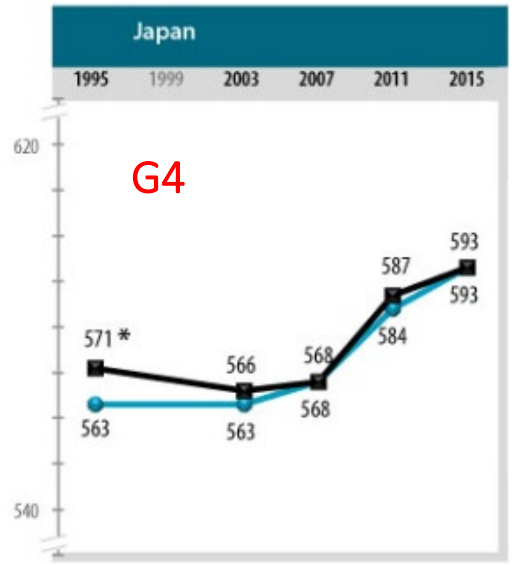
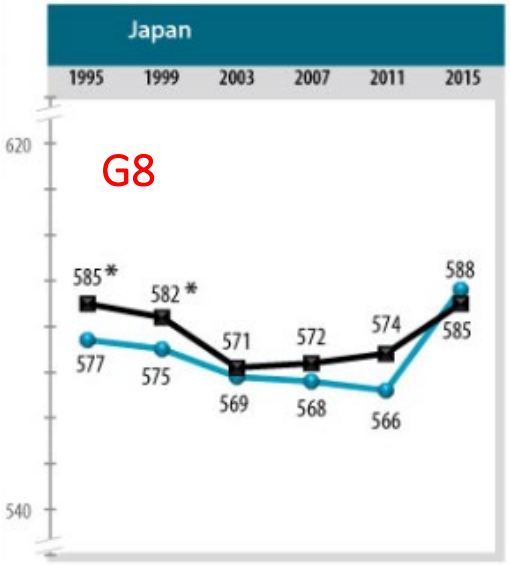
The Gender Gap in STEM:
The Evidence from Japanese High School Mock-Exam

Hideaki Ishikura (Keio University)
Ryohei Hayashi (Kochi University of Technology)
Makiko Nakamuro (Keio University)

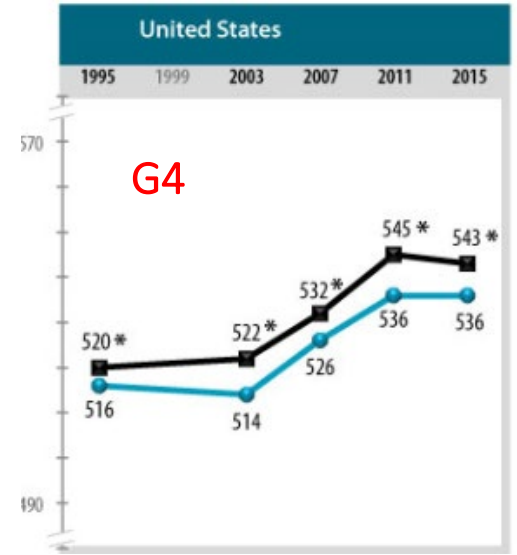
Is girls' academic performance in math really lower than that of boys?

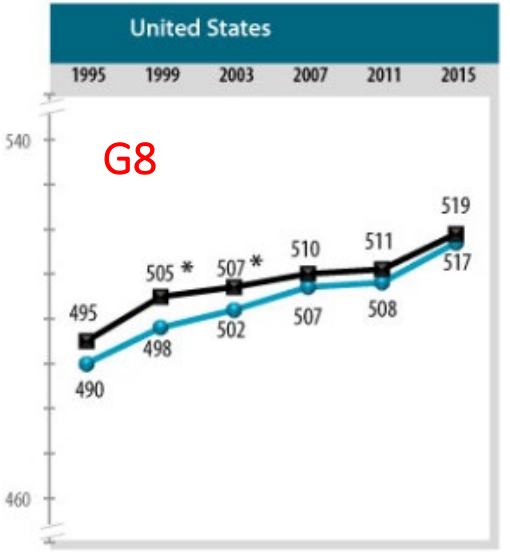
Trends in math achievements by gender (in Japan)





Trends in math achievements by gender (in US)





