Discussion Paper Series – CRC TR 224

Discussion Paper No. 633
Project A 02

Relative Grades and Gender Differences in STEM Enrollment

Larissa Fuchs[1]
Pia Pinger[2]
Philipp Seegers[3]

January 2025

[1]University of Cologne, Email: larissa.fuchs@wiso.uni-koeln.de
[2]University of Cologne, Email: pinger@uni-koeln.de
[3]Maastricht University, Email: pks@candidate-select.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Relative Grades and Gender Differences in STEM Enrollment*

Larissa Fuchs, Pia Pinger, Philipp Seegers**

December 19, 2024

Abstract

Based on novel administrative and survey data from Germany, this study investigates the importance of relative STEM performance in high school for the gender gap in STEM enrollment. We first document that males display a higher relative STEM performance than females, which however mainly emerges from females' stronger achievement in non-STEM subjects. Our findings further reveal that a one-standard-deviation increase in grade-based STEM advantage raises the likelihood of pursuing a STEM degree by approximately 19 percentage points for males, but only by half as much for females. A decomposition analysis shows that 26% of the STEM gender gap could be attributed to differences in grade-based STEM performance if major preferences resembled those of males. However, relative grades are largely unimportant in an environment where preferences mirror those of females. This suggests that STEM performance differences have limited influence on females' decisions to pursue STEM degrees. While STEM advantage significantly impacts observed gender gaps in STEM enrollment, this effect is primarily driven by males.

JEL Codes: *I21, I24, J16, J24*
Keywords: *gender gap, STEM enrollment, relative grades, ranks*

*We would like to express our gratitude to the "Qualitäts- und UnterstützungsAgentur – Landesinstitut für Schule (QUA-LiS NRW)" for their support throughout this process and for providing the necessary data. Financial support through the German Research Foundation Collaborative Research Centre TR 224 (Project A02) is gratefully acknowledged. Pia Pinger thankfully acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866.

**Larissa Fuchs: University of Cologne, larissa.fuchs@wiso.uni-koeln.de, Pia Pinger: University of Cologne, pinger@uni-koeln.de, Philipp Seegers: Research Fellow Maastricht University, pks@candidate-select.de

1. Introduction

Women are underrepresented in math-intensive fields. In Germany, only 22% of graduates in science, technology, engineering, and mathematics (STEM) are female, compared to 32% in the OECD overall. Moreover, the gender gap in STEM attendance has not decreased in most developed countries over the past decades (OECD 2024). The underrepresentation of women in STEM is concerning to the extent that STEM graduates tend to earn high wages and have above-average career prospects (e.g., Anger and Plünnecke 2022, Blau and Kahn 2017). Moreover, females shying away from STEM-related fields may limit the talent pool in occupations that are often viewed as key contributors to a country's growth and national competitiveness (Carnevale et al. 2011, Bianchi and Giorcelli 2020, Del Carpio and Guadalupe 2022).

Why is it that women are so much less likely to choose a math-intensive field of study? The underlying reasons are manifold, ranging from a difference in preferences and expectations (Zafar 2011, 2013, Wiswall and Zafar 2021, Niederle and Vesterlund 2010), via norms (Guiso et al. 2008, Nosek et al. 2009, Nollenberger et al. 2016, Del Carpio and Guadalupe 2022, Carlana 2019, Terrier 2020, Nicoletti et al. 2022), to a lack of female role models (Breda et al. 2023, Bettinger and Long 2005, Winters et al. 2013, Dee 2005, 2007, Canaan and Mouganie 2023), or peer effects (Murphy and Weinhardt 2020, Elsner et al. 2021, Elsner and Isphording 2017, Duflo et al. 2011). PISA results reveal persistent gender differences in academic performance across OECD countries (OECD 2019a,b). First, boys tend to do slightly better in STEM subjects on average (1.4% higher scores), although this difference does not suffice to explain prevailing gender disparities in math-intensive professions (Ceci et al. 2014). Second, girls outperform boys in terms of verbal abilities (6% higher scores), and they tend to outperform boys in school, overall.

As a consequence, a girl that performs well in STEM is likely to perform even better in non-STEM subjects. She thus receives more positive signals about her non-STEM abilities than about her STEM abilities compared to a boy, and may conclude that a non-STEM occupation suits her abilities best. To the extent that such perceptions lead to disparities in enrollment decisions, this could explain persistent gender differences in STEM choices and human-capital investments. Recent evidence (Breda and Napp 2019, Goulas et al. 2022) suggests that relative performance in STEM versus non-STEM can indeed have important implications for female intentions to continue an education in STEM and that relatively stricter grading policies in STEM courses might reinforce this tendency (Ahn et al. 2024). However, there remains a gap in our understanding of how these relative performance differences and their interpretation affect actual decision-making, beyond mere intentions. Our study addresses this gap by investigating three key questions: First, how does relative performance in STEM versus non-STEM subjects, measured by grades or ranks at the end of high school, influence the decision to pursue STEM-related subjects in higher education? Second, to what extent do males and females differ in their assessment of these relative performance indicators? At an aggregate level, how much of the gender gap in STEM enrollment at university can be attributed to differences in relative performance?

To answer these questions, we rely on two sources of data from Germany. First, administrative data documenting grade distributions from upper secondary education, including both overall performance and subject-specific achievement in the final exit exams. Second, survey data that contains information on background, university enrollment and performance, high school grade point averages (GPAs) of exit exams, and subjects chosen in high school exit exams. Based on these data, we construct two measures of relative STEM performance following Goulas et al. (2022). First, *grade-based STEM advantage* is calculated as the ratio of STEM over non-STEM GPA achieved in final exit exams, minus one. A grade-based STEM advantage greater than zero indicates that an individual has a higher GPA in STEM subjects than in non-STEM subjects, reflecting a relative proficiency in STEM based on grades. Second, *rank-based STEM advantage* is computed as the ratio of the school-cohort rank of STEM GPA to the school-cohort rank of non-STEM GPA, minus one. This describes an individual's relative grade position as compared to her classmates, that is, considering the individual's position within the school and the year based on their grades. Our analysis is based on a sample of 573 observations that allow us to link these measures of relative STEM advantage of upper secondary education school leavers to enrollment choices in tertiary education. We also conduct a decomposition analysis to quantify the extent to which gender differences in STEM enrollment can be attributed to variations in grade- and rank-based performance indicators.

Germany offers a perfect setting to study the effect of ability signals on human-capital investment and selection. First, Germany offers an educational landscape where over 90% of institutions, including schools and universities, are publicly funded and tuition-free. Public schools maintain exceptionally high quality, with private institutions holding no marked advantage. Consequently, financial constraints exert little influence over educational choices. Second, uniform compensation schemes for teachers and standardized curricula yield consistent educational quality across schools of a particular type. Variations in schooling levels and educational intensity primarily emanate from school tracking, which is transparent to students, parents, educators, and researchers alike. Lastly, gender-based disparities in tertiary-education outcomes are particularly persistent in Germany (OECD 2024), which may reflect substantial non-monetary variations in educational decision-making across groups.

We present two sets of results. First, we provide descriptive evidence of a significant gender gap in high school grades between STEM and non-STEM subjects. Females exhibit smaller STEM to non-STEM grade differences compared to males, which we refer to as "grade-based STEM advantage". This advantage stems from females achieving comparable grades in STEM subjects while outperforming in non-STEM subjects, aligning with existing literature (OECD 2019b, Breda and Napp 2019, Goulas et al. 2022). In our sample, we also identify a 24% gender gap in STEM enrollment in higher education programs. Our analytical findings reveal that grade-based STEM advantage increases the likelihood of choosing a STEM subject for both genders, but with a notably smaller effect for females. A one-standard-deviation increase in grade-based STEM advantage raises the probability of pursuing a STEM degree by 19 percentage points for males, but only half as much for females. A one-standard-deviation increase in rank-based STEM advantage raises the probability of pursuing a STEM degree by 4.2 percentage points for males, but there is no effect for females. Second, a decomposition of the STEM enrollment gap into relative-STEM-performance-related differences and differences in preferences reveals that if female major preferences resembled those of males, 26% of the gender gap in STEM enrollment could be attributed to disparities in grade-based performance indicators. However, rank-based performance indicators do not significantly affect the gender gap in STEM choices. In a scenario with female choice preferences, neither grade-based nor rank-based STEM performance differentials significantly influence STEM enrollment differences. We find that males are more likely than females to specialize in STEM fields if they have a relative advantage in STEM-related subjects, whether based on grades or ranks. This suggests that non-performance-related factors, such as preferences or anticipated discrimination, may discourage females from choosing STEM occupations despite positive ability signals and we provide suggestive evidence that females enrolled in STEM subjects indeed anticipate more gender-based discrimination in their future careers.

