis showing that the exposure effect to female role models in primary education does not improve test scores. On the other hand, in secondary education, such effects are nearly universally positive. For instance, Winters et al. (2013) found that performance in middle and high schools is slightly affected by same-gender teachers, but this effect was not observed in elementary schools. Lim and Meer (2017) found that middle-school female students in South Korea performed better in classes taught by female teachers, with no detrimental effects for male students. Similarly, Dee (2007) found that both boys and girls benefited from same-gender teachers in terms of objective and subjective academic performance evaluated in the 1988 National Education Longitudinal Study on 8th graders. Muralidharan and Sheth (2016) demonstrated that the positive effect of female teachers on reducing the gender education gap is also evident in developing countries, such as the Indian state of Andhra Pradesh.

Few studies have investigated the role of professor gender at the post-secondary level, and all face either causal identification challenges due to selection bias or generalizability issues arising from specific settings or small sample sizes. For example, Hoffmann and Oreopoulos (2009) exploits within-student and within-instructor variation to estimate the impact of same-sex teachers on grades and the probability of dropping a class, finding small effects. Price (2010) shows that female students are *less* likely to persist in STEM when more of their courses are taught by female instructors, though this study is also likely affected by selection biases.

A study most relevant to our approach is Carrell et al. (2010), who use random assignment of students to professors at the U.S. Air Force Academy and find that exposure to female professors positively affects female students' performance in math and science, with high-performing students being more likely to major in STEM fields.

Other related studies focus on enrollment and career choices. Porter and Serra (2020) reports that introductory undergraduate courses taught by female instructors positively influenced students' decisions to pursue economics majors. Additionally, Mansour et al. (2022) examines students at the U.S. Air Force Academy and finds that being assigned a female professor increases the probability of working in a STEM occupation and earning a STEM master's degree.

In this study, we examine the impact of female role models, represented by female professors teaching STEM undergraduate courses at Sapienza University, on the academic achievement of female STEM students. Specifically, our aim is to assess whether having female professors significantly and positively influences female students' academic performance, both in individual courses and throughout their entire degree program. At the same time, we can evaluate whether and how this exposure affects male peers.

We analyze data from STEM Bachelor's degree programs at Sapienza University, the largest university in Europe, focusing on matched student-professor data for the period 2017-2022. We leverage the fact that large classes in bachelor's programs are often divided into multiple parallel sections, with students randomly assigned to these sections based on the initial letter of their last name - a process henceforth referred to as *channeling.* This method effectively yields a quasi-random assignment of students to instructors, with the potential for gender variation among the assigned professors. By controlling for a battery of fixed effects, we first examine the impact of female students' exposure to female professors on the exam grades in the relevant subjects. We then investigate how this female-to-female exposure influences academic outcomes in subsequent courses, as well as its effect on the likelihood of dropping out, both within the same first academic year and in subsequent years.

Our analysis reveals that randomly assigned female professors significantly improve female students' performance without negatively affecting their male peers. As a result, female students are able to close the performance gap with their male counterparts, which is present in male-taught sections. In particular, role model effects are estimated to be between 5% and 8% of one standard deviation (SD) in final exam grade, and they also increase the probability of success on the first attempt by slightly less than 10% - which is equivalent to around 3-4 percentage points.

Subsample heterogeneity in the results suggests that we are actually able to identify role model effects since we find stronger results in the presence of low GPA students and students with a non-scientific background. The fact that the largest effects are found among lower-performing students without prior preparation in math and science is particularly noteworthy, as this group of women is arguably the most likely to drop out and exacerbate the gender gap. Additionally, the effect is stronger for younger professors, which further supports the interpretation of a role-model effect. We also find no significant effect when limiting the analysis to the COVID-19 period, which serves as a placebo test for the role-model hypothesis.

When examining the long-term association between exposure to female professors on female students' academic careers, we find no significant interaction effects on GPA and probability of graduating on time. However, importantly, we observe that such exposure is significantly linked to lower dropout rates by approximately 2 percentage points - a decrease of about 25% relative to the baseline - and to an increase in the probability of obtaining at least 20 credits in the first and second year in the program. These effects tend to dissipate over time, and cease being significant two years following the exposure. These results align with those of Carrell et al. (2010), who found that same-gender instructors have only limited impact on male students, but it has 5% of a SD 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.

In our setting, the random assignment of students combined with a focus on pure STEM disciplines at Europe's largest public university, along with mandatory course enrollment, allows us to investigate the impact of professor gender on student outcomes while avoiding the self-selection and attrition biases that have affected a relevant share of the existing literature.

The paper is structured as follows. The next Section reviews some general characteristics of college education in Italy and some stylized facts about STEM programs; hence, this section can be skipped by readers who are familiar with the Italian university system. We recommend Section 3 to acquire a statistical description of the STEM at Sapienza and on the dataset that we use is our study. Section 4 illustrates our empirical setup and the regression model describing the quasi-experimental framework. The results are reported and discussed in Section 5 where we also report various refinements as robustness checks. Section 6 summarizes the main message of the analysis and concludes with policy suggestions for university officials.

2 Tertiary Education and STEM in Italy

The following section can be omitted by readers familiar with the university system in Italy, while Subsection 2.2 provides some national information on the STEM fields in Italian universities.

2.1 Institutional Background

In 1999, the harmonization of European tertiary education introduced a two-tier system: first-level Bachelor's degrees (BA), typically completed in three years, and second-level degrees (MA), designed to be completed in two years.[2]

Similar to other countries (e.g., the United States), access to higher education is not contingent upon specific high school formation. Instead, entry tests for each discipline are required. Therefore, students with any type of high school diploma, including five-year vocational schools or artistic high schools, are eligible to take the entry tests for STEM degree programs.

The entry tests, known as TOLC, are standardized nationwide and mandatory for admission. However, in the absence of caps for first-year students, they do not serve as a barrier to entry.

When enrollment caps are set, performance on entry tests determines access to the degree programs. In particular, the 1999 Law (264/1999) mandates that national caps be established for certain disciplines like Architecture, Dentistry, and Medicine. Universities must adhere to the national caps while also setting their own limits based on institutional constraints, such as the availability of specialized laboratories, IT infrastructure, and other technical resources. At Sapienza University, enrollment caps are applied to selected STEM degrees, including Architecture (in accordance with the national cap) and most Engineering programs. In contrast, fields like Physics, Chemistry, and Mathematics typically remain uncapped.

To complete a degree, students are required to complete a specified number of exams (20 for the BA and 12 for the MA) and fulfill additional requirements, such as internships or practical training.[3]

[2]More recently, some two-year technical degrees have been introduced; however, they still represent a relatively small share of the tertiary student population.

[3]Activities and exams are allocated varying ECTS credits based on the workload involved, with 180 ECTS needed for a BA and 120 ECTS for an MA.

With the exception of a few courses that involve projects, attendance is not mandatory, and there is no formal attendance bonus. Consequently, the grading system and scale are uniform for both attending and non-attending students. Evaluation is typically based solely on a final exam, especially in compulsory first- and second-year classes. The grading scale ranges from a minimum passing score of 18 to a maximum of 30, with *cum laude* awarded for exceptional performance exceeding 100%.

Exams are conducted at the end of each course but are offered multiple times throughout the academic year (up to ten sittings), allowing students the flexibility to retake exams, improve their grades, or defer assessments. Notably, students are not required to complete first-year courses before enrolling in second- or third-year classes, which often results in students delaying their first-year exams.

The delay affects the timely completion rates for both BA and MA degrees, as reported in the Almalaurea[4] data presented in Table 1.

Table 1: Timeliness of program completion among graduates in Italy

	On-time		Avg. Duration	
	BA's	MA's	BA's	MA's
All Courses				
2017	50.8	58.6	4.3	2.8
2018	53.9	60.1	4.2	2.7
2019	56.1	61.0	4.2	2.8
2020	57.7	64.3	4.2	2.8
2021	60.1	67.0	4.1	2.8
2022	62.4	66.4	4.0	2.7
STEM Courses				
2017	41.4	51.4	4.7	2.9
2018	45.9	52.7	4.4	2.8
2019	47.9	54.1	4.5	2.9
2020	50.7	58.7	4.4	2.9
2021	51.9	61.1	4.4	2.8
2022	54.0	60.4	4.3	2.8

Note: The data come from Almalaurea - Graduates' Profile Survey. Columns 2 and 3 display the share of on-time graduates on total graduates in a specific year. Columns 4 and 5 report the average time needed to graduate measured in years.

While the share of on-time graduates has shown an upward trend in recent years, overall

[4] The Graduates' Profile Survey by Almalaurea covers around 90% of all graduates from Italian universities, excluding fully online universities. Universities in the North-West are slightly underrepresented, as Bocconi, Cattolica, and Milan Polytechnic have not joined the consortium.

completion rates remain very low, ranging from 50% to 62% for BA degrees and 58% to 66% for MA degrees. The picture is even more concerning in STEM programs (as shown in the lower panel of the table), where on-time graduation rates are several percentage points lower, and the average time to degree completion is longer across all levels.

2.2 STEM in Italy

This subsection provides an overview of the STEM landscape in Italy from 2017 to 2022, with a specific focus on bachelor's programs, as they form the core of our analysis. Data on student enrollment and faculty are sourced from the Italian National Register of Students and Graduates. Since this dataset does not provide information on student performance, we supplement it with the previously mentioned Almalaurea survey to examine graduates' outcomes.

Enrollment. As reported in Table 2, enrollment in STEM bachelor's programs has experienced a steady increase over the period of interest, with a cumulative growth rate of 9.5% between 2017 and 2021. A significant portion of this growth occurred in 2020, probably driven by the shift to remote lecture attendance during COVID-19 lockdowns. Despite the overall rise in enrollment, the share of female students has remained fairly stable, hovering around 38% of the total student population.

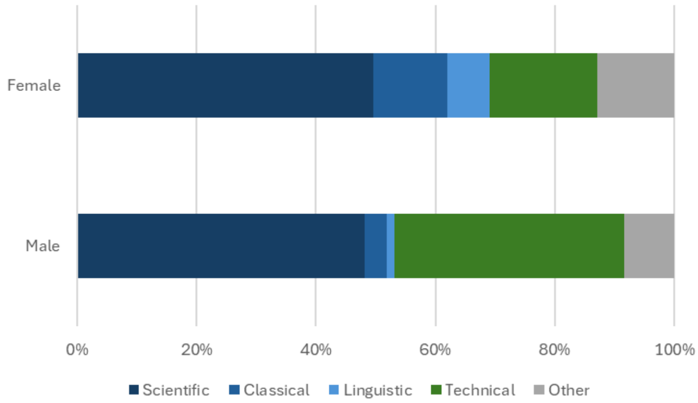
Table 2: Enrollment in STEM BA's in Italy

A.Y.	Male	Female	Total	Female Share
2017/2018	54831	32859	87690	0.37
2018/2019	55122	34013	89135	0.38
2019/2020	57645	35289	92934	0.38
2020/2021	59683	37425	97108	0.39
2021/2022	59435	36621	96056	0.38

Note: The data come from Italian National Register of Students and Graduates. The table reports the number of enrolled students in STEM bachelor's courses in Italy. Courses included are those reported in Appendix D, with the addition of two degrees (degree class labels L-21 and L-28).

Male and female students enter tertiary education with different backgrounds, as shown in Figure 1. A larger proportion of female students comes from classical, scientific, and linguistic high schools (represented in different shades of blue), whereas male students are predominantly drawn from technical high schools.

Figure 1: The High School Background of STEM Students in Italy



Note: The data come from Italian National Register of Students and Graduates. The figure reports the high school background of students enrolled in Italian STEM bachelor's courses in the 2021/2022 academic year. These patterns are largely consistent with those observed in previous cohorts.