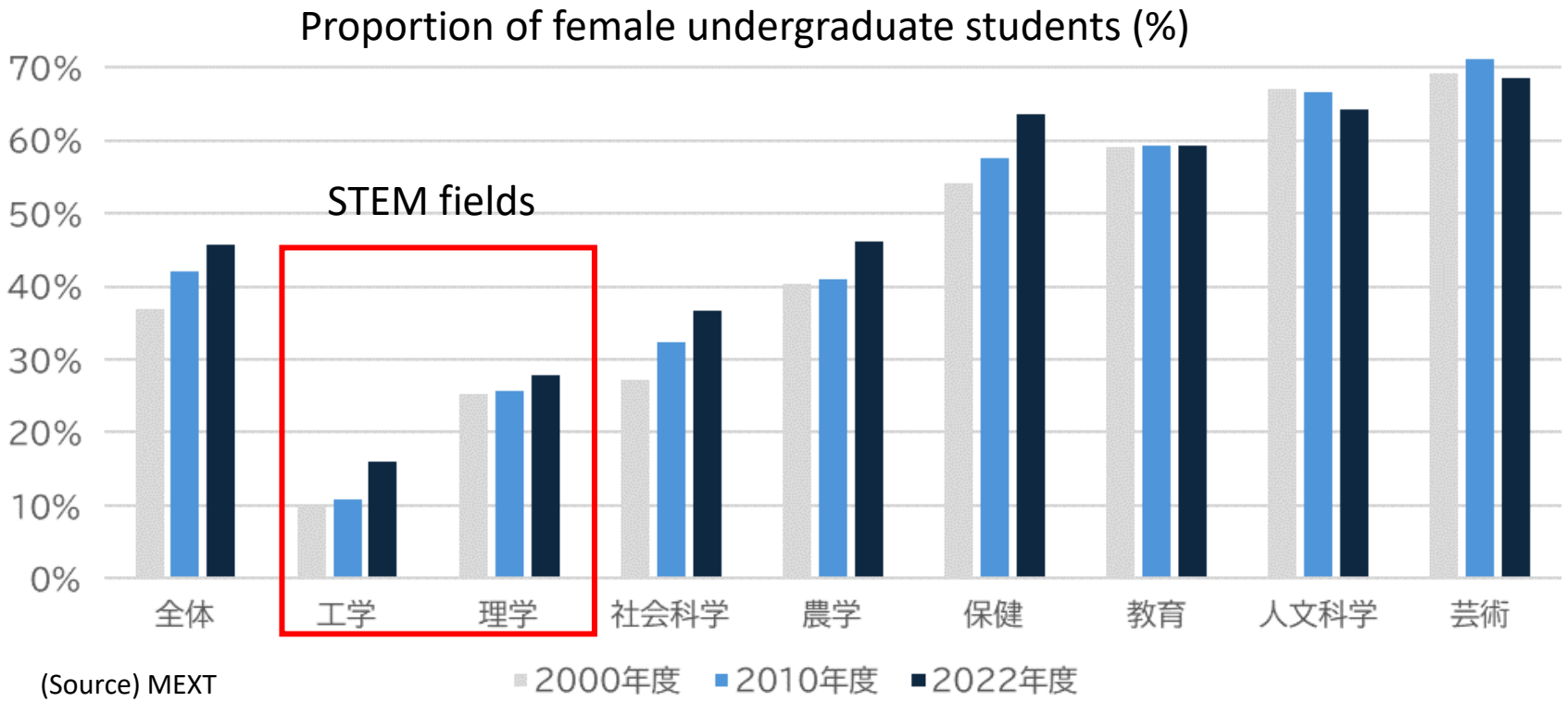
Average math achievement by gender

Country	Girls		Boys		Difference (Absolute Value)	Gender Difference	
	Percent of Students	Average Scale Score	Percent of Students	Average Scale Score		Girls Scored Higher	Boys Scored Higher
ψ Oman	48 (1.7)	420 (2.9)	52 (1.7)	388 (3.5)	32 (4.6)		
ψ Botswana (9)	51 (0.6)	400 (2.5)	49 (0.6)	381 (2.5)	19 (2.9)		
ж Jordan	50 (2.6)	395 (4.0)	50 (2.6)	376 (5.4)	19 (7.0)		
Thailand	54 (1.5)	440 (5.2)	46 (1.5)	422 (5.7)	18 (5.5)		
Bahrain	48 (0.9)	462 (2.4)	52 (0.9)	446 (2.2)	16 (3.6)		
ж Saudi Arabia	51 (1.6)	375 (5.1)	49 (1.6)	360 (7.1)	14 (8.2)		
United Arab Emirates	50 (2.5)	471 (3.5)	50 (2.5)	459 (4.0)	12 (6.4)		
Malaysia	50 (1.8)	470 (3.8)	50 (1.8)	461 (3.8)	9 (2.8)		
2 Singapore	49 (0.6)	626 (3.4)	51 (0.6)	616 (3.8)	9 (3.5)		
ψ Egypt	53 (2.3)	397 (5.5)	47 (2.3)	387 (5.1)	9 (6.7)		
ж South Africa (9)	51 (1.1)	376 (5.3)	49 (1.1)	369 (4.6)	7 (4.1)		
ψ Kuwait	50 (2.5)	396 (4.6)	50 (2.5)	389 (7.1)	7 (7.5)		
ψ Qatar	50 (3.0)	440 (3.2)	50 (3.0)	434 (4.5)	7 (4.9)		
Turkey	48 (0.8)	461 (4.8)	52 (0.8)	455 (5.3)	6 (3.6)		
Kazakhstan	49 (0.9)	531 (5.8)	51 (0.9)	525 (5.3)	6 (3.7)		
ψ Iran, Islamic Rep. of	48 (0.9)	438 (5.0)	52 (0.9)	435 (7.5)	3 (8.9)		
England	51 (1.6)	520 (5.2)	49 (1.6)	517 (4.8)	3 (5.6)		
Malta	49 (0.3)	495 (1.8)	51 (0.3)	492 (1.6)	3 (2.8)		