Our study contributes to the existing literature in at least four ways. First, we extend research on STEM advantage in educational decisions by examining actual choices rather than intentions (Breda and Napp 2019, Goulas et al. 2022). Second, we provide the first analysis of grade- and rank-based performance indicators' relative importance across genders in Germany, a setting, where grades are crucial for university enrollment.[1] Third, there is evidence that shows that students have imperfect knowledge of their own ability (Zafar 2011, Stinebrickner and Stinebrickner 2012, 2014, Bobba and Frisancho 2016) and are uncertain about their returns to education (Jensen 2010, Attanasio and Kaufmann 2014, Wiswall and Zafar 2015). We are able to show that in their education decisions, female students seemingly place too little weight on their relative advantage. In our decomposition exercise, we are able to delineate effects that stem from (gender) preferences to those from performance differences. This approach enables us to quantify the contribution of observed performance differences in STEM and non-STEM fields on the overall gender STEM-enrollment gap. Extending the literature (e.g., Delaney and Devereux 2019, Card and Payne 2021, Riegle-Crumb et al. 2012), we are able to add an explanation on the paradox of women selecting lower-wage non-STEM fields despite demonstrating equal or superior academic performance across disciplines. Lastly, we extend research on ability cues in decision-making (Stinebrickner and Stinebrickner 2012, Murphy and Weinhardt 2020, Elsner et al. 2021, Bond et al. 2018, Li and Xia 2024, Tan 2023) and gender differences in grade responsiveness. Prior work shows females' persistence in subjects correlates with strong performance (Owen 2010), yet they exit male-dominated and STEM fields more readily after poor performance than males do (Kugler et al. 2021, Rask and Tiefenthaler 2008). While existing studies examine absolute grades in individual subjects, we demonstrate that females respond less than males to relative performance differences across subjects. Overall, women may thus require stronger signals than males to decide for a career in STEM.

The remainder of the paper is organized as follows. In the next section, we provide information on the institutional setting, the data, measures, and descriptive statistics. In Section 3, we present the main results. Section 4 concludes.

[1] Access to tertiary education is determined by the acquisition of the high school degree only. Admission restrictions in competitive fields such as business and administration, psychology, or medicine, are generally determined by the final high school GPA.

2. Institutional Setting, Data, and Descriptive Statistics

2.1. The German School System

The German education system distinguishes itself by assigning the responsibility for education to each federal state. In this study, we investigate GPAs in central examinations from high schools situated within the federal state of North Rhine-Westphalia (NRW), using data of high school leavers between 2010 to 2019. These central exams make up an important fraction of the upper secondary degree GPA. It opens doors for future education and career paths by determining eligibility for tertiary education.

The exams are centrally provided by the federal state of NRW, aiming to enhance comparability and to ensure equitable treatment for all students. The grade information we observe in our sample stems from standardized exams across all upper secondary schools. In the final examination, students select four subjects, consisting of three written exams and one oral exam. For grading consistency, we focus exclusively on the GPAs obtained from written exams. The GPAs range from 0 to 15 points, with 15 denoting the highest grade and 0 the lowest. The minimum passing grade is 4 points. The final high school GPA, computed from these points, then ranges from 1.0 to 4.0, where 1.0 is the best grade and 4.0 the lowest one.[2] In 2005, there was a shift in the educational system from the G9 to the G8 system. The G8 system reduces time spent at at school from 13 years to 12 years. Since the first cohort participating in the G8 system finished upper secondary school in 2012, we need to account for graduation-year effects. The institutional background is presented in more detail in Section 5.1 of the Appendix.

2.2. Data

Our dataset combines survey data from the German student study "Fachkraft 2013" with administrative records of GPA distributions from NRW. The survey, conducted in March 2021 and March 2022, collected comprehensive information about students' background, university enrollment, performance, and for a subsample, detailed high school information including course selection, grades, and IQ scores. Students were recruited through a major nationwide job board platform, with

[2] See APO-GOSt in the version of 12 March 2009 [Article 1, Paragraph 20].

participation incentivized through Amazon vouchers.[3] The sample closely compares to the overall population of German students in terms of region, university type, study fields, and likelihood to hold a student job (Hemkes et al. 2016). The administrative data, obtained from the Qualitäts- und UnterstützungsAgentur – Landesinstitut für Schule (QUA-LiS NRW), covers GPA distributions from central final exams and school-leaving grades for all upper secondary schools in NRW from 2010-2019. These records include school characteristics (legal status and type) and GPA frequency distributions across 22 subjects[4], ranging from STEM fields like mathematics and physics to humanities and arts.

Merging these administrative records with our survey data yields a final sample of 573 individuals with information on high school performance and major choices in higher education.

2.3. Measures

Tertiary Education Sorting Students in Germany directly enroll for a field of study when they first enter university. We elicited the current study field as a choice out of a list of 14 majors.[5] We adopt a STEM definition that emphasizes strong quantitative rigor. For the purposes of classifying tertiary education choices, we consider the following disciplines as STEM-related: computer sciences, engineering, mathematics, chemistry, and physics.

Secondary Education and Grades We identify five subjects as STEM subjects in high school – computer sciences, mathematics, physics, and chemistry – in alignment with the STEM classification used for categorizing tertiary education choices. Since it is compulsory to take at least one subject from a STEM field, we are able to observe STEM and non-STEM GPAs for all of our sample. For simplicity, we reversed the order in our analysis such that a higher always GPA indicates better grades. To assess individual competence in STEM relative to non-STEM subjects, we follow Goulas et al. (2022). Our first measure of relative performance, *grade-based STEM advantage*, is

[3] The job board jobmensa.de is operated by Studitemps GmbH (jobvalley) and is the largest platform for student jobs. Participation was incentivized using Amazon vouchers amounting to 1,950 EUR (29 x 50, 1 x 500 vouchers).

[4] Specifically, the subjects include math, chemistry, physics, computer science, technology, German, English, French, Dutch, biology, history, geology, social sciences, Chinese, educational science, art, Latin, music, Spanish, sport, psychology, and business administration.

[5] Majors comprise educational sciences, computer sciences, engineering, art, music, mathematics, media sciences, medicine, health sciences, natural sciences, psychology, legal sciences, social sciences, humanities, sports science, linguistics, cultural studies, and economics.

based on grades and constructed for each student i in the following way:

$$\textit{Grade-based STEM advantage}_i = \frac{\textit{STEM GPA}_i}{\textit{Non-STEM GPA}_i} - 1 \tag{1}$$

A grade-based STEM advantage exceeding 0 indicates that individual i has achieved a higher GPA in STEM subjects than in non-STEM subjects, signifying a relative proficiency in STEM based on grades. A negative value would be interpreted inversely.

To construct our second measure of relative performance, *rank-based STEM advantage*, we need to construct two separate rank measures – one based on STEM GPA and another based on non-STEM GPA – to capture students' relative standing within each domain.. Since school cohorts and classes vary in size, we do not use the raw rank of students in each subject s in their school cohort c but transform the rank position (n_{ijsc}) into a local percentile rank (R_{ijsc}) to make it comparable across schools j, following Murphy and Weinhardt (2020).

$$R_{ijsc} = \frac{n_{ijsc} - 1}{N_{jsc} - 1} \times 100 \tag{2}$$

where N_{jsc} is the cohort size of school j in cohort c of subject s. We multiply this measure by 100, resulting in a rank scale from 0 to 100, where the lowest-ranked student in each school cohort has R = 0 and the highest-ranked student has R = 100. In the case of ties, both students are assigned the lower rank.

Rank-based STEM advantage is defined for each student i in the following way:

$$\textit{Rank-based STEM advantage}_i = \frac{R_{ijsc} \textit{ of STEM GPA}}{R_{ijsc} \textit{ of non-STEM GPA}} - 1 \tag{3}$$

where R_{ijsc} are the local percentile ranks we compute in Equation 2 where $s \in \{STEM, \textit{non-STEM}\}$.

2.4. Descriptive Statistics

Our final sample is drawn from North Rhine-Westphalia's student population. NRW, being the largest federal state in Germany and comparable in size to the Netherlands, offers a rich context for examining key characteristics of German pupils, including their university preferences, fields of study, and regional distribution. For a more detailed explanation of the variables, please refer to Section 5.2 of the Appendix.