Performance. As previously mentioned, there has been a remarkable increase in the share of on-time graduates between the first and last year of analysis, a trend observed for both male and female students. Although less pronounced, improvements are also observed in GPA, final (graduation) grade, and the overall number of graduates. Moreover, female graduates consistently outperform their male counterparts across all metrics reported in Table 3, both at the beginning and at the end of the period.

Table 3: Performance measures of STEM graduates in Italy

	Male				Female			
Year	Graduates	On-Time	GPA	Final Grade	Graduates	On-Time	GPA	Final Grade
2017	24413	40.8	24.8	97.5	15749	42.3	25.2	99.3
2022	25237	52.3	25.2	99	16558	56.6	25.6	100.9

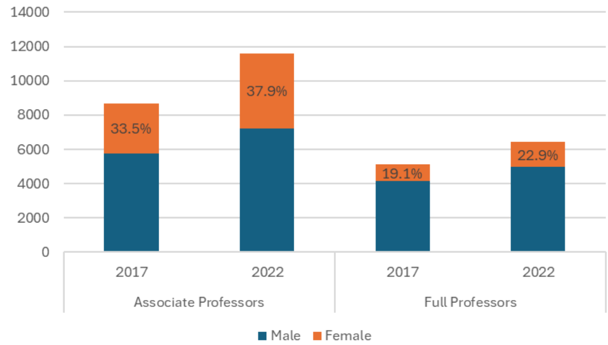
Note: The data come from Almalaurea - Graduates' Profile Survey. The table reports several performance measures of STEM bachelor's graduates in Italy for the indicated solar years. GPA and Final Grade are respectively out of 30 and out of 110. The increases remain fairly smooth over the omitted years, with the exception of the number of graduates where the peak is reached in 2020 for males and in 2021 for females.

Professors. The total number of professors increased between 2017 and 2022 (Figure 2).[5]

[5] However, this increase does not necessarily reflect a rise in number of instructors, as some assistant professors

In contrast to the trend observed among female students, the share of female associate and full professors has seen a slight improvement over the years. Nevertheless, these shares remain low. Female associate professors constitute less than 40% of the total, while the share of female full professors remains below 25%.

Figure 2: The Profile of STEM Professors in Italy



Note: The data come from Italian National Register of Students and Graduates. The figure reports the number of professors in academic disciplines primarily associated with STEM courses (*Area 01:* Mathematical and Computer Sciences; *Area 02:* Physical Sciences; *Area 03:* Chemical Sciences; *Area 04:* Earth Sciences; *Area 05:* Biological Sciences; *Area 08:* Civil Engineering and Architecture; *Area 09:* Industrial and Information Engineering), excluding Economics and Statistics (*Area 13*). The observed increases remain consistent across the omitted years.

3 STEM at Sapienza University

In this section, we refer to administrative data from Sapienza University to provide a comprehensive analysis of the STEM landscape at this major university. A subset of this data is used in the econometric analysis presented in the following sections. The focus remains solely on STEM bachelor's degrees, and data are relative to students who enrolled between the 2017/2018 and 2021/2022 academic years.

Enrollment. Each year, approximately 6,000 students enroll in STEM bachelor's programs at Sapienza, resulting in a total of 30,844 observed individuals over the years of interest. Similar to the national trend, the share of female students has remained relatively stable, albeit at a slightly higher level, consistently comprising around 41% of the total student population.

advanced to associate professors, and some associate advanced to full professorships during the observation period.

More heterogeneity is detected when examining enrollment data by school, as reported in Table 4. The number of enrolled students varies significantly across schools, primarily due to the enrollment caps mentioned in Section 2. The Mathematics, Physics, and Natural Sciences (MPNS) School gets the largest share of students (35%), while the School of Architecture enrolls the smallest proportion, comprising only 4% of undergraduates.

Also, the share of female students varies significantly across STEM schools, as previously highlighted in a report by Assolombarda et al. (2020) for the Italian case and in Bettinger and Long (2005) for public colleges in Ohio. While the female share is relatively balanced in the Architecture and MPNS Schools, women are underrepresented in Engineering but overrepresented in Pharmacy and Medicine. Regarding foreign students, their share remains below 10% across all schools, with the only notable variation being a lower proportion (3%) in the MPNS school.

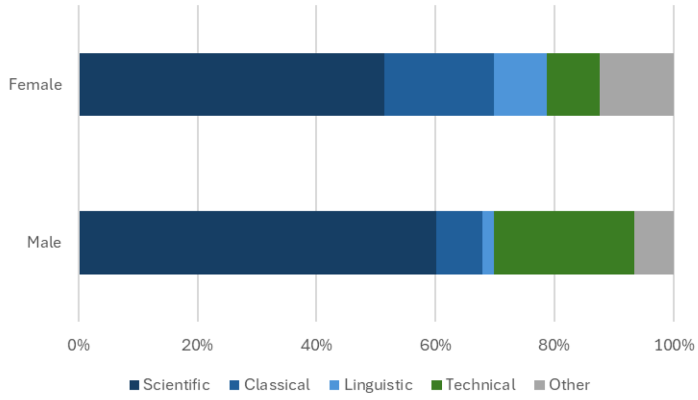
Table 4: STEM Enrollment by School at Sapienza University

School	Students	School Share	Female Share	Foreign Share
Architecture	1201	0.04	0.56	0.07
Civil/Industrial Engineering	7141	0.23	0.33	0.08
Information Engineering/Informatics/Statistics	8784	0.28	0.23	0.06
Mathematical/Physical/Natural Sciences (MNPS)	10763	0.35	0.52	0.03
Pharmacy and Medicine	2955	0.10	0.73	0.07

Note: The table reports information retrieved from Sapienza administrative data relative to all enrolled students in STEM bachelor's at Sapienza between the 2017/2018 and 2021/2022 academic years. Yearly figures show some minor variability and are available upon request.

Considering the high school background of students, we observe a similar gender-based distribution as seen in the national context discussed in Subsection 2.2 and reported in Figure 1. As shown in Figure 3, female students at Sapienza are more likely to come from classical and linguistic tracks compared to their male counterparts, whereas male students are predominantly from technical high schools, with scientific high schools comprising the second largest group. Both female and male STEM students at Sapienza are more likely to come from a lyceum background (named classical, scientific, linguistic in Figure 3 and represented in different shades of blue) more frequently than their Italian counterparts, but the highest increase is for males (almost 70% in Sapienza versus 53% nationally).

Figure 3: STEM Students' High School Background at Sapienza University



Note: The figure reports information retrieved from Sapienza administrative data relative to the high school background of students enrolled in STEM bachelor's programs at Sapienza University between the 2017/2018 and 2021/2022 academic years. The distribution remains stable across academic years, with no significant variation observed.

Performance Despite the guidance provided through high school orientation and the TOLC entry tests, dropout rates in STEM degrees remain notably high. In our sample, around 19.5% of enrolled students have not taken any exams. As shown in Table 5, dropout rates are unevenly distributed over the years of study, with a marked concentration in the first year. This trend is particularly pronounced among female students, whose first-year dropout rate accounts for over half of their total dropout rate and exceeds the corresponding male rate by 4 percentage points. This pattern suggests that female students are disproportionally affected by their first-year performance.

Considering dropout rates across all years, female students tend to drop out less than males in all schools except for the MPNS school, which enrolls the lion's share of students. This concentration ensures that the overall dropout rate remains balanced between genders. Surprisingly, dropout rates are lowest in Engineering Schools for both genders, an unexpected outcome given the significant underrepresentation of female students in this field. Conversely, dropout rates exceed 30% for both genders in the Pharmacy and Medicine school, a figure that may reflect strategic behavior by students seeking eventual entry into Medical schools.[6]

[6] Indeed, the share of transfers to other Schools is markedly higher in the Pharmacy and Medicine School compared to other schools. This pattern is largely driven by students who initially enroll in Pharmacy STEM programs after failing to secure admission to Medical School. These students often use Pharmacy degrees to

Table 5: STEM Dropout Rates by School at Sapienza University

Male Students				
School	All years	1st	2nd	3rd
Architecture	0.22	0.08	0.07	0.06
Civil/Industrial Engineering	0.17	0.07	0.06	0.04
Information Engineering/Informatics/Statistics	0.19	0.07	0.07	0.05
Mathematical/Physical/Natural Sciences (MNPS)	0.22	0.11	0.07	0.04
Pharmacy and Medicine	0.37	0.23	0.12	0.05
All Schools	0.20	0.09	0.07	0.05
Female Students				
Architecture	0.14	0.07	0.04	0.03
Civil/Industrial Engineering	0.12	0.06	0.04	0.02
Information Engineering/Informatics/Statistics	0.12	0.05	0.04	0.03
Mathematical/Physical/Natural Sciences (MNPS)	0.24	0.15	0.06	0.03
Pharmacy and Medicine	0.35	0.26	0.08	0.04
All Schools	0.21	0.13	0.05	0.03

Notes: The table reports information retrieved from Sapienza administrative data relative to the dropout rates of STEM bachelor's students at Sapienza. The adopted definition of dropout includes official abandonments, transfers, and absence of enrollment renewals. The 2021/2022 cohort is absent in the 3rd career year statistics.

We restrict our sample to students who have at least one recorded exam in their transcripts, excluding those without any exam recorded. This reduces our sample from 30,844 to 24,826 students, encompassing 301,192 exam records. Furthermore, we drop all records referring to the 2022/2023 academic year (13% of total exam records) since the information is partial, including information up to July 2022. Additionally, we exclude records unrelated to formal exams (e.g., supplementary activities or theses), which represented around 14% of the remaining observations. After implementing minor adjustments to align student and professor data, our final sample comprises 23,535 students, and 222,532 matched student-professor exam records. All statistics relating to students' performance and the subsequent econometric analysis are based on this final dataset.

As reported in Table 6, the average performance of first-year students is rather poor, with both the mean and median credits earned falling below 40, a threshold typically associated with satisfactory academic progress. Moreover, the 75th percentile of credits earned remains far below the 60-credit mark, which represents one-third of the 180 credits required for completing a bachelor's degree.[7]

The situation improves notably in the second and third years of the degree, where the

take exams and earn credits that they can later be transferred to Medical School upon successful admission in subsequent attempts.

[7]This picture is further compounded by the fact that we only consider students who have earned at least one credit, which may skew the results, as the inclusion of overachievers can inflate the average values.

Table 6: STEM students' performance by career year at Sapienza University

Career Year	Min	p25	p50	p75	Max	Mean
1st						
Exams	1	2	4	5	10	3.78
Credits	6	18	36	48	81	33.24
2nd						
Exams	1	3	5	7	11	4.81
Credits	6	26	42	57	97	40.01
3rd						
Exams	1	4	6	8	18	6.12
Credits	6	33	51	60	144	46.72

Note: The table reports information retrieved from Sapienza administrative data relative to performance measures for STEM bachelor's students at Sapienza. Only students with at least one exam registered in a career are included. Information for career years above the 3rd is omitted. The 2021/2022 cohort is absent in the 2nd and 3rd career year statistics. The 2020/2021 cohort is absent in the 3rd career year statistics.

distributions of both exam performance and credits move towards higher values, likely reflecting a selection effect. Specifically, students who underperform in the first year are more prone to dropping out, leaving a higher proportion of stronger performers in subsequent years.

The average on-time graduation rate (within three years) stands at around 30%, with 15% of students graduating with an extra year. As confirmed in Table 7, among on-time graduates, women outperform men across all schools except for MPNS. This difference is sufficient to balance the overall graduation rates between both genders, given the higher enrollment rate in MPNS. Notably, the highest graduation rates are observed in the Engineering schools, while the lowest are found in the Pharmacy and Medicine faculty. A slightly greater variability is observed when considering those who graduate with an additional year, probably due to selection effects.