† New Zealand	51 (2.0)	494 (3.2)	49 (2.0)	491 (4.6)	3 (4.2)		
Japan	51 (1.0)	588 (3.1)	49 (1.0)	585 (3.0)	2 (4.2)		
ж Morocco	46 (0.7)	385 (2.3)	54 (0.7)	384 (2.6)	2 (2.0)		
1 2 Georgia	47 (0.9)	454 (3.9)	53 (0.9)	453 (4.0)	1 (4.0)		
Chinese Taipei	49 (0.8)	599 (2.6)	51 (0.8)	599 (3.0)	0 (2.8)		
Korea, Rep. of	47 (0.5)	605 (2.6)	53 (0.5)	606 (3.1)	1 (2.7)		
Norway (9)	50 (0.7)	511 (2.5)	50 (0.7)	512 (2.7)	1 (2.6)		
† United States	50 (0.6)	517 (3.3)	50 (0.6)	519 (3.2)	2 (2.0)		
Australia	51 (1.6)	504 (3.8)	49 (1.6)	506 (3.5)	2 (4.0)		
3 Israel	49 (1.2)	510 (4.3)	51 (1.2)	512 (4.8)	2 (3.9)		
Slovenia	48 (0.7)	515 (2.4)	52 (0.7)	518 (2.5)	2 (2.4)		
Lebanon	53 (1.6)	441 (3.7)	47 (1.6)	444 (4.5)	3 (3.9)		
2 Lithuania	50 (0.8)	510 (3.4)	50 (0.8)	513 (3.1)	3 (3.4)		
1 † Canada	51 (1.0)	525 (2.0)	49 (1.0)	530 (2.7)	4 (2.0)		
Ireland	50 (1.1)	521 (2.6)	50 (1.1)	526 (4.0)	5 (3.9)		
Hong Kong SAR	47 (2.1)	591 (4.7)	53 (2.1)	597 (6.0)	5 (5.7)		
2 Italy	49 (0.8)	491 (3.0)	51 (0.8)	498 (2.8)	7 (2.8)		
Sweden	48 (1.0)	497 (3.3)	52 (1.0)	504 (3.1)	7 (3.2)		
Hungary	50 (0.9)	510 (4.3)	50 (0.9)	519 (4.0)	9 (3.4)		
Russian Federation	49 (0.9)	533 (5.1)	51 (0.9)	543 (4.6)	9 (2.9)		
ψ Chile	48 (1.8)	418 (3.7)	52 (1.8)	436 (4.2)	18 (4.9)		
International Avg.	50 (0.2)	483 (0.6)	50 (0.2)	480 (0.7)			