Table 1: Summary Statistics

	Females		*Males*		*Females - Males*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Mean	SD	Mean	SD	Norm.Δ	Abs.Δ	*p*-val
A. Performance in high school							
High school GPA	2.713	0.622	2.634	0.588	0.09	0.08	0.12
STEM GPA	9.204	3.239	9.470	3.259	-0.06	-0.27	0.33
Non-STEM GPA	11.015	2.218	10.594	2.262	0.13	0.42	0.03
B. Constructed variables in high school							
Grade-based STEM advantage	-0.140	0.309	-0.071	0.346	-0.15	-0.07	0.01
Rank-based STEM advantage	0.067	1.346	0.169	1.900	-0.04	-0.10	0.45
Rank STEM GPA	58.304	27.525	59.991	27.655	-0.04	-1.69	0.47
Rank non-STEM GPA	68.298	24.024	66.056	24.477	0.07	2.24	0.28
C. Background variables							
High school GPA (cohort)	2.561	0.180	2.533	0.172	0.11	0.03	0.07
STEM GPA (cohort)	7.954	1.505	7.967	1.629	-0.01	-0.01	0.92
Non-STEM GPA (cohort)	8.805	1.060	8.472	1.062	0.22	0.33	0.00
Low SES	0.292	0.455	0.474	0.500	-0.27	-0.18	0.00
Migration status	0.088	0.284	0.094	0.292	-0.01	-0.01	0.82
IQ	2.310	1.691	1.782	1.896	0.21	0.53	0.00
D. Tertiary education							
STEM degree	0.198	0.399	0.435	0.497	-0.37	-0.24	0.00
Law degree	0.041	0.199	0.039	0.194	0.01	0.00	0.88
Economics and Business degree	0.109	0.312	0.220	0.415	-0.21	-0.11	0.00
Humanities and Social Sciences degree	0.469	0.500	0.228	0.421	0.37	0.24	0.00
Health degree	0.183	0.387	0.078	0.268	0.22	0.11	0.00
University GPA	2.859	0.592	2.700	0.528	0.20	0.16	0.00
University STEM GPA	2.674	0.565	2.532	0.510	0.19	0.14	0.11
University non-STEM GPA	2.899	0.592	2.805	0.514	0.12	0.09	0.11

Note: This table reports statistics of variables by gender for a set of 573 observations. Columns 1 and 3 show the mean for each group, while Columns 2 and 4 present the standard deviation (sd). Column 5 reports normalized differences between females and males (Imbens and Wooldridge 2009). Normalized differences are calculated as averages by group status scaled by the square root of the sum of the variances. Column 6 presents the absolute differences, while Column 7 provides the *p*-values from a two-sided *t*-test for comparing means.

Panel A of Table 1 shows that female students perform better than male students as regards their overall high school GPA. Moreover, while females display a slightly lower performance in STEM subjects (0.06 *sd*), they achieve 0.13 *sd* better grades in non-STEM subjects compared to males. Gender differences in high school GPAs are statistically significant for non-STEM subjects ($p<.05$) but not for STEM subjects. The probability density distributions in Figure 1 confirm this pattern, showing significant gender disparities in distributions of non-STEM GPAs ($p<.05$) but smaller, non-significant differences in STEM fields.

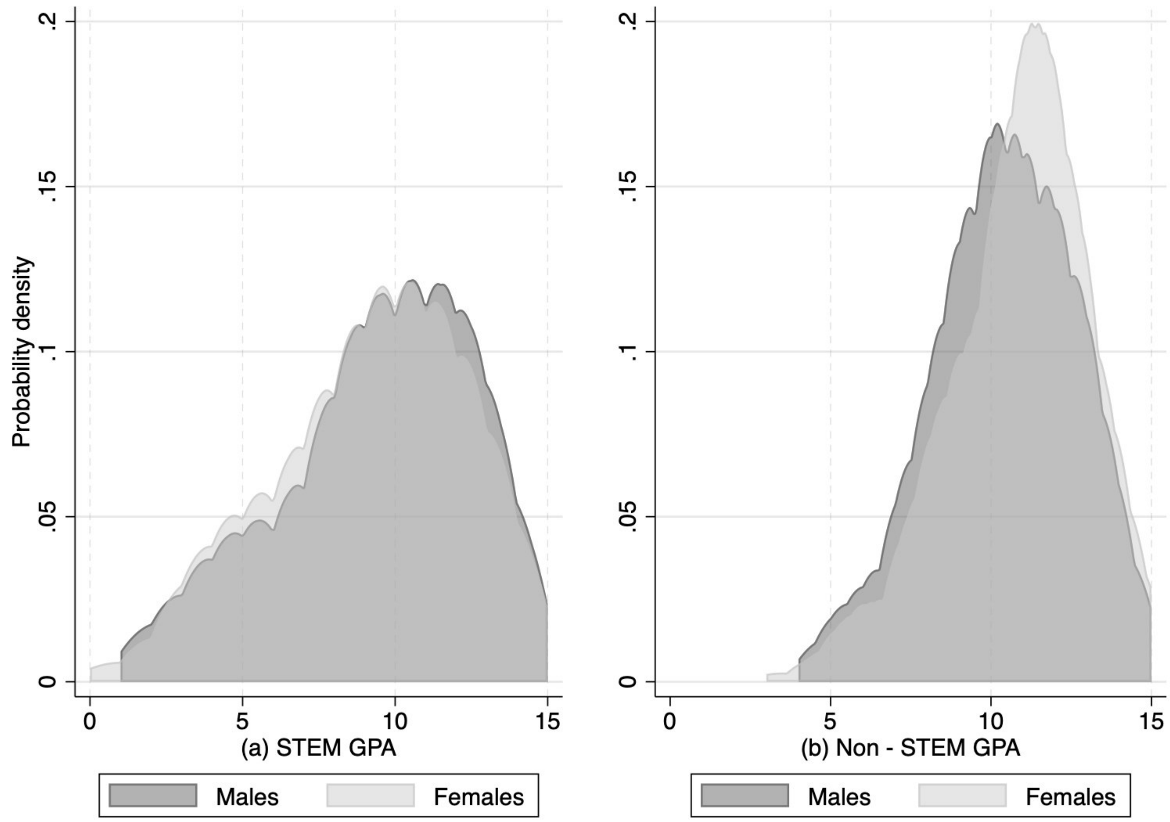

Figure 1: Kernel Densities of STEM and Non-STEM Performance

Note: The figure displays kernel density plots of the number of points achieved in STEM and non-STEM subjects in the final exams of upper secondary school for males and females, respectively. Kernel = Epanechnikov; a two-sample Kolmogorov-Smirnov test indicates a significant difference between distributions of non-STEM GPA by gender ($p<.05$), rejecting the null hypothesis of equal distributions; the test fails to reject the null hypothesis for the distributions of STEM GPA by gender, indicating no significant difference between distributions; (a) optimal bandwidth = 0.896 for males, 0.906 for females; (b) optimal bandwidth = 0.672 for males, 0.614 for females.

Panel B of Table 1 presents gender-based disparities in scholarly achievement of our constructed metrics. The outcomes reveal a significant male advantage in STEM subjects relative to non-STEM subjects, measured by grade-based STEM advantage. Males have a 0.15 *sd* higher STEM advantage ($p<.01$) based on grades. This finding is substantiated by the empirical evidence presented in Figure 2. Importantly, the higher grade-based STEM advantage for males is mainly driven by *worse performance in non-STEM GPAs of males compared to females.* Furthermore, we find no significant difference in rank-based STEM advantage. As shown in Figure 1, the variation in grade- versus rank-based STEM advantage comes from the fact that there is much more mass at the upper end of the non-STEM GPA distribution, indicating that it is easier to get a top grade in these subjects when compared to STEM subjects. We find no significant gender differences in

STEM or non-STEM GPA ranks. Since ranks do not vary across subjects, the grade-based STEM advantage seems to be driven by generally higher grades in non-STEM subjects compared to STEM subjects.