Finally, considering class grades as a measure of student performance, we observe that the GPA distribution of both sexes gradually shifts toward higher values and becomes more concentrated in the second and third years, reflecting the effect of selection on the intensive margin. As shown in Figure 4a, during the first year, female students tend to achieve medium-high grades more frequently than their male counterparts, who, conversely, are more likely to attain very outstanding GPAs. However, this picture shifts in favor of female students over the years, and by the third year, the female GPA distribution stochastically dominates that of

Table 7: STEM graduation rates by School at Sapienza University

School	Males		Females	
	3 Years	3+1 Years	3 Years	3+1 Years
Architecture	0.24	0.20	0.39	0.14
Civil/Industrial Engineering	0.31	0.18	0.39	0.21
Information Engineering/Informatics/Statistics	0.35	0.14	0.50	0.17
Mathematical/Physical/Natural Sciences (MNPS)	0.32	0.15	0.28	0.16
Pharmacy and Medicine	0.14	0.10	0.16	0.12
All Schools	0.31	0.15	0.32	0.16

Note: The table reports information retrieved from Sapienza administrative data relative to graduation rates for STEM bachelor's students at Sapienza. A bachelor degree is supposed to be completed in 3 years, however many students exceed this time limit, hence we include information on students graduating in 4 years. The 2021/2022 and 2020/2021 cohorts are absent in the 3-year statistics. The 2021/2022, 2020/2021, and 2019/2020 cohorts are absent in the 4-year statistics.

males for all values exceeding a GPA of 24, as shown in Figure 4c.

Professors. During the period of interest, a total of 1,348 distinct professors taught STEM bachelor's courses at Sapienza University, with 36.5% of them being female. Since professors can teach across multiple schools, the faculty compositions reported in Table 8 show a higher total number of instructors (1,802), as professors are counted multiple times for each school they teach in. Despite this repetition, the faculty composition by class instructor used in our analysis is more appropriate for capturing the exposure of students to faculty members of varying genders, aligning with the focus of our study.

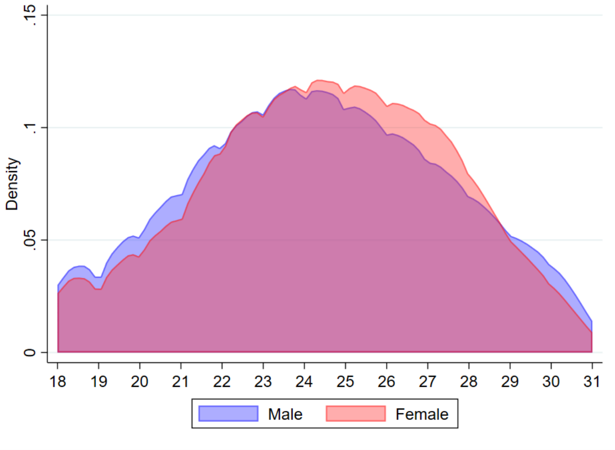
Table 8: Professors in STEM Bachelor's Courses at Sapienza University

School	Professors	Female	Foreign	Male Age	Female Age
Architecture	174	0.39	0.01	53.6	52.4
Civil/Industrial Engineering	403	0.30	0.02	53.0	52.7
Information Engineering/Informatics/Statistics	404	0.31	0.03	51.9	52.1
Mathematical/Physical/Natural Sciences (MNPS)	618	0.39	0.03	53.9	53.0
Pharmacy and Medicine	203	0.47	0.04	52.8	53.1
All Schools	1802	0.36	0.03	53.1	52.7

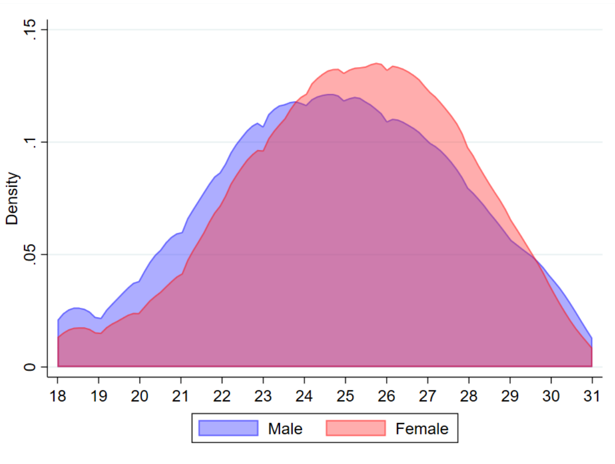
Note: The table reports information retrieved from Sapienza administrative data relative to professors teaching in STEM bachelor's courses at Sapienza between 2017/2018 and 2021/2022 academic years. Professors teaching in more than one school are counted multiple times. Professors' age is firstly averaged at the individual level and secondly at the school level.

The number of faculty members is approximately proportional to student enrollment across the various schools, with the lowest count in the School of Architecture (174) and the highest in the Mathematics, Physics, and Natural Sciences School (618). The percentage of female

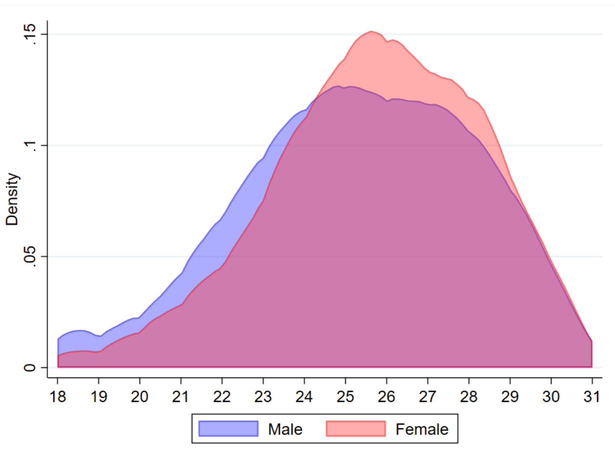
Figure 4: STEM GPA Distribution by Career Year at Sapienza University



(a) 1st Year



(b) 2nd Year



(c) 3rd Year

Note: The figure reports information retrieved from Sapienza administrative data relative to the GPA distribution of STEM bachelor's students at Sapienza University. A GPA of 31 corresponds to a score of 30 *cum laude*. The 2021/2022 cohort is not represented in the 2nd and 3rd year statistics, while the 2020/2021 cohort is absent from the 3rd year statistics.

professors is below 50% in all schools, with the highest value of 47% in the School of Pharmacy and Medicine.

Similarly to the distribution of female students, female professors are notably underrepresented in Engineering schools, a disparity that may foster gender stereotypes and influence the career choices of young women considering entering this field.

The share of foreign professors is rather low across all schools, consistently below 5%. The highest share is recorded in Pharmacy and Medicine schools, while the School of Architecture reports only 1% of foreign-born faculty members. In general, male professors are slightly older than their female counterparts, with an average age of 53.1 years across all schools compared to 52.7 years for women. However, this age gap varies significantly across different schools.[8]

3.1 Summary of Stylized Facts at Sapienza

We can summarize the main stylized facts by focusing on the gender composition of students and professors in the STEM at Sapienza University as follows:

- The share of female students varies significantly across STEM schools
- Female students are more likely to come from classical and linguistic tracks compared to their male counterparts
- First-year dropout rate of female students accounts for over half of the total dropout rate, but considering dropout rates across all years, female students tend to drop out less than males in all schools except for the MPNS school
- Among on-time graduates, female students outperform male students across all schools except for MPNS
- The female GPA distribution stochastically dominates that of males for all values exceeding a GPA of 24
- The number of faculty members is approximately proportional to student enrollment across the various schools

[8]Some heterogeneity across schools is observed, with female professors being, on average, older in the School of Pharmacy and Medicine and in one of the two Engineering schools.

- The share of foreign professors is below 5% in all schools.

4 Empirical setup

The primary objective of our study is to evaluate whether the academic performance of female students, both within individual courses and throughout their entire degree program, is significantly and positively influenced by the presence of female professors.

The presence of female faculty in STEM Bachelor's programs could be associated with role models' effects, which could positively influence the academic performance of female students both in terms of grades and persistence, reducing the probability of dropping out.

To investigate this hypothesis, we analyze the impact of female student-professor interactions by exploiting a unique quasi-random assignment of students to instructors based on the initial letter of their last name. The assignment process, henceforth referred to as *channeling*, provides an exogenous source of variation in the gender of the assigned professors. A similar quasi-experimental setting has been analyzed in Mengel et al. (2018), who, however, explore students' evaluations of instructors as the main outcome variable.

4.1 Focus on first-year students

We use matched student-professor data from all STEM Bachelor programs at Sapienza University for the period from September 2017 to December 2022, hence including five students' cohorts (2017-18, 2018-19, 2019-20, 2020-21, 2021-22).

We focus on first-year students, as they exhibit the highest dropout rates and are particularly susceptible to external influences during their transition from high school to the academic environment. At this formative stage, students are more inclined to seek motivational cues, including the presence of role models, to guide their academic trajectories.

Moreover, the structure of STEM programs in the first year is largely standardized, with limited or no elective options. This stands in contrast to the later years of the programs, which typically offer greater flexibility through a broader range of elective courses. Such uniformity minimizes the scope for selection bias, providing a more controlled and consistent academic framework to analyze the impact of female role models.

First-year classes are also characterized by their large size, with an average of 63% of the students assigned to multiple sections, as we describe thoroughly in the following Subsection 4.2 as it is a feature we take advantage of for our identification.

As previously reported, students can delay taking exams without incurring penalties under the program's requirements.[9] However, those who defer their exams beyond the scheduled time frame may ultimately attend courses taught by a professor different from their initially assigned instructor. To mitigate the potential measurement error arising from this variation, we restrict our analysis to scheduled first-year exams that are passed on time, that is, during the first year of a student's career.[10] This accounts for 86% of first-year scheduled exams.

4.2 *Channeling*: our quasi-experimental setting

Although most first-year courses are mandatory, academically underprepared students may adopt coping strategies to navigate the perceived academic rigor. These strategies often involve selecting schools, disciplines, and courses taught by professors whose grading standards are perceived as more lenient. Consequently, students may self-select into specific degree programs, potentially influenced by the proportion of female faculty, and into particular classes that are systematically assigned to female professors.

If this selection process is not orthogonal to the gender of either students or professors, the interaction between student and professor gender risks being confounded by a selection bias. Consequently, estimating the effect of exposure to female professors without adequately controlling for these selection dynamics could introduce endogeneity, leading to biased results.

To mitigate this issue, we exploit the *channeling* mechanism, i.e. the assignment of students to multiple sections of the same course based on the alphabetical order of the last names. For instance, in the case of two sections, students with last names starting with *A*-*L* are assigned to one section, while those with *M*-*Z* are assigned to another. Each section is taught by a distinct professor who is responsible for delivering the entirety of the course's instructional

[9] In many programs, students who arrive at the final test for graduation within three years receive a bonus in points for the graduation grade, but no penalties are present for laggards.

[10] Notice that this restriction is only introduced to impute the correct student-professor match. During one career year, the student can still take the same exam multiple times, as we will see in the following.

hours. Occasionally, teaching assistants are allocated additional hours for review sessions or exercise, though these supplementary contributions generally do not exceed one hour per week.

A similar approach is adopted in Mengel et al. (2018) where the authors study gender bias in professor evaluations in the Netherlands. This quasi-random assignment process ensures that students are distributed across instructors who may differ in gender, enabling us to isolate the causal impact of female student-professor interactions on academic achievement from potential confounding factors associated with non-random selection.

On average, 63% of first-year students attend classes distributed across multiple sections, with 46% of these sections being taught by professors of mixed gender, as documented in Table 9. These figures vary significantly across different schools.