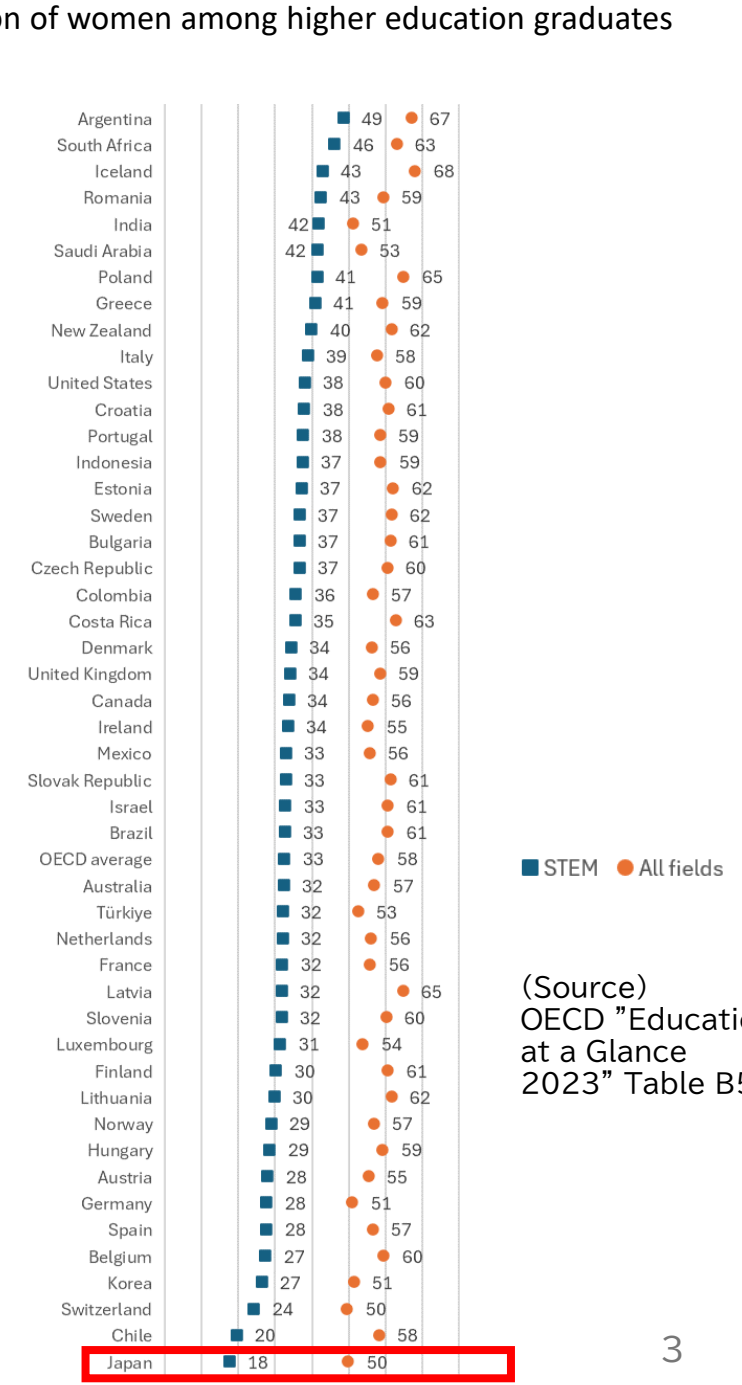
(Source) IEA TIMSS & PIRLS
https://timssandpirls.bc.edu/timss2015/international-results/timss-2015/mathematics/student-achievement/trends-in-mathematics-achievement-by-gender/

Gender gap in STEM majors in higher education

- The proportion of female undergraduates is low in STEM, whereas it is high in schools of education, the humanities, and the arts.
- The share of women among higher-education graduates in STEM fields ranks last out of 47 countries, and the gap between this share and women's share across all fields is also larger than in other countries.



(Source) MEXT



(Source)
OECD "Education at a Glance 2023" Table B5.4

Motivation and Research Question

- The gender gap in math performance was little explained by the low proportion of women who pursue STEM majors.

- Could the determinants of pursuing STEM majors differ between males and females?

- Although both male and female took high school education under the same institutional setting and are formally on equal footing between genders, there may exist environments, educational quality, or psychological pressure within the current institutional setting that, in practice, deter female students from pursuing STEM?