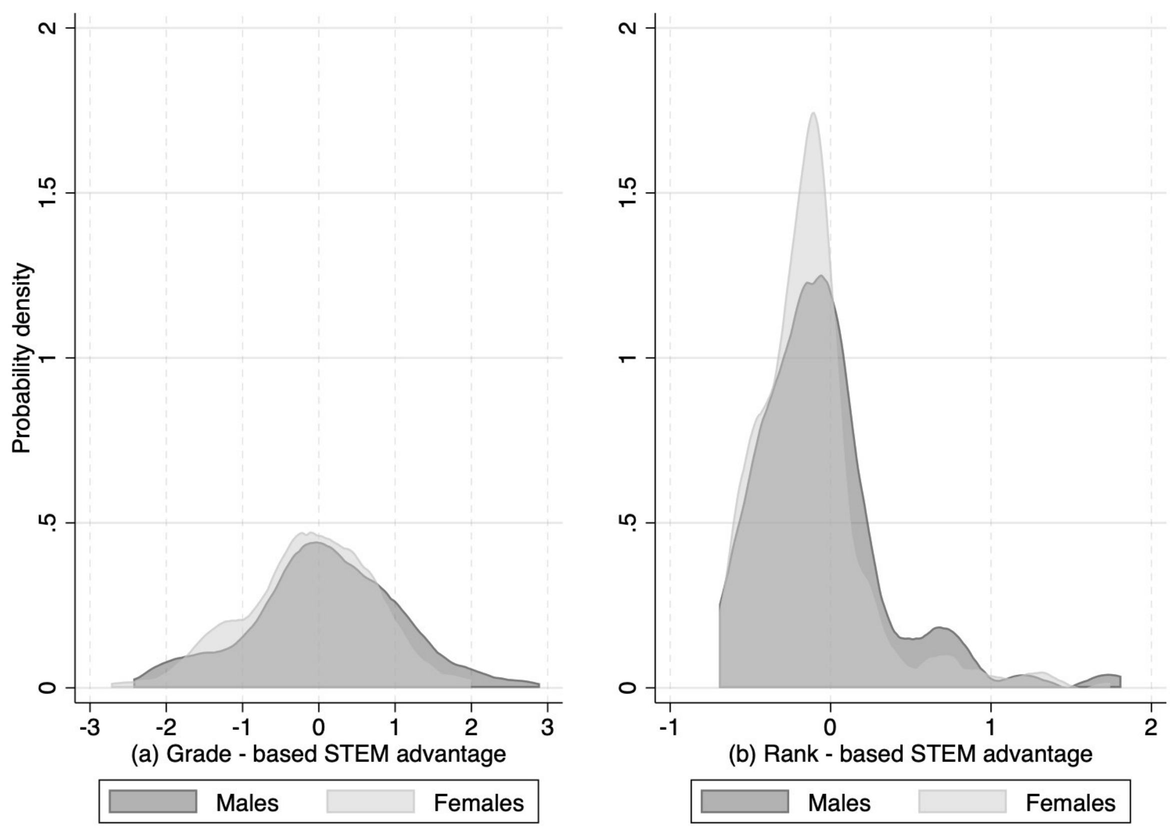

Figure 2: Kernel Densities of Grade- and Rank-Based STEM Advantage

Note: The figure displays kernel density plots of our measures of relative STEM advantage for males and females, respectively. Kernel = Epanechnikov; a two-sample Kolmogorov-Smirnov test indicates a significant difference between distributions of a grade-based STEM advantage by gender($p<.10$) and rank-based STEM advantage ($p<.05$) , rejecting the null hypothesis of equal distributions; (a) optimal bandwidth = 0.267 for males, 0.222 for females; (b) optimal bandwidth = 0.097 for males, 0.068 for females. We remove 14 outliers by dropping observations where grade-based STEM advantage exceeds 3, or rank-based STEM advantage exceeds 2. Both measures are standardized.

Panel C of Table 1 displays gender differences in background variables. While we observe considerable variation in our key variables of interest, the selected nature of our sample – individuals from upper secondary education who have enrolled in tertiary education – may limit representativeness. In particular, the selection process for university enrollment could differ between males and females. To address potential sample selection issues, we control for cohort-level performance using average GPAs in overall high school performance, STEM subjects, and non-STEM subjects.

Furthermore, we find low SES and IQ to be statistically significant factors, so we also control for these background characteristics in our empirical analysis.

Panel D of Table 1 provides summary statistics on major choice and academic performance. It shows the distribution of male and female students across STEM, Law, Economics and Business, Humanities and Social Sciences, and Health and Natural Sciences degree programs. STEM majors, characterized by their math-intensive nature, attract a significantly larger proportion of male students, resulting in a gender gap of 24%. In contrast, there is a reversed gender gap of 24% for the choice of majors in Humanities and Social Sciences, where there is a substantially higher representation of female students. Further, we see that female students outperform their male counterparts in terms of academic grades (0.20 *sd*).

Figure 3 graphically illustrates average marginal effects of STEM GPA and grade-based STEM advantage on the probability to pursue a STEM degree by gender. We observe a consistently higher likelihood of males choosing STEM majors over females based on STEM GPA and grade-based STEM advantage. Males' likelihood of choosing a STEM degree increases with their STEM advantage, while females show a more modest response.

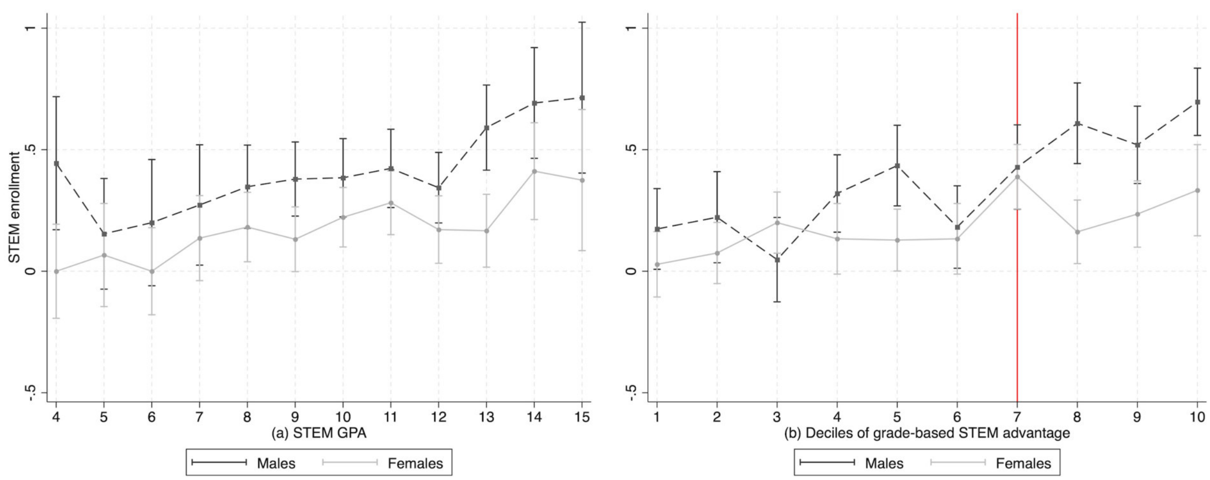

Figure 3: STEM Enrollment by STEM GPA and Grade-Based STEM Advantage

Note: For (a) STEM GPA, we predict average marginal effects on the probability of studying STEM across a range of values from 4 (minimum passing grade) to 15 points, separately for males and females. For (b) deciles of grade-based STEM advantage, we regress STEM enrollment on the deciles of grade-based STEM advantage interacted with a female indicator. We then calculate predicted probabilities of STEM enrollment across the deciles of grade-based STEM advantage by gender. The vertical line at the 7$^{\text{th}}$ decile represents the threshold, where the right-hand side indicates having a positive grade-based STEM advantage.

3. Empirical Results

3.1. The Relationship between STEM Enrollment and Relative Performance Indicators

We estimate a linear probability model to investigate the relationship between our relative STEM performance indicators and the choice of a STEM major.

$$
\begin{aligned}
Y_i = \alpha + \beta_0 \times Female_i + \\
\beta_1 \times Performance\ indicator_{pi} + \\
\beta_2\, Female_i \times Performance\ indicator_{pi} + \\
\gamma' X_i + \delta_t + \varepsilon_i
\end{aligned} \tag{4}
$$

Y_i is a binary variable indicating whether individual i enrolled into a STEM major. $Female_i$ is a dummy variable equal to one when i is female, zero if male. $Performance\ indicator_{pi}$ is a placeholder for either grade- or rank-based STEM advantage. We introduce an interaction term of performance indicators and the female dummy. Both performance measures capture the proficiency of individual i in STEM subjects compared to non-STEM subjects, with grade-based STEM advantage focusing on performance differences based on grades and rank-based STEM advantage being based on information from local percentile ranks. Our control variables, denoted as X_i, include both GPAs and ranks in STEM and non-STEM subjects to account for students' absolute performance levels. We incorporate school-cohort performance by including average STEM and non-STEM GPAs, along with high school GPAs. To control for ability, we include measures of IQ[6] and the individual's high school GPA. Additionally, we include personal background information on socioeconomic status and migration status. To accommodate potential graduation-year characteristics, we introduce graduation-year dummies denoted as δ_t. The error term is represented by ε_i. We estimate the model using robust standard errors. Both performance indicators are standardized in order to compare effect sizes across variables.