Table 9: STEM students' exposure to channeling in 1st year courses at Sapienza University by school

School	1 Ch.	2 Ch.	3+ Ch.	Mixed Gender
Architecture	0.60	0.37	0.02	0.59
Civil/Industrial Engineering	0.52	0.48	0.00	0.43
Information Engineering/Informatics/Statistics	0.44	0.56	0.00	0.37
Mathematical/Physical/Natural Sciences (MNPS)	0.19	0.28	0.53	0.54
Pharmacy and Medicine	0.35	0.25	0.40	0.36
All Schools	0.37	0.40	0.23	0.46

Note: The table reports information retrieved from Sapienza administrative data relative to students' exposure to channeling in 1st-year courses of STEM bachelor's at Sapienza. Columns report the average share of exams with a specified number of sections on total exams given by students. Students' values are averaged at the school level. The Mixed Gender column reports the share of channeled courses (those with more than 1 Ch.) taught by mixed-gender professors.

Students in the MPNS school attend, on average, 81% of their first-year exams in multiple channels, with 53% of these exams distributed across three or more sections and 54% being taught by mixed-gender faculty. In contrast, Architecture students experience less exposure to the channeling process, with only 39% of their first-year classes divided into multiple sections. However, the share of mixed-gender faculty in these classes is the highest across all schools, at 59%. Conversely, students in Pharmacy and Medicine, where 65% of first-year classes are divided into multiple sections, encounter the lowest share of mixed-gender faculty, with only 36% of these sections being taught by professors of both genders.

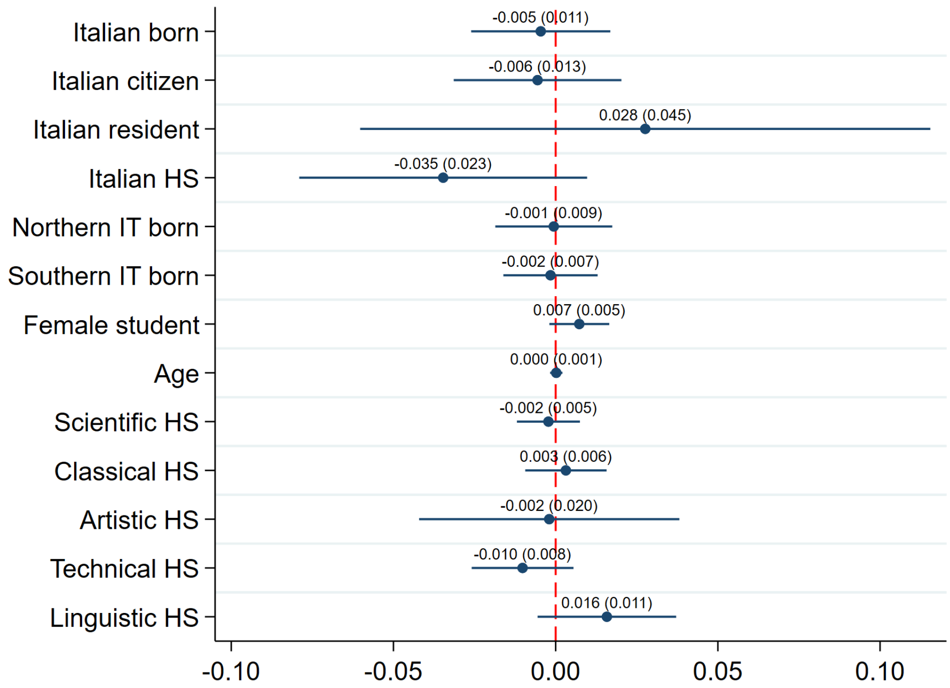
Balance Testing. While the assignment process based on alphabetical order provides an apparent exogenous source of variation in the exposure of students to professors of different genders, it is important to critically assess the assumption that this process constitutes a fully random assignment. In theory, randomization would imply that there are no systematic differences between students assigned to professors of different genders. However, a number of factors may introduce subtle biases that could undermine this assumption. First, while the channeling mechanism allocates students to professors of different genders based on their last name, there may be unobserved factors related to the student's name, such as cultural or socio-economic background, that influence academic outcomes. Additionally, the geographical origin of the student may have an influence on the student's last name (influencing selection into treatment) and be correlated with their high school type choice, which clearly affects future academic achievement.

Given these potential biases, we conduct balance tests to ensure that the assignment of students to professors is random and that no significant differences exist between students assigned to male or female professors in terms of observable characteristics. Specifically, our balance test in Figure 5 compares the baseline characteristics of students across the assignment groups (i.e., those assigned to male professors versus those assigned to female professors) within mixed-gender channeled courses. The test examines whether key observable covariates – such as citizenship, country of origin and residence, age, gender, and the type of secondary school attended – are balanced across the treatment groups. The results indicate that the assignment mechanism does not introduce systematic selection bias, thereby validating the use of this quasi-random setting to estimate the causal effect of female role models on student outcomes.

In contrast, when examining the complementary scenario to mixed-gender channeling - instances where students are enrolled in single-section courses or multiple-section courses with professors of the same gender - the balance test raises significant concerns. In those cases, it is not possible to compare the performance of students in the same course but assigned to professors of different genders. This context introduces potential selection bias, as it is plausible that female students with different characteristics may ex-ante opt for programs or courses with a greater representation of female instructors.[11]

[11] The results from the corresponding balance tests (Figure A1 in Appendix A) confirm this concern, showing that students exposed to female professors in this sample are disproportionately foreign-born, foreign citizens, and have different high-school backgrounds.

Figure 5: Balance test - *Channeled* Courses with Mixed Gender Professors



Note: The figure presents the results of a balancing test based on Sapienza administrative data, where we regress the gender of the instructor ($female = 1$) on pre-determined student characteristics. The sample includes all first-year *channeled* exams taken on time in courses with mixed-gender professors, consistent with our quasi-experimental setting. Each coefficient estimate originates from a separate model, which includes course fixed effects. Standard errors are clustered at the section level.

These imbalances highlight the limitations of comparisons between courses for isolating teacher gender causal effects, as the lack of randomization undermines the ability to disentangle the role of female instructors from pre-existing selection dynamics.

4.3 Setting up the Empirical Model

Given the quasi-random assignment mechanism, we leverage the exogenous variation in professor gender to estimate the causal effect of exposure to female professors on student outcomes. Specifically, the allocation process allows us to isolate the impact of professor gender while mitigating potential selection biases that would otherwise confound the relationship between student performance and faculty gender.

As discussed before, we focus on exams taken on time by first-year students in courses taught in multiple sections with professors of mixed-gender, in all STEM Bachelor's programs at Sapienza University from September 2017 to December 2022.[12] To compare exams taken by female students exposed to female professors as opposed to male students, we estimate the following model specification:

$$\begin{aligned} y_{ijc} = &\beta_0 + \beta_1 Female\,Student_i + \beta_2 Female\,Prof_j + \\ &\beta_3 Female\,Student_i \times Female\,Prof_j + \alpha^{'} X_i + \gamma^{'} Z_j + \delta_c + \epsilon_{ijc} \end{aligned} \tag{1}$$

where y_{ijc} represents the outcome of student i matched with professor j in (first-year) course c. We refer to each course as a specific year-program-course combination. Hence, for simplicity in notation, we omit the subscripts for school, program, and time since they are implicit for each given student-professor-course combination.

The outcome y_{ijc} is represented by two alternative measures. One is the final grade on the exam in course c obtained by student i ranging between 0-31, the highest being the *cum laude* grade. The second outcome variable is defined as a dummy indicating whether the exam was passed on the first attempt with a grade above the median of the course (henceforth referred to as *Success on first attempt*, for brevity). Students can retake the same exam indefinitely, either due to failure or dissatisfaction with their grade. As a result, the number of attempts may reflect both low achievement and high ambition. To isolate successful outcomes, we define a dummy variable equal to 1 if the exam is passed on the first attempt with a high grade. We avoided setting an arbitrary fixed grade benchmark to define good performance, as grading standards vary across courses. Instead, we use course-specific median benchmarks, considering a grade above the median as high. This also allows us to introduce more variability to the outcome variable as the overall probability of passing the exam on the first attempt is above 85% in the sample.

The dummy variables $Female\,Student_i$ and $Female\,Prof_j$ identify, respectively, the female

[12] In Table A1 in Appendix B, we report the estimates when considering all exams without restricting the sample to channeled courses with mixed-gender professors only. The results differ since in that setting we no longer isolate the selection bias, evidenced in the results of the balance test reported in Figure A1 in Appendix A, and in a regular OLS fixed-effects setting elicit the exposure effect, conflating student and teacher individual characteristics with same-gender exposure effect.

gender of the student and of the instructor/professor that teaches the course c. Student characteristics - such as high-school background, nationality, etc. - as well as professor characteristics are included in the vectors X_i and Z_j, respectively. Course fixed effects are captured by δ_c, and they allow to exploit within the course - i.e., between sections - variability, while ϵ_{ijc} represents the error term.

The coefficient β_1 can be interpreted as the difference between female and male students' performance when assigned to a male professor section. Analogously, β_2 measures the difference between having a female vs. male professor for male students. Finally, β_3, our primary parameter of interest, can be interpreted as the variation in the gap between female and male students' performance when switching to a female-taught section. Alternatively, it represents how the female professor effect on performance changes between female and male students.

Our identifying assumption is that within courses divided into multiple sections with professors of mixed gender, student characteristics - both observables and nonobservables - are balanced between sections (see Figure 5). Hence, any systematic difference in performance between sections can be entirely attributed to the gender of the instructor. Our coefficient of interest, β_3, should then capture role model effects, as well as preferential treatments by professors with respect to students of the same gender. Concerning this matter, as previously suggested in Hoffmann and Oreopoulos (2009), the focus on large first-year undergraduate courses should allow for more proper isolation of role model effects since, at that stage, students do not typically receive differential treatment from professors, and professors often rely on teaching assistants to grade students' exams – but that do not lecture regularly. To tackle this concern further, we also include professor fixed effects in our final model specification.

Beyond this *short-term* effect, we also examine the *long-term* implications of such exposure. The corresponding empirical approach and model specification are detailed in Subsection 5.3. Specifically, we analyze outcomes such as the pace of progression through the program (measured by accumulated credits) and the likelihood of dropping out of the degree program, thereby providing a broader understanding of the sustained influence of female faculty on female students' academic success.

5 Results

In what follows, in Subsection 5.1, we first discuss the results relative to our quasi-experimental setting, looking at female student-professor interaction effects on final exam grade and success on the first attempt. In Subsection 5.2, we explore the heterogeneity of the results across subsamples. Importantly, in Subsection 5.3, we discuss the empirical framework and results concerning the long-term effects of exposure to female professors. We refer to the Table A1 for a simple OLS specification that does not rely on the random allocation mechanism and perform additional checks to reinforce the interpretation of the student-professor fixed effects specification.

5.1 Evidence from the Quasi-Experimental Setting

The results of our analysis on both the *exam grade* and *success at first attempt* are presented respectively in Table 10 and Table 11, employing in each table four model specifications that range from the most parsimonious to the one including the full set of control variables and fixed effects. Following the specification of Mengel et al. (2018), each model includes course-fixed effects in order to exploit within-course variability, and standard errors are clustered at the section level. Additionally, professor-fixed effects are added in column (4) in order to account for unobserved leniency in grading standards, which might be correlated with their gender - or with one of their students.

Focusing on final exam grade, the β_1 coefficient across all specifications of Table 10 shows that, on average, female students achieve lower grades compared to their male counterparts in sections taught by a male professor. The size of this negative gap is between 5% to 6% of a standard deviation in final exam grades and is strongly significant. The different gender of instructors does not seem to significantly alter the final grades of male students, as suggested by the estimated β_2 coefficient. The presence of female professors seems to slightly worsen their grades, yet these effects are not of practical relevance - at most 0.043 grade less, in a context in which grades vary by units - and are estimated to be statistically insignificant.