- What is the main driver of the gender gap in STEM choices in higher education?

Relevant Literature (comparative advantage?)

- Parents' educational attainment is low, performance gaps in STEM subjects have little influence on the choice of STEM majors in higher education (Contini et al., 2025).
- Using data from public high schools in Ontario, where students are assigned into different tracks, the gender gap in pursuing STEM majors at university can be largely explained by differences in the share of students who are "STEM-ready," meaning they have taken—and earned adequate grades in—the mathematics and science courses required in the final year of high school. Even among female students whose achievement was already high in Grade 11, they are less likely to take required STEM courses such as physics or calculus in Grade 12 (Card & Abigail, 2021). → Need to focus more on **the path prior to choosing the majors (At a majority of high schools in Japan, students will choose either science or liberal arts tracks at G11)**.

Two important drivers to female students away from STEM majors at high schools:

(1) Female ratio

- A 10-percentage-point increase in the female share of a student's incoming cohort lowers women's probability of choosing STEM majors by 1.4 percentage points and raises men's by 0.9. A higher female share leads girls to underestimate their own **comparative advantage** and avoid STEM majors. Girls with a female role model—specifically, a mother working in a STEM field—are less susceptible to such peer effects (Brenøe & Zölitz, 2020).

(2) Local percentile rank within the school

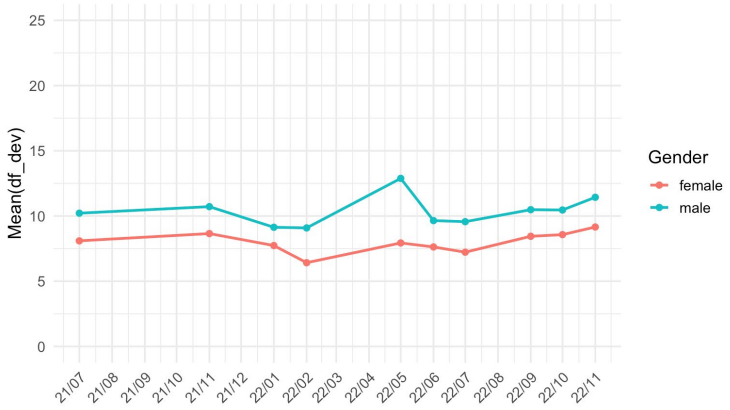
- A higher within-school rank in mathematics increases the likelihood of choosing a STEM major (Delaney & Devereux, 2021).After controlling absolute ability, female students whose STEM performance rank relatively higher within their class are more likely to choose a STEM track—both choosing the science track in high school and a STEM major at university. For boys, this effect is small or not statistically significant. This **comparative advantage** mechanism explains 12–18 percent of the roughly 34-percentage-point gender gap in STEM track choice in high school (Goulas et al., 2024).

Relevant Literature (confidence?)

Female students are less likely than males to pursue for highly selective institutions after controlling for absolute ability. A 10-percentile drop in confidence reduces the probability of applying to an elite preparatory class (CPGE) by 3.2 percentage points (Hakimov et al., 2023).

Figure 1: Difference between the deviation value (*hensa-chi*) of a student's first-choice university and the student's actual deviation value by gender (=a proxy of **confidence**)

all



STEM = Students who choose STEM preference at Grade10

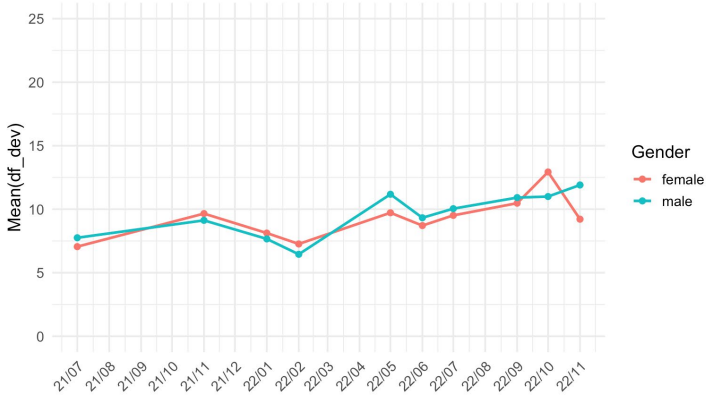
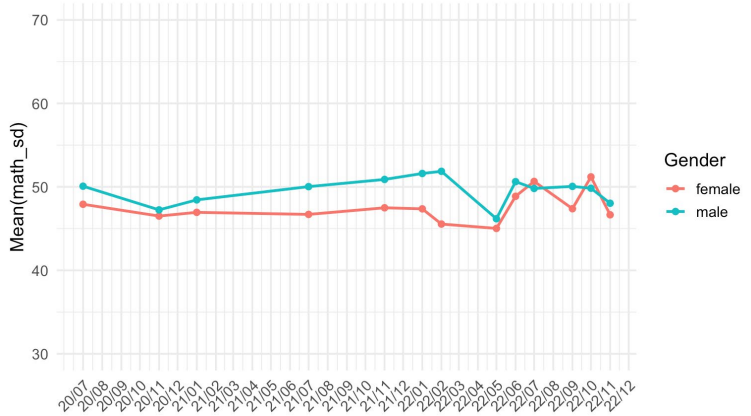