The coefficient of primary interest is β_2, denoting potential heterogeneity in the importance of our performance indicators for the likelihood of pursuing a STEM degree across genders. The

[6] We measured IQ based on ten items from a Raven-type matrices IQ test (Raven and Court 1998).

results from estimating this model are displayed in Table 2. The full table is displayed in Table A.2 of the Appendix, Section 5.3.

First, we inspect the results without interaction terms in Columns 1 and 2 for grade-based STEM advantage and Columns 4 and 5 for rank-based STEM advantage. Following this, we explore gender differences in the effects of grade-based STEM advantage in Column 3 and rank-based STEM advantage in Column 6. The results indicate a substantial and statistically significant gender gap of approximately 17 percentage points in STEM enrollment across all specifications ($p<.01$). A one-standard-deviation increase in grade-based STEM advantage boosts the probability of pursuing a STEM degree by 15-19 percentage points ($p<.01$), ceteris paribus.

The results from the interaction model displayed in Column 3, indicate that being female reduces the positive effect of a one-standard-deviation increase in grade-based STEM advantage by 9.7 percentage points relative to males ($p<.01$). Thus, while grade-based STEM advantage increases the likelihood of choosing a STEM subject for both genders, the effect is around 50% smaller for females. Hence, although females show similar performance in STEM subjects, the influence of grade-based STEM advantage on choosing a STEM major is disproportionately smaller for them compared to males. That is, females require significantly stronger signals of relative ability based on grades in STEM compared to non-STEM subjects to pursue a STEM degree. In fact, to attain an equivalent probability of pursuing a STEM degree as males, females require a grade-based STEM advantage that is almost four standard deviations higher.[7] In Column 6, the interaction term (Female $\times$ Rank-based STEM advantage) largely offsets or even reverses the main effect. This is a discouraging finding: Females' relative STEM performance ranking seems to have no influence on their likelihood of choosing a STEM major.

Our results point towards substantial selection costs of choosing a STEM occupation among females. Further, one interpretation of our findings is that non-performance-related factors, such as field preferences, perceived future working conditions, and perceived discrimination strongly discourage females from choosing a STEM occupation, even if they obtain very positive signals about their STEM abilities. As a consequence, even strong signals about STEM performance hardly affect female choices compared to males.

[7] We want to set the probability of males to pursue a STEM degree equal to the probability for females conditional on grade-based STEM advantage. Given Table 2, we solve for x in $0.188=-0.160+x(0.188-0.097)$, which leads to an $x=3.8$.

Table 2: STEM Enrollment and Relative Performance Indicators

	Grade-based STEM advantage			*Rank-based STEM advantage*		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.184***	-0.165***	-0.160***	-0.200***	-0.169***	-0.168***
	(0.036)	(0.039)	(0.039)	(0.037)	(0.039)	(0.039)
Grade-based STEM advantage	0.174***	0.152***	0.188***			
	(0.055)	(0.054)	(0.052)			
Female×Grade-based STEM advantage			-0.097***			
			(0.034)			
Rank-based STEM advantage				0.023	0.020	0.042***
				(0.016)	(0.015)	(0.011)
Female×Rank-based STEM advantage						-0.058**
						(0.028)
Grades	Yes	Yes	Yes	No	No	No
Ranks	No	No	No	Yes	Yes	Yes
Other controls	No	Yes	Yes	No	Yes	Yes
Observations	573	573	573	573	573	573
Adjusted R^2	0.134	0.168	0.178	0.109	0.151	0.153

Note: Columns 1-2 and Columns 4-5 present estimated effects of grade- or rank-based STEM advantage on STEM enrollment in tertiary education, respectively. Columns 3 interacts grade-based STEM advantage with gender to identify heterogeneous effects. Column 6 repeats this analysis for rank-based STEM advantage. Regressions are estimated with a constant, control for STEM and non-STEM GPAs, and ranks of STEM and non-STEM GPA. Other controls include school-cohort performance as measured by STEM, non-STEM, and high school GPA, individual ability approximated by IQ and high school GPA, personal background such as socioeconomic status, migration status, and graduation-year dummies, which are omitted in this table for brevity. We use robust standard errors, reported in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.2. Quantifying Decision-Making Differences: Male vs. Female Choice Worlds

In the previous sections, we identified gender disparities in STEM performance and observed varying effects of STEM advantage across genders, indicating that strong signals of having a grade-based STEM advantage hardly induce females to choose a STEM major. In this section, we conduct a twofold Kitagawa-Oaxaca-Blinder-type decomposition (Jann 2008) to assess the contributions of performance differences to the gender gap. Our goal is to run the following thought experiment: How would observed performance differences among males and females affect the sorting into STEM-related fields, in a world where female (male) major preferences in STEM-related fields resembled that of males (females)? To assess the relative importance of performance measures in a non-discriminatory "male-choice" or a discriminatory "female-choice" world, we respectively categorize potential drivers of the STEM gender gap into two groups. The first group represents the grade-based performance indicators, incorporating grade-based STEM advantage, STEM GPA, and non-STEM GPA. The second group represents the rank-based performance indicators, comprising rank-based STEM advantage, rank of STEM GPA, and rank of non-STEM GPA.

Consider two groups, *male* and *female*, with our outcome variable Y representing STEM major, and our set of predictors X driving the STEM gender gap categorized above in grade- and rank-based performance indicators. We define the mean outcome difference as follows,

$$R = E(Y_{male}) - E(Y_{female}). \tag{5}$$

Given $E(Y)$ as the expected value of the outcome variable, we seek to understand the extent to which group differences in predictors contribute to the mean outcome difference. The twofold decomposition dissects outcome differences into an explained part ("quantity effect") and an unexplained component (which we term "preferences effect").

Consider the following linear model:

$$Y_l = X_l' \beta_l + \epsilon_l. \tag{6}$$

Assuming $E(\epsilon_l) = 0$ for each group $l \in$ (male, female), within the linear model framework, where X represents a vector containing predictors and a constant, β encompasses slope parameters and intercept, and ϵ denotes the error term, the mean outcome difference in the twofold decomposition can be expressed as:

$$\begin{aligned} R =& \{E(X_{male}) - E(X_{female})\}'\beta^* \\ &+ \{E(X_{male})'(\beta_{male} - \beta^*) + E(X_{female})'(\beta^* - \beta_{female})\}. \end{aligned} \tag{7}$$

The first component,

$$Q = \{E(X_{male}) - E(X_{female})\}'\beta^*, \tag{8}$$

is the part of the outcome differential that is explained by group differences in the predictors, the *quantity effect*. The second component,

$$U = \{E(X_{male})'(\beta_{male} - \beta^*) + E(X_{female})'(\beta^* - \beta_{female})\}, \tag{9}$$

is the *preferences effect*, i.e., the part that reflects differences in decision-making that can be due to (anticipated) discrimination, considerations about fit, unobserved factors or preference heterogeneity between males and females.

Table 3 presents two set of results. In Columns 1 and 2, following the literature on STEM disparities, we posit that the STEM enrollment gap is biased against women rather than men.[8] Thus, in Equation 7, we use the coefficients of males, denoted as β_{male}, for β^*, evaluating the relative significance of performance measures in a non-discriminatory "male-choice" world. In Columns 2 and 3, we set $\beta_{female} = \beta^*$ to analyse the influence of performance measures in a "female-choice" scenario.

[8]Determining the components of the twofold decomposition, as shown in Equation 7, requires an estimate for the unknown non-discriminatory coefficients vector β^*. Oaxaca (1973) suggests that β^* can equal either β_{male} or β_{female} based on the direction of discrimination towards a particular group.