Table 10: Quasi-Experiment Estimates - Final Exam Grade

	(1)	(2)	(3)	(4)
Female student (β_1)	-0.229***	-0.245***	-0.256***	-0.191**
	(0.085)	(0.086)	(0.086)	(0.084)
Female prof (β_2)	-0.031	-0.040	-0.043	
	(0.105)	(0.108)	(0.109)	
Female student × Female prof (β_3)	0.314***	0.332***	0.341***	0.207*
	(0.120)	(0.122)	(0.122)	(0.117)
Student age		-0.046***	-0.022*	-0.022*
		(0.013)	(0.012)	(0.012)
Foreign born student		-0.381***	-0.213*	-0.243*
		(0.120)	(0.123)	(0.124)
Prof age		-0.010*	-0.009	
		(0.006)	(0.006)	
Foreign born prof		0.058	0.046	
		(0.233)	(0.227)	
Classical HS			-0.093	-0.100
			(0.066)	(0.067)
Technical HS			-1.092***	-1.112***
			(0.092)	(0.090)
Linguistic HS			-0.731***	-0.729***
			(0.122)	(0.122)
Foreign HS			-1.571***	-1.525***
			(0.271)	(0.277)
Other HS			-1.163***	-1.160***
			(0.118)	(0.117)
Constant	25.135***	26.647***	26.318***	25.798***
	(0.073)	(0.392)	(0.387)	(0.248)
Mean of Dep. Variable	25.09	25.12	25.12	25.12
SD of Dep. Variable	3.89	3.88	3.88	3.88
Course FE	Yes	Yes	Yes	Yes
Professor FE	No	No	No	Yes
Observations	25964	25388	25388	25385
($\beta_1 + \beta_3$)	0.085	0.087	0.085	0.016
P-value	0.298	0.295	0.314	0.848

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Estimates are based on Sapienza administrative data and refer to exams subject to the channeling process with mixed-gender professors and passed on time in the first year of all STEM bachelor degrees. The outcome in each model is the final grade obtained in the exam, which can range between 18 and 31. The reference category for the high-school type is Scientific HS. Course fixed effects refer to each year-program-course combination. Standard errors are clustered at the section level.

Finally, the estimated β_3 indicates how the gap between female and male students significantly reduces in sections taught by a female professor. The estimated role model effects allow female students to completely close the gap with their male peers, as shown in the lower panel of the table by the sum $(\beta_1 + \beta_3)$, up to the point that this gap becomes positive, indicating a slightly better performance of female students with respect to males in female taught sections. Nonetheless, the gap becomes statistically insignificant. The size of the role model effect is estimated to be around 5% to 8% of a standard deviation in final exam grades.

The inclusion of additional controls in columns (2) and (3) does not alter the main conclusion at hand while slightly inflating the estimated coefficients. Interestingly, professors' demographic characteristics do not seem to alter students' outcomes in a significant way, except for a weak negative relation detected between professor age and final exam grade in column (2).

On the other hand, most students' traits are significantly correlated with their outcomes. Students' age appears to be negatively linked to their grades, which could be explained by the fact that students enrolling later might do so because they took more years to graduate from high school or because they changed programs. Also, being born in a country different than Italy appears to negatively influence final exam results.

Nonetheless, high school background appears to be the strongest determinant of student performance. Using Scientific high school as the reference category, we find that all other backgrounds—except for Classical high school—are negatively and significantly associated with the outcome. These effects are most pronounced for students from foreign high schools, who receive, on average, grades that are 1.6 points lower than their peers from Italian Scientific high schools. This highlights the challenges foreign students may face in adapting to different teaching styles and language barriers.

Even though the inclusion of professor fixed effects in column (4) shrinks the coefficients of interest, reducing the significance level of both β_1 and β_3, our main conclusions remain unaltered. Moreover, it is important to consider whether some professors' unobservable time-invariant factors, such as professor ability and teaching style, should be considered transmission channels of the role model effects rather than confounders. If they function as transmission channels, allowing them to vary across professors rather than controlling for them would provide a more accurate representation of the overall effect of interest. Conversely, if these characteristics are regarded as potential confounding factors, such as different grading standards, incorporating professor-fixed effects enables us to obtain cleaner results.

Next, we investigate the probability of passing the exam on the first attempt. As explained earlier, to distinguish between retakes due to failure and successful first attempts, we define the outcome variable as 1 for first attempts resulting in a grade above the course median.

The results presented in Table 11 are perfectly in line with those presented just above and indicate analogous dynamics. While the probability of success is almost 5 percentage points lower for female students compared to their male peers in male-taught sections, this gap is almost entirely closed in female-taught sections, where β_3 is estimated to be between 3 and 4 percentage points - which accounts for slightly less than 10% of the average probability in the sample. The estimated gap between female and male students in female-taught sections ($\beta_1 + \beta_3$) is negative but statistically insignificant. Also, in this case, male students do not seem to be influenced in a meaningful manner by the gender of their instructor. As for professor traits, the negative relation between professor age and students' performance is now significant at the 10% level, also in column (3), while there is no evidence of meaningful links between students' age and their own performance. High school background and place of birth remain the most critical factors influencing a student's probability of success. The inclusion of professor fixed effects is accompanied by the same remarks previously mentioned.

These findings underscore the nuanced dynamics of gender interactions in higher education, revealing that while female students tend to underperform, on average, when assigned to male professors, exposure to female professors in first-year courses can exert a positive and significant impact on their academic outcomes, allowing them to bridge the gap with their male peers. This result carries substantial implications, as performance in first-year courses serves as a strong predictor of both degree completion and broader academic success, as discussed in the following. This evidence highlights the potential role of female instructors as influential role models, fostering improved academic achievement and persistence among female students during the critical transition into STEM higher education.

Table 11: Quasi-Experiment Estimates - Success on First Attempt

	(1)	(2)	(3)	(4)
Female student (β_1)	-0.046***	-0.048***	-0.049***	-0.043***
	(0.011)	(0.012)	(0.012)	(0.011)
Female prof (β_2)	0.010	0.010	0.010	
	(0.013)	(0.013)	(0.013)	
Female student × Female prof (β_3)	0.037**	0.039**	0.040**	0.027*
	(0.016)	(0.016)	(0.016)	(0.016)
Student age		-0.002	0.001	0.001
		(0.002)	(0.002)	(0.002)
Foreign born student		-0.058***	-0.040**	-0.043***
		(0.015)	(0.016)	(0.016)
Prof age		-0.001**	-0.001*	
		(0.001)	(0.001)	
Foreign born prof		-0.024	-0.025	
		(0.035)	(0.035)	
Classical HS			-0.014	-0.016*
			(0.009)	(0.009)
Technical HS			-0.125***	-0.129***
			(0.013)	(0.013)
Linguistic HS			-0.076***	-0.077***
			(0.017)	(0.017)
Foreign HS			-0.162***	-0.158***
			(0.035)	(0.036)
Other HS			-0.149***	-0.150***
			(0.016)	(0.016)
Constant	0.430***	0.552***	0.517***	0.449***
	(0.009)	(0.051)	(0.050)	(0.032)
Mean of Dep. Variable	0.42	0.42	0.42	0.42
SD of Dep. Variable	0.49	0.49	0.49	0.49
Course FE	Yes	Yes	Yes	Yes
Professor FE	No	No	No	Yes
Observations	25988	25410	25410	25407
($\beta_1 + \beta_3$)	-0.009	-0.009	-0.009	-0.015
P-value	0.384	0.403	0.413	0.156

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Estimates based on Sapienza administrative data refer to exams subject to the channeling process with mixed-gender professors and passed on time in the first year of all STEM bachelor degrees. The outcome in each model is a dummy variable which takes value 1 if the exam has been passed on the first attempt with a grade above the median of the course. The reference category for the high-school type is Scientific HS. Course fixed effects refer to each year-program-course combination. Standard errors are clustered at the section level.

Finally, one might consider applying individual student fixed effects to account for personal characteristics that may be simultaneously correlated with both the student's and the professor's gender. While conceptually appealing, this approach comes with a series of limitations. Since we focus on first-year students and exams taken on time in mixed-gender courses, we rarely observe students obtaining multiple grades. Specifically, first-year students typically take between 1 and 8 exams, with a median of 5. However, when restricting the sample to exams in mixed-gender sections, the range narrows to 1–6 exams, with a median of only 2, and 79% of students complete no more than three exams. Imposing student fixed effects reduces the sample size by approximately 22%, notably excluding students who took only one exam, likely lower achievers who, according to our heterogeneity analysis, could represent the group benefiting the most from same-gender professors. Results relating to this restricted sample are reported in Appendix B, and it is possible to see how the effects of interest become non-significant also across all specifications of the model which do not include students' fixed effects.

As a final remark concerning all results, clustering standard errors at different levels does not alter our conclusions in any notable way.

5.2 Heterogeneous effects

We conduct a series of subsample analyses to verify whether the heterogeneity in the results across different types of individuals aligns with the notion that we are capturing role model effects.

Focusing on the sole coefficient of the interaction term of interest describing the differential effect of female student exposure to female professors, Table 12 reveals interesting patterns relative to the final grade obtained at the exam. In particular, role model effects do not appear to have had a significant impact during the COVID-19 pandemic. The COVID setting serves as a placebo test, as the distancing measures implemented in Italy during that period forced students to attend lectures remotely. The lack of actual contact with female professors has likely weakened the role model effects for female students (second panel), which, on the other hand, are estimated to be much stronger in periods without the remote learning option (top panel).

Table 12: Interaction Term Across Different Subsamples - Final Exam Grade

	Covid Years	Young Prof.	Low GPA Stud.	Scientific HS
No	0.315**	0.009	0.185*	0.360
	(0.148)	(0.165)	(0.107)	(0.250)
Observations	14776	13083	12870	5467
Yes	0.052	0.385**	0.257**	0.153
	(0.191)	(0.166)	(0.125)	(0.128)
Observations	10609	12297	12501	19906

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the estimates based on Sapienza administrative data of the interaction term (β_3) in our quasi-experimental setting. The outcome in each model is the final grade obtained in the exam, which can range between 18 and 31. Estimates are based on separate models, where students are divided into subsamples according to: *i* exposure outside (top/**No**) vs. during (bottom/**Yes**) the COVID pandemic period; *ii* exposure to professors with age above (top/**No**) vs. below (bottom/**Yes**) the median; *iii* exposure for students with first-year GPA above (top/**No**) vs. below (bottom/**Yes**) the median; *iv* exposure for students without (top/**No**) vs. with (bottom/**Yes**) scientific high school background. Each model accounts for course and professor fixed effects. Standard errors are clustered at the section level.

One additional piece of evidence is linked to the age of the instructor. To construct the age cutoff, we first calculate the average age of each professor over time and use the median of these averages as the threshold. Professors with an average age below this median are classified as young. The resulting threshold is slightly below 54 years old, which may not seem particularly young but is consistent with the relatively late retirement ages for scholars and the relatively slow career process in Italy. According to the second column of Table 12, the role model effects seem to act in a significant manner only for young professors (second panel), coherently with the idea that female students might respond more strongly to female role models when they perceive them as closer and more relatable.

Additionally, the interaction term is estimated to be higher for students with low GPA, namely below the median GPA in the first year, indicating how female role model effects might act in a more powerful way among female students who are facing difficulties during their first year in the program. To reinforce this opinion, the estimated interaction term is more than twice as large for students without a scientific or technical background (first panel). However, this effect is not statistically significant due to the lower sample size with respect to the other cases, which leads to inflated standard errors.