Figure 2: Students' actual deviation value by gender (=a proxy of **absolute ability**)

all



STEM = Students who choose STEM preference at Grade10

Data

- In anonymous Prefecture A, Company B administered private mock-exams for preparing university entrance exams. 12 schools agreed to use their data for this research. Company B's mock exams are widely administered nationwide in Japan and has a large number of examinees. Company B provides the mock-exams 3 times a year for Grade 10, 4 times a year for Grade 11, and 6 times a year in Grade 12, which mock tests are taken—and how many times—are decided at the school level.

- For this study, we constructed a panel dataset for students enrolled in Grades 10–12 during academic years of 2020–2022. Note that the sample is not nationally or Prefecture representative.

- The dataset includes the choice-list of which universities and departments students wish to attend at each time of mock-exam and the list of university and departments they actually applied, and finally enrolled in. It also includes their initial aspirations regarding track choice at G10—for example, whether they "firmly decided to choose science track at G10" (called *Sharp* in this study) and "Wants to choose the science track, but still undecided at G10" (called *Fuzzy* in this study).

- The advantage of this dataset are that (1) it allows us to observe how students changed their track from Grade 10 through Grade 12, and (2) it reveals at which stage and for what reasons students altered their intended track.

Variable Description

	Valuales	Description
	stem_start	Students who choose STEM preference at Grade10 =1 （named "STEM"）
	female_ratio	The Ratio of female students in the class (take the value between 0 and 1)
Absolute ability in math	math_sd	Standard deviation value for math in each mock exam
Absolute STEM advantage	relative_math_sd	Deviation value for math / Deviation value for all subjects excluding math in each mock exam
Confidence	df_dev	Deviation value of first preference university - Deviation value of all subjects in mock exams
	local_ave	Average local_percentile_rank_math at grade 10
	local_math_ave	Average local_relative_math_rank at grade 10
Local percentile rank in math	local_percentile_rank_math	The relative ranking of math scores within the school
Comparative STEM advantage	local_relative_math_rank	The relative ranking of relative math sd within the school
	stem_exam	Students who took the STEM department exam = 1
	stem_pass	Students who passed the STEM department exam = 1
	stem_enter	Students who enrolled in the STEM department = 1
	drop_stem_all	Students who withdrew from the STEM preference at some point = 1
	drop_stem_after	Students who withdrew from the STEM preference after Grade11 = 1
	to_liveral arts_all	Students who changed their STEM preference to liberal arts at some point = 1
	to_liveral arts_after	Students who changed their STEM preference to liberal arts after Grade11 = 1
	negative_local	Average local_percentile_rank_math after grades 2 and is lower than the average local_percentile_rank_math at grade 1 = 1
	negative_relative_math_sd	Average relative_math_sd after grades 2 and is lower than the average relative_math_sd at grade 1 = 1
	negative_local_relative_math_rank	Average local_relative_math_rank after grades 2 and is lower than the average local_relative_math_rank at grade 1 = 1

- The deviation value what is called “hensa-chi”: a statistical index (mean=50, sd=10) widely used in Japan to express how a student’s test score compares to the national (or test-taking group’s) average. It is conceptually similar to a standard score (z-score) but is rescaled for easier interpretation for teachers and students.
- The definition of STEM majors: in the field of science or engineering.

(Note) Brenøe and Zölitz (2020) define STEM according to the International Standard Classification of Education (ISCED): natural sciences, mathematics, and statistics (ISCED-05); information and communication technologies (ISCED-06); and engineering, manufacturing, and construction (ISCED-07).