Table 3: Kitagawa-Oaxaca-Blinder Decomposition of the STEM Gender Gap

	Male coefficients		*Female coefficients*	
	(1) Absolute	(2) Share	(3) Absolute	(4) Share
Difference	0.211***	100.000	0.211***	100.000
	(0.039)		(0.039)	
Explained difference	0.045*	21.327	0.059***	27.962
	(0.025)		(0.017)	
Composition effects attributable to				
(A) Grade-based performance indicator	0.054**	25.592	0.013	6.161
	(0.023)		(0.010)	
(B) Rank-based performance indicator	-0.010	4.739	0.001	0.474
	(0.009)		(0.007)	
Control variables	-0.001	0.474	0.045***	21.327
	(0.015)		(0.015)	
Observations	573		573	

Notes: This table decomposes differences in STEM subject choice in tertiary education attributable to differences in grade- and rank-based performance indicators using a twofold Kitagawa-Oaxaca-Blinder decomposition. We control for IQ, socioeconomic status, and migration status. Columns 1 and 2 use male coefficients for the unknown non-discriminatory coefficients vector β^*, while Columns 3 and 4 use female coefficients. For each decomposition, we also present the share of the difference that is attributable to the respective component. Robust standard errors in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Conceptually, the β-coefficients can be interpreted as preference parameters, reflecting the decision-making tendencies of males and females when confronted with distinct abilities and constraints. By using male coefficients in Columns 1 and 2, we gain insights into the role of gender differences in decision-making processes. Assuming that females may decide differently due to preference-related factors such as discouragement, lack of role models, or concerns about penalties related to family responsibilities in STEM fields, we can assess the extent to which relative STEM performance disparities would persist if these barriers were eliminated.

Abstracting from these barriers, 21% of the STEM gender gap stems from group differences in predictors (quantity effect). Our decomposition reveals that gender differences in grade-based performance indicators account for 26% (p<.05) of the STEM enrollment gap, while rank-based metrics show no significant impact. When examining a counterfactual scenario where preferences are "female" (Columns 3-4), neither grade- nor rank-based performance disparities significantly influence STEM-field selection. Instead, the gender gap in STEM enrollment is largely driven by differences in our control variables: IQ, migration status, and socioeconomic status. This observation aligns with our previous finding presented in Table 2 , where rank-based STEM advantage does not demonstrate significant economic relevance in relation to the STEM enrollment gap. Since the coefficients of rank-based STEM advantage and its interaction with gender offset each other, we do not find an overall effect of rank-based STEM advantage on STEM choice in a "female-preference" world.

To address potential concerns about the choice of reference coefficients in decomposition analyses (Neumark 1988, Oaxaca and Ransom 1994), we additionally estimate a pooled model where coefficients are derived from a regression. Results presented in Table A.3 of the Appendix largely align with our main specifications: the explained portion of the gender gap remains substantial at 21%, with grade-based performance indicators continuing to be the primary driver, accounting for 14% of the gap ($p<.05$). The insignificant role of rank-based measures persists across all specifications. This is unsurprising given that there are no male-female differences in rank-based STEM advantage.

Given that performance differentials minimally influence female STEM enrollment, we examine alternative drivers. Prior literature suggests anticipated gender discrimination may deter STEM pursuit (e.g., Porter and Serra 2020). Using a linear probability model (Table 4), we examine whether female students in STEM programs report higher levels of anticipated discrimination. Our preferred specification (Column 4) shows females experience a 32 percentage point higher probability of anticipated gender discrimination ($p<.01$), with an additional 19.8 percentage point increase among STEM-enrolled females ($p<.05$). This implies that females pursuing STEM degrees expect additional obstacles. To the extent that only those women select into STEM fields who expect less discrimination in a STEM-related occupation, our estimates provide a lower bound estimate of perceived barriers or discrimination in the STEM occupations. For detailed table contents, we refer to Table A.4 of the Appendix in Section 5.3.

Table 4: Anticipated Discrimination and STEM Enrollment

	Anticipated gender-based discrimination			
	(1)	(2)	(3)	(4)
Female	0.382***	0.375***	0.336***	0.319***
	(0.036)	(0.038)	(0.041)	(0.043)
STEM major	0.043	0.037	-0.036	-0.057
	(0.040)	(0.044)	(0.041)	(0.043)
Female×STEM major			0.169**	0.198**
			(0.082)	(0.087)
Relative STEM advantages	No	Yes	No	Yes
Grades	No	Yes	No	Yes
Ranks	No	Yes	No	Yes
Other controls	No	Yes	No	Yes
Observations	573	573	573	573
Adjusted R^2	0.150	0.138	0.154	0.144

Notes: Columns 1-2 present estimated effects of gender and STEM enrollment on the expectation of gender-based discrimination. Columns 3-4 interact STEM enrollment with gender to identify heterogeneous effects. Regressions are estimated with a constant. 'Relative STEM advantages' includes our grade- and rank-based measures of STEM advantage. 'Grades' controls for STEM and non-STEM GPA and 'Ranks' controls for rank of STEM and non-STEM GPA. Other controls include school-cohort performance as measured by STEM, non-STEM, and high school GPA, individual ability approximated by IQ and high school GPA, personal background such as socioeconomic status, migration status, and graduation-year dummies, which are omitted in this table for brevity. We use robust standard errors, reported in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4. Conclusion

Gender disparities in relative performance across STEM and non-STEM fields have long-lasting effects, potentially affecting not only educational decisions, but also leading to wage disparities (e.g., Blau and Kahn 2017, Kleven et al. 2019), especially in sectors such as science and technology (Goldin 2014) and countries like Germany where occupational mobility is low (SOEP 2022). The scarcity of STEM graduates, particularly among women, poses a significant challenge to the tech industry and innovation in general (Carnevale et al. 2011, Bianchi and Giorcelli 2020, Del Carpio and Guadalupe 2022, Coff 1997).

Our study examines how performance indicators affect human-capital investment and selection into STEM-related fields in Germany. Germany's publicly funded, tuition-free education system, coupled with standardized curricula and compensation schemes for teachers, minimizes financial constraints on educational choices. Persistent gender-based disparities in tertiary-education choices thus seemingly reflect significant non-monetary factors in educational decision-making.

Our study identifies important gender dynamics in educational decision-making. Women graduate from high school with better GPAs. Although there are no significant or large gender disparities in STEM subject performance, males perform significantly worse than females in non-STEM subjects. Additionally, we observed a 24% gender gap in male-female STEM major choices, aligning with trends observed across OECD countries, including Germany (OECD 2024). Our analysis reveals that an grade-based STEM advantage is positively associated with STEM enrollment among both genders, though the impact is much smaller for females compared to males. This findings indicate that, despite similar proficiency in STEM subjects, females require substantially stronger grade-based performance signals in STEM relative to non-STEM subjects to pursue a STEM degree. However, women who do enter STEM programs ultimately outperform their male peers in terms of GPA.

Our decomposition analysis examines how differences in relative performance indicators and control variables contribute to the STEM gender gap under different preference scenarios. We show that in a male-choice world, where female preferences mirror those of males, 21% of the STEM gender gap can be attributed to group differences in predictors. Specifically, differences in grade-based STEM advantage and performance gaps across STEM and non-STEM subjects account for 26% of the gender gap, while rank-based differences explain only 5%. Conversely, in a female-choice world grade- and rank-based performance differences account for merely 6% and 0.5%, respectively. These results underscore that performance variations play a minimal role in females' STEM choices, with rank-based STEM advantage showing consistently low economic significance across specifications.

We interpret these findings of being indicative of substantial perceived selection costs to entering a STEM occupation among females and we show that women who selected into STEM majors indeed expect more gender-related on-the-job discrimination.

Our analysis is informative as regards policies that address the underrepresentation of women in STEM and the ongoing discourse on gender inequality in education and labor markets (Goldin 2014, Marianne 2011, Blau and Kahn 2017, Kleven and Landais 2017, Goulas et al. 2022, Breda and Napp 2019, Breda et al. 2019, 2018, 2023, Francesconi and Parey 2018, Zafar 2013, Wiswall and Zafar 2018). Our results suggest that a change in STEM grades or respective grading policies in secondary school will have little impact on reducing the gender gap in STEM. On the positive side, this also implies that "easier grades" as observed mostly in non-STEM subjects might not systematically drive females out of STEM fields, although this could differ when comes to grades in higher education (Ahn et al. 2024).

Our results further indicate that, despite similar proficiency in STEM subjects, females seem to be held back by differential preferences or barriers when it comes to pursuing a STEM degree. Non-performance-related factors such as field preference, perceived future working conditions, and perceived discrimination seemingly discourage females from choosing a STEM occupation even if they obtain very positive signals about their STEM abilities. Specifically, females require a STEM advantage that is four standard deviations higher than males to have the same probability of studying a STEM subject, highlighting the need for more encouragement and support for women in science. In line with Breda et al. (2023), our findings suggest significant overselection costs, emphasizing the importance of addressing gender-specific barriers in STEM fields to ensure equitable opportunities for all aspiring students.