The broad picture is the same when turning the attention to the probability of success, as

Table 13: Interaction Term Across Different Subsamples - Success on First Attempt

	Covid Years	Young Prof.	Low GPA Stud.	Scientific HS
No	0.036*	0.005	0.041**	0.039
	(0.020)	(0.021)	(0.019)	(0.033)
Observations	14792	13090	12887	5475
Yes	0.015	0.049**	0.017	0.022
	(0.025)	(0.023)	(0.015)	(0.017)
Observations	10615	12312	12506	19920

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the estimates based on Sapienza administrative data of the interaction term (β_3) in our quasi-experimental setting. The outcome in each model is a dummy variable which takes value 1 if the exam has been passed on the first attempt with a grade above the median of the course. Estimates are based on separate models, where students are divided into subsamples according to: *i* exposure outside (top/**No**) vs. during (bottom/**Yes**) the COVID pandemic period; *ii* exposure to professors with age above (top/**No**) vs. below (bottom/**Yes**) the median; *iii* exposure for students with first-year GPA above (top/**No**) vs. below (bottom/**Yes**) the median; *iv* exposure for students without (top/**No**) vs. with (bottom/**Yes**) scientific high school background. Each model accounts for course and professor fixed effects. Standard errors are clustered at the section level.

reported in Table 13. The effect is not distinguishable from zero during the years of remote teaching due to COVID. A similar pattern to Table 12 is observed for the stronger role-model effect of young professors, as well as for students with a non-scientific high school background. The only notable difference emerges for the interaction term for students with a high GPA, for whom the effect of exposure to female professors appears stronger. This discrepancy may stem from a definitional issue, as the outcome is defined as the probability of passing the exam on the first attempt with a grade above the course median. Therefore, it is more likely to find students meeting this requirement among those with a high GPA.

5.3 K-year-ahead effects

Although the immediate benefits of exposure to female professors may seem promising, there is a possibility that these effects may dissipate once that exposure ends, resulting in a lack of persistence in positive outcomes. To investigate this limitation, we extend our empirical framework to a descriptive follow-up exercise, in order to examine how first-year exposure to a female professor is related to aggregate academic results in both the current (indexed by $k = 0$) and subsequent academic years (indexed by $k = 1, 2$). For each student, we estimate the following empirical model:

$$
\begin{aligned}
y_{ikp} = & \beta_{0,k} + \beta_{1,k} Female\,Student_i + \beta_{2,k} Female\,Prof_i^0 + \\
& \beta_{3,k} Female\,Student_i \times Female\,Prof_i^0 + \alpha' X_i + \delta_p + \epsilon_{ikp}
\end{aligned}
\tag{2}
$$

The outcome variable y_{ikp} is no longer an exam-level measure, but rather a student-specific aggregate measure capturing the academic performance of student i in his k-th career year in program p. We refer to each program as a specific year-program combination. For simplicity in notation, subscripts for school (s) and year (t) are omitted, since they are implicit for each student-career year-program combination.

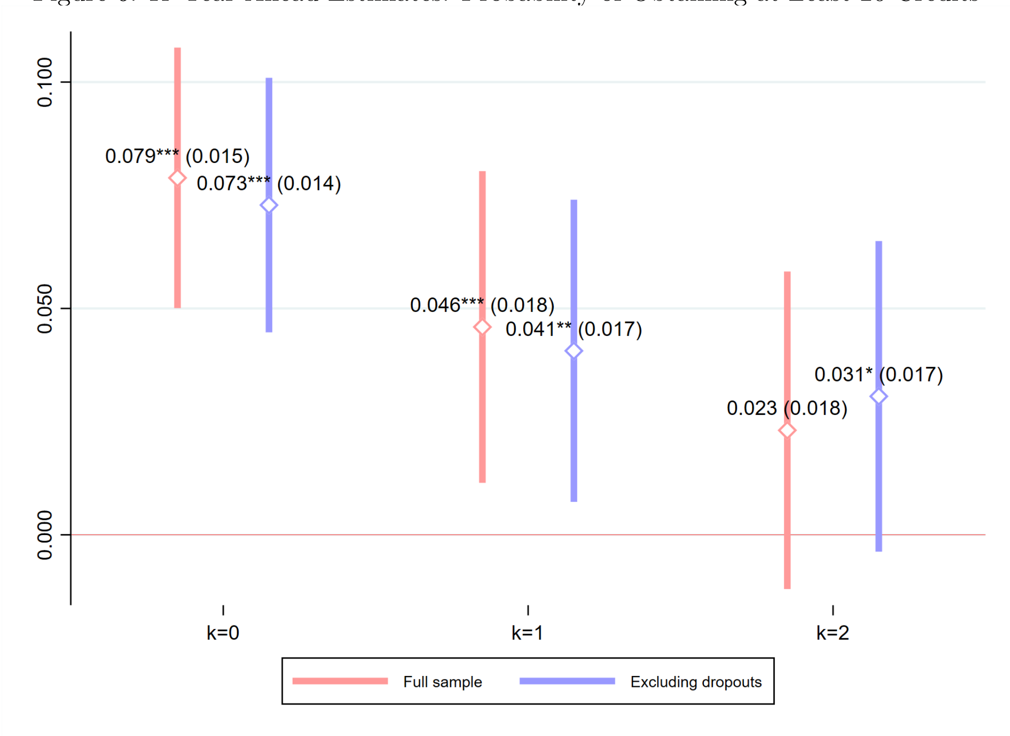
The outcomes considered are: (a) the probability of maintaining program progression, defined as acquiring at least 20 credits in each academic year, and (b) the probability of dropping out at any point during the three-year program. The dummy variable $Female\,Student_i$ retains the same definition outlined in Section 4, taking a value of 1 if student i is female.

Conversely, the dummy $Female\,Prof_i^0$ captures the quasi-random exposure of student i to a female professor in any first-year course where multiple sections are taught by professors of different genders. The dummy takes value 0 when the student was quasi-randomly assigned to a male-taught section. Hence, in this part of the analysis, we only investigate the entire careers of students who have taken at least one exam subject to the *channeling* process during the first year, where the sections were taught by professors of different sexes.

Student characteristics are included in the vector X_i. Program fixed effects are captured by δ_p since, at the aggregate level, it is no longer possible to identify the specific courses and include their relative fixed effects. The inclusion of δ_p allows us to compare students within the same program cohort. Ultimately, ϵ_{ikp} represents the error term. Our primary parameters of interest are the coefficients $\beta_{3,k}$, which quantify how first-year female role model exposure is related to female students' investigated outcomes in the same ($k = 0$) and subsequent two years ($k = 1, 2$).

Figure 6 presents the results relative to the probability of acquiring at least 20 credits in the year of exposure and in the subsequent two years. For each period, the figure provides two sets of estimates: the first includes all students in the selected sample, while the second excludes those who drop out during that year. This is done in order to distinguish whether the accumulated credits only work as a predictor of a future dropout or whether they are affected by role model effects even when focusing solely on students who persist in their academic careers.

Figure 6: K-Year-Ahead Estimates: Probability of Obtaining at Least 20 Credits



Note: The estimates are based on Sapienza administrative data. The figure illustrates the interaction term ($\beta_{3,k}$) between the female student dummy and a dummy for having been exposed to a female professor in at least one first-year course with mixed-gender channeling. The dependent variable in each model is a binary indicator equal to 1 if the student attains at least 20 credits in the *(k+1)*-th program year. For each period k, the first estimate (red) pertains to all students, while the second (blue) excludes those who drop out during the k-th period. Program fixed effects are included, as well as controls for student's age, country of birth, and high school background. Standard errors are clustered at the student level.

The findings highlight that the increased likelihood of meeting the minimum satisfactory pace of acquiring 20 credits is not solely attributable to students who eventually drop out and take fewer exams. Instead, the positive relationship persists among students who remain enrolled in the program. As reported in the extended results of Table A4 in the appendix, the quasi-random exposure to female professors is beneficial for both male and female students's probability of meeting the 20 credits threshold. Nonetheless, females benefit more strongly from this exposure with respect to males and gain 7.9 additional percentage points in the year of exposure and around 4.6 additional percentage points in the immediately following year. These effects amount to around 9% and 5% of the average probability of meeting the threshold in the respective samples. Female role model effects dissipate over time and approach the threshold of statistical insignificance in the second year post-exposure (corresponding to the third year of the program).

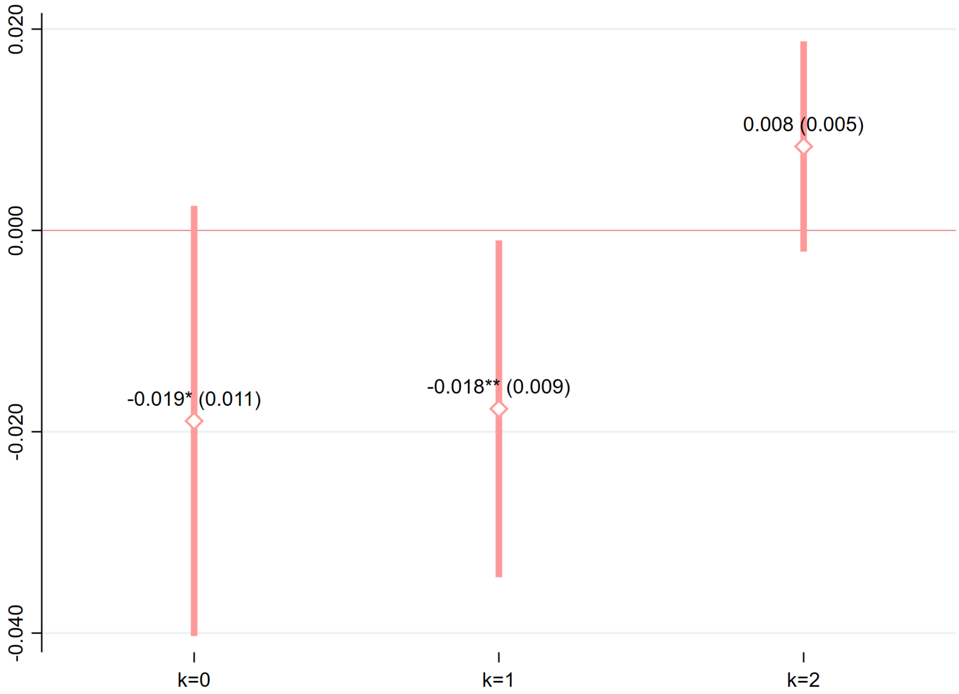
A similar conclusion can be drawn when focusing on the dropout probability, as shown in Figure 7 and the relative extended results in Table A5. Once more, exposure to female role models is beneficial - i.e. is associated to reduced dropout rates - for both male and female students, but female students are the ones benefiting relatively more. The interaction term indicates an additional reduction in dropout probability of around 2 percentage points for females, which accounts for almost 25% of the average probability in the sample. The estimated coefficient is only significant at the 10% level in the year of exposure, while it remains stable and gains statistical significance thanks to the reduced standard errors in the second year. Once more, role model effects seem to dissipate in the third year of the program.

We also constructed and tested additional outcome measures, such as GPA, the probability of acquiring at least 40 credits, and the probability of graduating on time. We do not find any statistically significant results for these measures, suggesting that female exposure does not appear to have a meaningful impact on these outcomes, both in the first year and in subsequent years of the program. When focusing on the 40 credit threshold — an indicator of students' good performance — we retrieve a similar temporal pattern in the magnitudes of the coefficient estimates, but the effect is much less pronounced. This is in line with the idea that the observed role model effects are helpful in ensuring a baseline level of academic progression rather than driving top-tier performance. These results are omitted for brevity but are available upon request.

6 Conclusion

The gender bias in STEM programs characterizes both the demand and the supply side since both female students and professors are underrepresented. In this paper, we investigate whether the teacher-learner interaction in the early stage of STEM programs between female subjects could boost the performance of female students in different dimensions.

Figure 7: Effect of first-year exposure to female professors on female students' dropout rate: k-year horizon



Note: The estimates are based on Sapienza administrative data. The figure illustrates the interaction term ($\beta_{3,k}$) between the female student dummy and a dummy for having been exposed to a female professor in at least one first-year course with mixed-gender channeling. The dependent variable in each model is a binary indicator equal to 1 if the student drops out in the *(k+1)*-th program year. Program fixed effects are included, as well as controls for student's age, country of birth, and high school background. Standard errors are clustered at the student level.

We take advantage of the pseudo-random course sectioning at Sapienza (referred to as *channeling*) due to the large enrollment in the first years of STEM bachelor programs, where students are assigned to different sections of the same course according to the first letter of their last name, hence orthogonally with respect to their gender.

In the first part of the analysis, we concentrate on the direct effects of exposure to professors of different genders, exploring how students' performance varies in our quasi-experimental setting. The interaction between female students and female professors in first-year classes contributes positively to various measures of performance. Depending on the adopted model specification, our results indicate a positive effect on final exam grades, with a magnitude varying between 5% and 8% of one SD and an increase in the probability of success on the first attempt by slightly less than 10% of the average value in the sample.

Importantly, these role model effects allow female students to close the gap with their male peers - which appears to be present in male-taught sections - without harming male students' performance. In addition, we evaluate whether this first-year exposure could also have indirect and long-lasting effects on students' careers. Our findings show how the presence of female professors is beneficial for both male and female students. Still, female students benefit significantly more than males in terms of increased probability of obtaining sufficient credits (20) and reduced dropout probability.

According to our evidence, an increase in the presence of female professors could trigger a virtuous effect on STEM female students. This should be acknowledged during professor recruiting processes, considering that hiring female professors who can also teach in first-year courses may have a positive effect on the performance of female students in STEM. Additionally, another cost-effective policy suggestion is to generally foster female instructors' presence in first-year classes, i.e. anticipating female-taught courses to earlier years when compatible with the program progression. These measures could contribute to reducing gender stereotypes for male peers but especially benefit female students in their university achievements and to close the gender pay gap in the labor market in the future, as reported by Bovini et al. (2024) and Arellano-Bover et al. (2024).

References

Arellano-Bover, J., N. Bianchi, S. Lattanzio, and M. Paradisi (2024). One cohort at a time: A new perspective on the declining gender pay gap. Technical report, National Bureau of Economic Research.

Assolombarda, C. S., O. T. Venture, and STEAMiamoci (2020). Osservatorio Talents Venture e STEAMiamoci sul Gender Gap nelle facoltà STEM. *Ricerca Talents Venture 03/2020*.

Bettinger, E. P. and B. T. Long (2005). Do faculty serve as role models? the impact of instructor gender on female students. *American Economic Review 95*(2), 152–157.

Bovini, G., M. De Philippis, and L. Rizzica (2024). The origins of the gender pay gap: education and job characteristics. *Available at SSRN 5021036*.

Breda, T., J. Grenet, M. Monnet, and C. Van Effenterre (2023). How Effective are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools. *The Economic Journal 133*(653), 1773–1809.

Carrell, S. E., M. E. Page, and J. E. West (2010). Sex and Science: How Professor Gender Perpetuates the Gender Gap. *The Quarterly Journal of Economics 125*(3), 1101–1144.

Commission, E., D.-G. for Research, and Innovation (2022). *Bridging the gender gap in STEM – Strengthening opportunities for women in research and innovation*. Publications Office of the European Union.

de Gendre, A., J. Feld, N. Salamanca, and U. Zölitz (2023). Same-sex role model effects in education. Working Paper 438, Zurich.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources XLII*(3), 528–554.

Hoffmann, F. and P. Oreopoulos (2009). A professor like me: The influence of instructor gender on college achievement. *Journal of human resources 44*(2), 479–494.

Kali Pal, K., K. Piaget, and S. Zahidi (2024). *Global Gender Gap Report 2024*. World Economic Forum.

Lim, J. and J. Meer (2017). The Impact of Teacher–Student Gender Matches: Random Assignment Evidence from South Korea. *The Journal of Human Resources 52*(4), 979–997.

Mansour, H., D. I. Rees, B. M. Rintala, and N. N. Wozny (2022). The effects of professor gender on the postgraduation outcomes of female students. *ILR Review 75*(3), 693–715.

Mengel, F., J. Sauermann, and U. Zölitz (2018, 02). Gender bias in teaching evaluations. *Journal of the European Economic Association 17*(2), 535–566.

Muralidharan, K. and K. Sheth (2016). Bridging education gender gaps in developing countries: The role of female teachers. *Journal of Human Resources 51*(2), 269–297.

Porter, C. and D. Serra (2020). Gender Differences in the Choice of Major: The Importance of Female Role Models. *American Economic Journal: Applied Economics 12*(3), 226–254.

Price, J. (2010). The effect of instructor race and gender on student persistence in stem fields. *Economics of Education Review 29*(6), 901–910.

Riise, J., B. Willage, and A. Willén (2022). Can Female Doctors Cure the Gender STEMM Gap? Evidence from Exogenously Assigned General Practitioners. *The Review of Economics and Statistics 104*(4), 621–635.

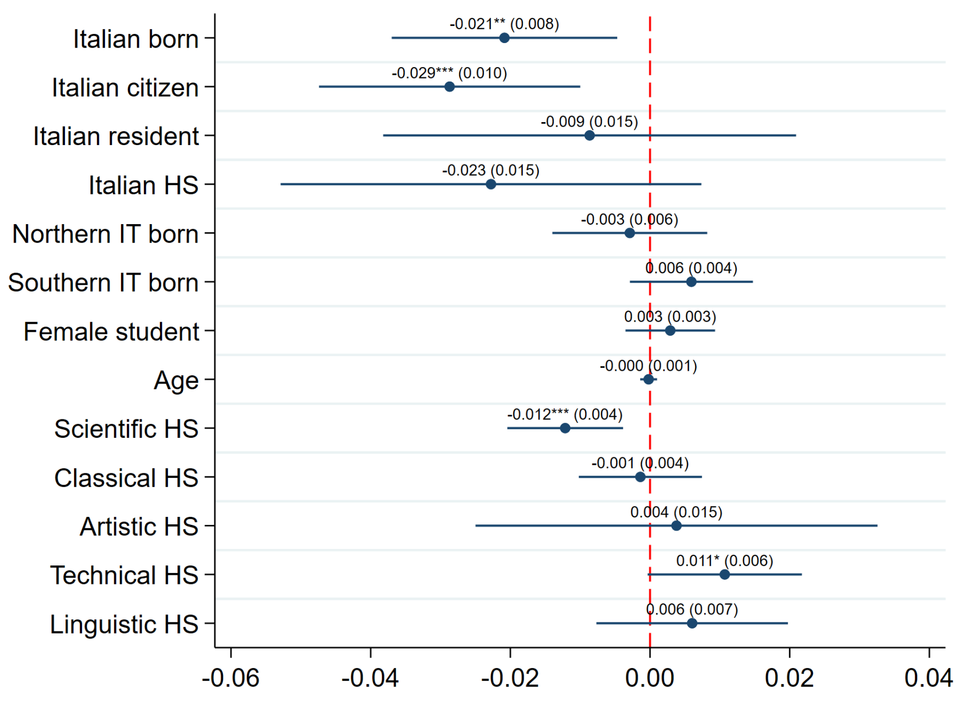
Winters, M. A., R. C. Haight, T. T. Swaim, and K. A. Pickering (2013). The effect of same-gender teacher assignment on student achievement in the elementary and secondary grades: Evidence from panel data. *Economics of Education Review 34*, 69–75.

Appendix

A Baseline Evidence

When there is no channeling, or in cases in which multiple sections of a course are assigned to professors of the same sex, comparing students in the same course whose only difference is the assigned section is not possible if we intend to investigate same-sex professor effects. Hence, we need to rely on the within program - i.e., between courses - variability in order to estimate the effects of interest. However, as shown in Figure A1, significant concerns arise when looking at the characteristics of students in courses taught by female professors, which appear to be systematically different than those of students in male-taught courses, indicating that self-selection dynamics are likely at play and could also affect unobservable variables.

Figure A1: Balancing test - other first-year courses



Note: The figure is based on Sapienza administrative data and presents the results of a balancing test, where we regress the gender of the instructor ($female = 1$) on pre-determined student characteristics. The sample includes all first-year exams taken on time in courses with professors of a single sex, hence deviating from our quasi-experimental setting. Each coefficient estimate originates from a separate model, which includes program-fixed effects. Standard errors are clustered at the section level.

For this reason, we tried to see how the results of the analysis change when we only include program fixed effects, hence exploiting between courses variability, both in the full sample - columns (1) to (3) of Table A1 - and in a restricted sample that excludes the exams complying with our quasi-experimental setting - columns (4) to (6).

Table A1: Baseline Estimates - Final Exam Grade

	(1)	(2)	(3)	(4)	(5)	(6)
Female student (β_1)	-0.205***	-0.216***	-0.215***	-0.150***	-0.155***	-0.147***
	(0.046)	(0.047)	(0.047)	(0.051)	(0.051)	(0.052)
Female prof (β_2)	-0.242***	-0.244***	-0.242***	-0.203*	-0.248**	-0.239**
	(0.086)	(0.086)	(0.087)	(0.117)	(0.117)	(0.118)
Female student × Female prof (β_3)	0.438***	0.470***	0.475***	0.426***	0.449***	0.447***
	(0.085)	(0.085)	(0.085)	(0.101)	(0.101)	(0.101)
Student age		-0.022***	-0.004		-0.018**	-0.000
		(0.007)	(0.006)		(0.008)	(0.007)
Foreign born student		-0.769***	-0.407***		-0.908***	-0.480***
		(0.070)	(0.068)		(0.083)	(0.080)
Prof age		-0.012***	-0.012***		-0.010**	-0.010**
		(0.004)	(0.004)		(0.005)	(0.005)
Foreign born prof		-0.051	-0.037		-0.280	-0.269
		(0.255)	(0.254)		(0.278)	(0.277)
Classical HS			-0.187***			-0.238***
			(0.041)			(0.052)
Technical HS			-0.933***			-0.915***
			(0.052)			(0.061)
Linguistic HS			-0.824***			-0.898***
			(0.070)			(0.084)
Foreign HS			-1.913***			-1.972***
			(0.142)			(0.162)
Other HS			-1.082***			-1.084***
			(0.068)			(0.079)
Constant	25.148***	26.301***	26.144***	25.093***	26.079***	25.966***
	(0.053)	(0.263)	(0.259)	(0.061)	(0.311)	(0.307)
Mean of Dep. Variable	25.05	25.06	25.06	25.03	25.04	25.04
SD of Dep. Variable	3.94	3.94	3.94	3.96	3.96	3.96
Program FE	Yes	Yes	Yes	Yes	Yes	Yes
Exclude Quasi-Experiment	No	No	No	Yes	Yes	Yes
Observations	86623	84546	84546	60635	59134	59134
($\beta_1 + \beta_3$)	0.233	0.254	0.260	0.276	0.294	0.301
P-value	0.000	0.000	0.000	0.000	0.000	0.000

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Estimates are based on Sapienza administrative data and refer to exams passed on time in the first year of all STEM bachelor degrees. The outcome in each model is the final grade obtained in the exam, which can range between 18 and 31. The reference category for high-school type is Scientific HS. Program fixed effects refer to each year-program combination. Columns (1) to (3) refer to all exams, while in columns (4) to (6) the exams consistent with our quasi-experimental setting - multiple sections with professors of mixed gender - are excluded. Standard errors are clustered at the section level. administrative data.

The results depict a completely different scenario. A *'Battle of sexes'* emerges, with female students still performing worse than their male peers in male-taught courses, while the female professor's impact on male students appears to be now significantly detrimental. Additionally, the estimated role model effects become so strong that they are able to completely reverse the gap between female and male students. In female-taught courses, female students significantly outperform males, and the gap is even greater in magnitude than the one present in male-taught courses - as indicated by $(\beta_1 + \beta_3)$ - especially when excluding the exams that fall into the quasi-experimental setting.