A systematic analysis of measures and programs aiming to effectively counter perceived on-the-job discrimination in STEM occupations would be an interesting endeavor for future research. In light of our findings, such policies could reduce the STEM-enrollment gap, improve the talent pool in STEM occupations, and may ultimately improve a country's growth and competitiveness.

References

Ahn, T., P. Arcidiacono, A. Hopson, and J. Thomas (2024). Equilibrium grading policies with implications for female interest in STEM courses. *Econometrica 92*(3), 849–880.

Anger, C. and A. Plünnecke (2022). MINT gewinnt. Hohe Löhne in den MINT-Berufen. Technical Report 106, IW-Kurzbericht, Cologne.

Attanasio, O. P. and K. M. Kaufmann (2014). Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics 109*, 203–216.

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.

Bianchi, N. and M. Giorcelli (2020). Scientific education and innovation: From technical diplomas to university stem degrees. *Journal of the European Economic Association 18*(5), 2608–2646.

Blau, F. D. and L. M. Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature 55*(3), 789–865.

Bobba, M. and V. Frisancho (2016). Learning about oneself: The effects of performance feedback on school choice. Unpublished manuscript, Inter-American Development Bank.

Bond, T. N., G. Bulman, X. Li, and J. Smith (2018). Updating human capital decisions: Evidence from SAT score shocks and college applications. *Journal of Labor Economics 36*(3), 807–839.

Breda, T., J. Grenet, M. Monnet, and C. Van Effenterre (2019). Female role models: Are they effective at encouraging girls to study science? Notes IPP, (45).

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.

Breda, T., E. Jouini, and C. Napp (2018). Societal inequalities amplify gender gaps in math. *Science 359*(6381), 1219–1220.

Breda, T., E. Jouini, and C. Napp (2023). Gender differences in the intention to study math increase with math performance. *Nature Communications 14*(1), 3664.

Breda, T. and C. Napp (2019). Girls' comparative advantage in reading can largely explain the gender gap in math-related fields. *Proceedings of the National Academy of Sciences 116*(31), 15435–15440.

Canaan, S. and P. Mouganie (2023). The impact of advisor gender on female students' STEM enrollment and persistence. *Journal of Human Resources 58*(2), 593–632.

Card, D. and A. A. Payne (2021). High school choices and the gender gap in STEM. *Economic Inquiry 59*(1), 9–28.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics 134*(3), 1163–1224.

Carnevale, A. P., N. Smith, and M. Melton (2011). Stem: Science technology engineering mathematics. Technical report, Georgetown University Center on Education and the Workforce. Accessed on July 25, 2024: `https://files.eric.ed.gov/fulltext/ED525297.pdf`.

Ceci, S. J., D. K. Ginther, S. Kahn, and W. M. Williams (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest 15*(3), 75–141.

Coff, R. W. (1997). Human assets and management dilemmas: Coping with hazards on the road to resource-based theory. *Academy of Management Review 22*(2), 374–402.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review 95*(2), 158–165.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human resources 42*(3), 528–554.

Del Carpio, L. and M. Guadalupe (2022). More women in tech? Evidence from a field experiment addressing social identity. *Management Science 68*(5), 3196–3218.

Delaney, J. M. and P. J. Devereux (2019). Understanding gender differences in STEM: Evidence from college applications. *Economics of Education Review 72*, 219–238.

Duflo, E., P. Dupas, and M. Kremer (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review 101*(5), 1739–74.

Elsner, B. and I. E. Isphording (2017). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics 35*(3), 787–828.

Elsner, B., I. E. Isphording, and U. Zölitz (2021). Achievement rank affects performance and major choices in college. *The Economic Journal 131*(640), 3182–3206.

Francesconi, M. and M. Parey (2018). Early gender gaps among university graduates. *European Economic Review 109*, 63–82.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Goulas, S., S. Griselda, and R. Megalokonomou (2022). Comparative advantage and gender gap in STEM. *Journal of Human Resources*, 0320–10781R2.

Guiso, L., F. Monte, P. Sapienza, and L. Zingales (2008). Culture, gender, and math. *Science 320*(5880), 1164–1165.

Hemkes, B., K. M. Wiesner, L. Borghans, P. Seegers, J. Bergerhoff, and S. Hartmann (2016). Studierendenbefragung zur attraktivität der beruflichen bildung. *Abschlussbericht des Entwicklungsprojekts 3*, 306.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal 8*(4), 453–479.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics 125*(2), 515–548.

Kleven, H. and C. Landais (2017). Gender inequality and economic development: Fertility, education and norms. *Economica 84*(334), 180–209.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Kugler, A. D., C. H. Tinsley, and O. Ukhaneva (2021). Choice of majors: Are women really different from men? *Economics of Education Review 81*, 102079.

Li, H. and X. Xia (2024). Grades as signals of comparative advantage: How letter grades affect major choices. *Journal of Economic Behavior & Organization 227*, 106717.

Marianne, B. (2011). New perspectives on gender. In *Handbook of Labor Economics*, Volume 4, pp. 1543–1590. Elsevier.

Murphy, R. and F. Weinhardt (2020). Top of the class: The importance of ordinal rank. *The Review of Economic Studies 87*(6), 2777–2826.

Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources 23*, 279–295.

Nicoletti, C., A. Sevilla, and V. Tonei (2022). Gender stereotypes in the family. IZA Discussion Paper.

Niederle, M. and L. Vesterlund (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives 24*(2), 129–144.

Nollenberger, N., N. Rodríguez-Planas, and A. Sevilla (2016). The math gender gap: The role of culture. *American Economic Review 106*(5), 257–261.

Nosek, B. A., F. L. Smyth, N. Sriram, N. M. Lindner, T. Devos, A. Ayala, Y. Bar-Anan, R. Bergh, H. Cai, K. Gonsalkorale, et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences 106*(26), 10593–10597.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 693–709.

Oaxaca, R. L. and M. R. Ransom (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics 61*(1), 5–21.

OECD (2017). *OECD Handbook for Internationally Comparative Education Statistics*.

OECD (2019a). *PISA 2018 Results (Volume I)*.

OECD (2019b). *PISA 2018 Results (Volume II)*.

OECD (2024). The OECD Going Digital Toolkit. Technical report. Accessed online on July 15, 2024: `https://goingdigital.oecd.org/en/indicator/43`.

Owen, A. L. (2010). Grades, gender, and encouragement: A regression discontinuity analysis. *The Journal of Economic Education 41*(3), 217–234.

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.

Rask, K. and J. Tiefenthaler (2008). The role of grade sensitivity in explaining the gender imbalance in undergraduate economics. *Economics of Education Review 27*(6), 676–687.

Raven, J. C. and J. H. Court (1998). *Raven's progressive matrices and vocabulary scales*. Oxford Psychologists Press Oxford.

Riegle-Crumb, C., B. King, E. Grodsky, and C. Muller (2012). The more things change, the more they stay the same? prior achievement fails to explain gender inequality in entry into STEM college majors over time. *American Educational Research Journal 49*(6), 1048–1073.

SOEP (2022). Own calculation. Unpublished raw data from socio-economic panel (soep) v37, German Institute for Economic Research. Accessed online on Feb 20, 2024: `https://paneldata.org/soep-core/datasets/biojob/occmove`.

Stinebrickner, R. and T. R. Stinebrickner (2014). A major in science? Initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies 81*(1), 426–472.

Stinebrickner, T. and R. Stinebrickner (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics 30*(4), 707–748.

Tan, B. J. (2023). The consequences of letter grades for labor market outcomes and student behavior. *Journal of Labor Economics 41*(3), 565–588.

Terrier, C. (2020). Boys lag behind: How teachers' gender biases affect student achievement. *Economics of Education Review 77*, 101981.

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.

Wiswall, M. and B. Zafar (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies 82*(2), 791–824.

Wiswall, M. and B. Zafar (2018). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics 133*(1), 457–507.

Wiswall, M. and B. Zafar (2021). New approaches to understanding choice of major. NBER Reporter (2): 18–21.

Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics 29*(2), 301–348.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources 48*(3), 545–595.