These findings highlight the need for more proper identification since it is likely that, as previously demonstrated, students in courses taught by female professors of mixed gender are systematically different from those in the male-taught courses and that these differences are correlated with student's gender, hence biasing the results at hand.

B Restricted Sample Results - Quasi-Experiment

Table A2: Quasi-Experiment Estimates - Final Exam Grade - Restricted Sample

	(1)	(2)	(3)	(4)	(5)
Female student (β_1)	-0.167*	-0.173**	-0.188**	-0.194**	
	(0.087)	(0.087)	(0.087)	(0.087)	
Female prof (β_2)	0.168	0.170*	0.173*		
	(0.102)	(0.103)	(0.103)		
Female student × Female prof (β_3)	0.123	0.121	0.128	0.145	0.145
	(0.118)	(0.118)	(0.119)	(0.120)	(0.109)
Student age		-0.050**	-0.031	-0.029	
		(0.021)	(0.019)	(0.019)	
Foreign born student		-0.287**	-0.134	-0.189	
		(0.140)	(0.141)	(0.142)	
Prof age		-0.007	-0.007		
		(0.006)	(0.006)		
Foreign born prof		-0.120	-0.102		
		(0.245)	(0.241)		
Classical HS			-0.022	-0.029	
			(0.072)	(0.073)	
Technical HS			-0.931***	-0.956***	
			(0.105)	(0.102)	
Linguistic HS			-0.621***	-0.613***	
			(0.151)	(0.151)	
Foreign HS			-1.762***	-1.718***	
			(0.302)	(0.305)	
Other HS			-1.128***	-1.125***	
			(0.143)	(0.143)	
Constant	25.434***	26.844***	26.572***	26.258***	25.433***
	(0.073)	(0.505)	(0.491)	(0.380)	(0.032)
Mean of Dep. Variable	25.47	25.47	25.47	25.47	25.47
SD of Dep. Variable	3.83	3.83	3.83	3.83	3.83
Course FE	Yes	Yes	Yes	Yes	Yes
Professor FE	No	No	No	Yes	Yes
Student FE	No	No	No	No	Yes
Observations	19842	19842	19842	19842	19842
($\beta_1 + \beta_3$)	-0.044	-0.052	-0.060	-0.049	0.145
P-value	0.582	0.513	0.467	0.562	0.182

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table is based on Sapienza administrative data and it replicates the results of Table 10 on the restricted number of observations from columns (5) referring to students taking multiple exams in the sample. All remarks concerning the replicated results apply here as well.

Table A3: Quasi-Experiment Estimates - Success on First Attempt - Restricted Sample

	(1)	(2)	(3)	(4)	(5)
Female student (β_1)	-0.048***	-0.048***	-0.050***	-0.052***	
	(0.013)	(0.013)	(0.013)	(0.013)	
Female prof (β_2)	0.029**	0.029**	0.030**		
	(0.013)	(0.013)	(0.013)		
Female student $\times$ Female prof (β_3)	0.016	0.016	0.017	0.021	0.014
	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)
Student age		-0.003	-0.000	-0.000	
		(0.002)	(0.002)	(0.002)	
Foreign born student		-0.057***	-0.041**	-0.047**	
		(0.018)	(0.019)	(0.019)	
Prof age		-0.001	-0.001		
		(0.001)	(0.001)		
Foreign born prof		-0.037	-0.035		
		(0.037)	(0.036)		
Classical HS			-0.005	-0.007	
			(0.010)	(0.010)	
Technical HS			-0.117***	-0.120***	
			(0.015)	(0.015)	
Linguistic HS			-0.055**	-0.055***	
			(0.021)	(0.021)	
Foreign HS			-0.163***	-0.157***	
			(0.040)	(0.040)	
Other HS			-0.159***	-0.160***	
			(0.021)	(0.021)	
Constant	0.459***	0.561***	0.526***	0.501***	0.452***
	(0.009)	(0.062)	(0.062)	(0.048)	(0.004)
Mean of Dep. Variable	0.46	0.46	0.46	0.46	0.46
SD of Dep. Variable	0.50	0.50	0.50	0.50	0.50
Course FE	Yes	Yes	Yes	Yes	Yes
Professor FE	No	No	No	Yes	Yes
Student FE	No	No	No	No	Yes
Observations	19858	19858	19858	19858	19858
($\beta_1 + \beta_3$)	-0.031	-0.032	-0.033	-0.032	0.014
P-value	0.004	0.004	0.003	0.007	0.358

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table is based on Sapienza administrative data and it replicates the results of Table 11 on the restricted number of observations from columns (5) referring to students taking multiple exams in the sample. All remarks concerning the replicated results apply here as well.

C Tables and Extended Results for K-year-ahead Effects

Table A4: Extended Results from Figure 6

	$k = 0$		$k = 1$		$k = 2$	
	Full	No Dpt.	Full	No Dpt.	Full	No Dpt.
Female student (β_1)	-0.059***	-0.057***	-0.017	-0.015	-0.007	-0.015
	(0.013)	(0.013)	(0.016)	(0.015)	(0.016)	(0.016)
Female prof (β_2)	0.129***	0.102***	0.064***	0.057***	0.039***	0.038***
	(0.010)	(0.010)	(0.013)	(0.012)	(0.012)	(0.012)
Female student × Female prof (β_3)	0.079***	0.073***	0.046***	0.041**	0.023	0.031*
	(0.015)	(0.014)	(0.018)	(0.017)	(0.018)	(0.017)
Student age	-0.008***	-0.009***	-0.005**	-0.006**	-0.012***	-0.012***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Foreign born student	0.010	0.005	-0.056**	-0.050**	-0.034	-0.042*
	(0.016)	(0.015)	(0.022)	(0.021)	(0.024)	(0.024)
Classical HS	-0.058***	-0.037***	-0.003	-0.000	0.021**	0.027***
	(0.010)	(0.009)	(0.011)	(0.010)	(0.010)	(0.009)
Technical HS	-0.037***	-0.018	-0.043***	-0.042***	-0.038**	-0.035**
	(0.012)	(0.011)	(0.015)	(0.015)	(0.016)	(0.016)
Linguistic HS	-0.058***	-0.034**	-0.039*	-0.020	-0.033	-0.039
	(0.017)	(0.016)	(0.022)	(0.021)	(0.024)	(0.024)
Foreign HS	-0.022	-0.004	-0.045	-0.049	-0.011	0.003
	(0.037)	(0.035)	(0.047)	(0.046)	(0.047)	(0.045)
Other HS	-0.080***	-0.065***	-0.039*	-0.037*	-0.053**	-0.043*
	(0.018)	(0.017)	(0.022)	(0.022)	(0.026)	(0.025)
Constant	0.936***	1.001***	0.933***	0.963***	1.149***	1.154***
	(0.041)	(0.042)	(0.052)	(0.052)	(0.067)	(0.066)
Mean of Dep. Variable	0.85	0.89	0.87	0.88	0.91	0.92
SD of Dep. Variable	0.36	0.32	0.34	0.32	0.29	0.28
Program FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13221	12207	8406	8230	6135	6077

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table refers to Sapienza administrative data and reports the extended regression output of Figure 6. The dependent variable in each model is a binary indicator equal to 1 if the student attains at least 20 credits in the $(k+1)$-th program year. For each period k, the first column pertains to all students, while the second excludes those who drop out during the k-th period. The reference category for high-school type is Scientific HS. Program fixed effects are included. Standard errors are clustered at the student level.

Table A5: Extended Results from Figure 7

	k=0	k=1	k=2
Female student (β_1)	0.018*	0.015*	-0.008*
	(0.010)	(0.008)	(0.004)
Female prof (β_2)	-0.038***	-0.009*	-0.002
	(0.007)	(0.005)	(0.005)
Female student × Female prof (β_3)	-0.019*	-0.018**	0.008
	(0.011)	(0.009)	(0.005)
Student age	-0.001	-0.000	0.000
	(0.002)	(0.001)	(0.001)
Foreign born student	-0.021*	0.005	-0.011***
	(0.011)	(0.009)	(0.003)
Classical HS	0.053***	-0.001	0.005
	(0.008)	(0.005)	(0.004)
Technical HS	0.018**	0.007	0.007
	(0.009)	(0.007)	(0.005)
Linguistic HS	0.037**	0.019	0.003
	(0.015)	(0.012)	(0.007)
Foreign HS	0.048	-0.010	0.014
	(0.030)	(0.015)	(0.020)
Other HS	0.027*	-0.001	0.017*
	(0.014)	(0.010)	(0.010)
Constant	0.115***	0.030*	0.009
	(0.032)	(0.017)	(0.015)
Mean of Dep. Variable	0.077	0.021	0.009
SD of Dep. Variable	0.266	0.143	0.097
Program FE	Yes	Yes	Yes
Observations	13221	8406	6135

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table refers to Sapienza administrative data and reports the extended regression output of Figure 7. The dependent variable in each model is a binary indicator equal to 1 if the student drops out in the $(k+1)$-th program year. The reference category for high-school type is Scientific HS. Program fixed effects are included. Standard errors are clustered at the student level.

D List of STEM BA Programs in Sapienza University of Rome

Table A6: STEM bachelor's programs in Sapienza

Field	Level	Code	Program Name
Pharmacy and Medicine	L	2	Biotechnology
Pharmacy and Medicine	L	2	Bioinformatics
Science (Math/Physics/Natural)	L	2	Agro-Industrial Biotechnology
Civil/Industrial Engineering	L	7	Environmental Engineering
Civil/Industrial Engineering	L	7	Civil Engineering
Civil/Industrial Engineering	L	7	Environmental Engineering for Sustainable Development
Information/Computer/Statistics	L	8	Electronic Engineering
Information/Computer/Statistics	L	8	Management Engineering
Information/Computer/Statistics	L	8	Information Engineering
Information/Computer/Statistics	L	8	Communication Engineering
Information/Computer/Statistics	L	8	Computer and Automation Engineering
Civil/Industrial Engineering	L	9	Environmental Engineering for Sustainable Development
Civil/Industrial Engineering	L	9	Electrical Energy Engineering
Civil/Industrial Engineering	L	9	Aerospace Engineering
Civil/Industrial Engineering	L	9	Energy Engineering
Civil/Industrial Engineering	L	9	Chemical Engineering
Civil/Industrial Engineering	L	9	Mechanical Engineering
Civil/Industrial Engineering	L	9	Clinical Engineering
Science (Math/Physics/Natural)	L	13	Biological Sciences
Architecture	L	17	Architecture Sciences
Civil/Industrial Engineering	L	23	Sustainable Building Engineering
Civil/Industrial Engineering	L	23	Techniques for Construction and Territory for the Surveyor Profession
Architecture	L	23	Construction Process Management - Project Management
Science (Math/Physics/Natural)	L	27	Chemical Sciences
Pharmacy and Medicine	L	29	Applied Pharmaceutical Sciences
Science (Math/Physics/Natural)	L	30	Physics
Information/Computer/Statistics	L	31	Computer Science
Information/Computer/Statistics	L	31	Computer Science
Information/Computer/Statistics	L	31	Applied Computer Science and Artificial Intelligence
Science (Math/Physics/Natural)	L	32	Environmental Sciences
Science (Math/Physics/Natural)	L	32	Natural Sciences
Science (Math/Physics/Natural)	L	34	Geological Sciences
Science (Math/Physics/Natural)	L	35	Mathematics
Information/Computer/Statistics	L	41	Management Statistics
Information/Computer/Statistics	L	41	Statistics, Economics, and Society
Information/Computer/Statistics	L	41	Statistics, Economics, Finance, and Insurance
Science (Math/Physics/Natural)	L	43	Technologies for the Conservation and Restoration of Cultural Heritage