5. Appendix

5.1. Institutional Background

Upon completing primary school, students are channeled into different secondary-school types. The four main types of secondary schools are: *Gymnasium* (academic secondary school/high school, ISCED Level 3), *Realschule* (intermediate secondary school, ISCED Level 2), *Hauptschule*, and *Gesamtschule* (comprehensive school, ISCED Level 2).[9] In our analysis we focus on the *Gymnasium*, which can either last eight (G8) or nine (G9) years depending on the state and cohort. It culminates in the *Abitur*, the highest secondary-school certificate and a prerequisite for admission to tertiary education. While transitions between these four types of tracks are theoretically possible at any time, the frequency and structure of such transitions vary by state, with most upward movements occurring after the completion of lower secondary programs. This tracking system plays a crucial role in shaping students' future academic and professional paths within the German education landscape. For those interested in further details, we direct their attention to the regulatory framework governing both the upper secondary level of *Gymnasium* and the *Abitur* examination in North Rhine-Westphalia. This framework is established by the "Verordnung über den Bildungsgang und die Abiturprüfung in der gymnasialen Oberstufe" (APO-GOSt).[10] Enacted on 5 October 1998, this foundational regulation provides the legal and structural basis for the educational processes and assessment methods analysed in our study. The APO-GOSt serves as a cornerstone document, outlining the curriculum structure, examination procedures, and qualification requirements for students in the upper secondary level of *Gymnasiums* in North Rhine-Westphalia.

---

[9] ISCED-97 definitions provided by the OECD (2017).

[10] This translates to "Ordinance on the Educational Path and *Abitur* Examination in Upper Secondary Education" and can be accessed here: `https://recht.nrw.de/lmi/owa/br_text_anzeigen?v_id=10000000000000000186`.

5.2. Variable Descriptions

Table A.1: Variable Definitions

Variable	Description
High school GPA	Between 1.0 and 4.0 (higher better)
STEM GPA	0-15 points (higher better)
Non-STEM GPA	0-15 points (higher better)
Rank of STEM GPA	Between 0 and 100 (higher better)
Rank of non-STEM GPA	Between 0 and 100 (higher better)
High school GPA (cohort)	Between 1.0 and 4.0 (higher better)
STEM GPA (cohort)	0-15 points (higher better)
Non-STEM GPA (cohort)	0-15 points (higher better)
Low SES	Dummy that equals 1 if at least one parent has a high school diploma, 0 else
Migration status	Dummy that equals 1 if individual grew up in another country, 0 else
University GPA	Between 1.0 and 4.0 (higher better)
University STEM GPA	Between 1.0 and 4.0 (higher better)
University non-STEM GPA	Between 1.0 and 4.0 (higher better)

5.3. Tables

Table A.2: STEM Enrollment and Relative Performance Indicators

	Grade-based STEM advantage			*Rank-based STEM advantage*		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.184***	-0.165***	-0.160***	-0.200***	-0.169***	-0.168***
	(0.036)	(0.039)	(0.039)	(0.037)	(0.039)	(0.039)
Grade-based STEM advantage	0.174***	0.152***	0.188***			
	(0.055)	(0.054)	(0.052)			
Female×Grade-based STEM advantage			-0.097***			
			(0.034)			
Rank-based STEM advantage				0.023	0.020	0.042***
				(0.016)	(0.015)	(0.011)
Female×Rank-based STEM advantage						-0.058**
						(0.028)
STEM GPA	-0.015	-0.003	0.003			
	(0.018)	(0.019)	(0.020)			
Non-STEM GPA	0.015	0.023	0.020			
	(0.016)	(0.017)	(0.017)			
Rank STEM GPA				0.004***	0.004***	0.004***
				(0.001)	(0.001)	(0.001)
Rank non-STEM GPA				-0.001	-0.001	-0.001
				(0.001)	(0.001)	(0.001)
STEM GPA (cohort)		0.005	0.006		0.044***	0.043***
		(0.015)	(0.015)		(0.016)	(0.016)

Continues next page

Table A.2: STEM Enrollment and Relative Performance Indicators (Continued)

	Grade-based STEM advantage			*Rank-based STEM advantage*		
	(1)	(2)	(3)	(4)	(5)	(6)
Non-STEM GPA (cohort)		-0.006	-0.007		-0.021	-0.020
		(0.020)	(0.020)		(0.020)	(0.020)
High school GPA (cohort)		-0.312**	-0.291*		-0.330**	-0.332**
		(0.155)	(0.153)		(0.157)	(0.155)
IQ		0.010	0.009		0.009	0.008
		(0.010)	(0.011)		(0.011)	(0.011)
High school GPA		-0.049	-0.053		-0.056	-0.055
		(0.048)	(0.048)		(0.044)	(0.044)
Low SES		0.094**	0.104***		0.093**	0.096**
		(0.040)	(0.040)		(0.041)	(0.041)
Migration status		0.099	0.120*		0.096	0.105
		(0.067)	(0.068)		(0.068)	(0.068)
Grad.-year	No	Yes	Yes	No	Yes	Yes
Observations	573	573	573	573	573	573
Adjusted R^2	0.134	0.168	0.178	0.109	0.151	0.153

Notes: Columns 1-2 and Columns 4-5 present estimated effects of grade- or rank-based STEM advantage on STEM enrollment in tertiary education, respectively. Columns 3 interacts grade-based STEM advantage with gender to identify heterogeneous effects. Column 6 repeats this analysis for rank-based STEM advantage. Regressions are estimated with a constant, control for STEM and non-STEM GPAs, and ranks of STEM and non-STEM GPA. Other controls include school-cohort performance as measured by STEM, non-STEM, and high school GPA, individual ability approximated by IQ and high school GPA, personal background such as socioeconomic status, migration status, and graduation-year dummies. Graduation-year dummies are omitted for brevity. We use robust standard errors, reported in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Kitagawa-Oaxaca-Blinder Decomposition of the STEM Gender Gap

	Pooled regression model	
	(1) Absolute	(2) Share
Difference	0.211***	100.000
	(0.038)	
Explained difference	0.045***	21.327
	(0.016)	
Composition effects attributable to		
(A) Grade-based performance indicator	0.030**	14.218
	(0.013)	
(B) Rank-based performance indicator	-0.002	0.948
	(0.007)	
Control variables	0.017*	8.057
	(0.010)	
Observations	573	

Notes: This table decomposes differences in STEM subject choice in tertiary education attributable to differences in absolute and relative performance indicators using twofold Kitagawa-Oaxaca-Blinder decomposition from a pooled regression model. We control for IQ, socioeconomic status, and migration status. For each decomposition, we also present the share of the difference that is attributable to the respective component. Robust standard errors in parenthesis.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Anticipated Discrimination and STEM Enrollment

	Anticipated gender-based discrimination			
	(1)	(2)	(3)	(4)
Female	0.382***	0.375***	0.336***	0.319***
	(0.036)	(0.038)	(0.041)	(0.043)
STEM major	0.043	0.037	-0.036	-0.057
	(0.040)	(0.044)	(0.041)	(0.043)
Female×STEM enrollment			0.169**	0.198**
			(0.082)	(0.087)
Grade-based STEM advantage		0.004		0.008
		(0.066)		(0.065)
Rank-based STEM advantage		0.029		0.031
		(0.025)		(0.026)
STEM GPA		0.031		0.033
		(0.032)		(0.031)
Non-STEM GPA		0.000		0.001
		(0.028)		(0.027)
Rank STEM GPA		-0.003		-0.004
		(0.003)		(0.003)
Rank non-STEM GPA		0.000		0.000
		(0.002)		(0.002)
STEM GPA (cohort)		-0.013		-0.017
		(0.027)		(0.027)
Non-STEM GPA (cohort)		0.009		0.004
		(0.025)		(0.025)
High school GPA (cohort)		-0.114		-0.097
		(0.175)		(0.174)
IQ		0.004		0.008
		(0.011)		(0.011)
High school GPA		0.034		0.035
		(0.055)		(0.055)
Low SES		0.007		-0.001
		(0.042)		(0.042)
Migration status		-0.068		-0.074
		(0.057)		(0.055)
Grad.-year	No	Yes	No	Yes
Observations	573	573	573	573
Adjusted R^2	0.150	0.138	0.154	0.144

Note: Columns 1-2 present estimated effects of gender and STEM enrollment on the expectation of gender-based discrimination. Columns 3-4 interact STEM enrollment with gender to identify heterogeneous effects. Regressions are estimated with a constant, control for grade- and rank-based STEM advantage, STEM and non-STEM GPAs, and ranks of STEM and non-STEM GPA. Other controls include cohort performance as measured by STEM, non-STEM, and high school GPA, individual ability approximated by IQ and high school GPA, and personal background such as socioeconomic status and migration status. Graduation-year dummies are omitted in this table for brevity. We use robust standard errors, reported in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$