Descriptive statistics (Whole Sample)

ALL

Men

Valuables	N	Mean	Min	Max	Sd
stem_start	46484	0.257	0.000	1.000	0.437
standard_deviation	52845	48.343	18.600	87.100	9.223
female_ratio	61262	0.414	0.000	0.833	0.144
math_sd	43289	49.673	22.100	86.100	9.007
relative_math_sd	43285	1.029	0.525	1.985	0.158
df_dev	34207	10.140	-34.600	50.100	9.140
local_ave	59321	0.554	0.000	1.000	0.255
local_math_ave	59321	0.526	0.000	0.994	0.242
local_percentile_rank_math	43289	0.551	0.000	1.000	0.292
local_relative_math_rank	43285	0.535	0.000	0.999	0.290
stem_exam	60318	0.455	0.000	1.000	0.498
stem_pass	60259	0.381	0.000	1.000	0.453
stem_enter	60223	0.289	0.000	1.000	0.486
drop_stem_all	57878	0.173	0.000	1.000	0.378
drop_stem_after	57878	0.096	0.000	1.000	0.295
to_liveral_arts_all	58031	0.129	0.000	1.000	0.336
to_liveral_arts_after11	58031	0.066	0.000	1.000	0.249
negative_local	61262	0.315	0.000	1.000	0.464
negative_relative_math_sd	61262	0.390	0.000	1.000	0.488
negative_local_relative_math_rank	61262	0.374	0.000	1.000	0.484

Women

Valuables	N	Mean	Min	Max	Sd
stem_start	44805	0.043	0.000	1.000	0.202
standard_deviation	54822	48.381	19.300	90.900	8.556
female_ratio	61057	0.594	0.091	1.000	0.230
math_sd	44113	47.281	25.500	79.700	7.621
relative_math_sd	44106	0.962	0.566	1.690	0.137
df_dev	36328	8.009	-35.900	45.900	8.896
local_ave	59863	0.510	0.000	1.000	0.239
local_math_ave	59867	0.443	0.000	0.982	0.226
local_percentile_rank_math	44105	0.499	0.000	1.000	0.279
local_relative_math_rank	44106	0.460	0.000	0.999	0.284
stem_exam	59421	0.285	0.000	1.000	0.451
stem_pass	59359	0.109	0.000	1.000	0.251
stem_enter	59291	0.068	0.000	1.000	0.311
drop_stem_all	59564	0.076	0.000	1.000	0.265
drop_stem_after	59564	0.046	0.000	1.000	0.210
to_liveral_arts_all	59518	0.053	0.000	1.000	0.225
to_liveral_arts_after11	59518	0.033	0.000	1.000	0.178
negative_local	61057	0.381	0.000	1.000	0.486
negative_relative_math_sd	61057	0.462	0.000	1.000	0.499
negative_local_relative_math_rank	61057	0.458	0.000	1.000	0.498

Descriptive statistics (STEM)

STEM

Men

Valuables	N	Mean	Min	Max	Sd
standard_deviation	10580	47.468	19.200	86.100	8.555
female_ratio	11941	0.378	0.000	0.750	0.136
math_sd	9888	49.862	27.900	86.100	8.934
relative_math_sd	9884	1.068	0.645	1.871	0.160
df_dev	6872	9.400	-34.600	44.800	8.885
local_ave	11656	0.637	0.010	0.997	0.242
local_math_ave	11656	0.596	0.019	0.993	0.236
local_percentile_rank_math	9888	0.605	0.000	1.000	0.283
local_relative_math_rank	9884	0.569	0.000	0.999	0.286
stem_exam	11784	0.703	0.000	1.000	0.457
stem_pass	11769	0.624	0.000	1.000	0.484
stem_enter	11660	0.512	0.000	1.000	0.500
drop_stem_all	11527	0.403	0.000	1.000	0.491
drop_stem_after	11290	0.163	0.000	1.000	0.369
to_bunkei_all	11344	0.450	0.000	1.000	0.497
to_bunkei_after	11344	0.110	0.000	1.000	0.313
negative_local	11941	0.318	0.000	1.000	0.466
negative_relative_math_sd	11941	0.444	0.000	1.000	0.497
negative_local_relative_math_rank	11941	0.369	0.000	1.000	0.482

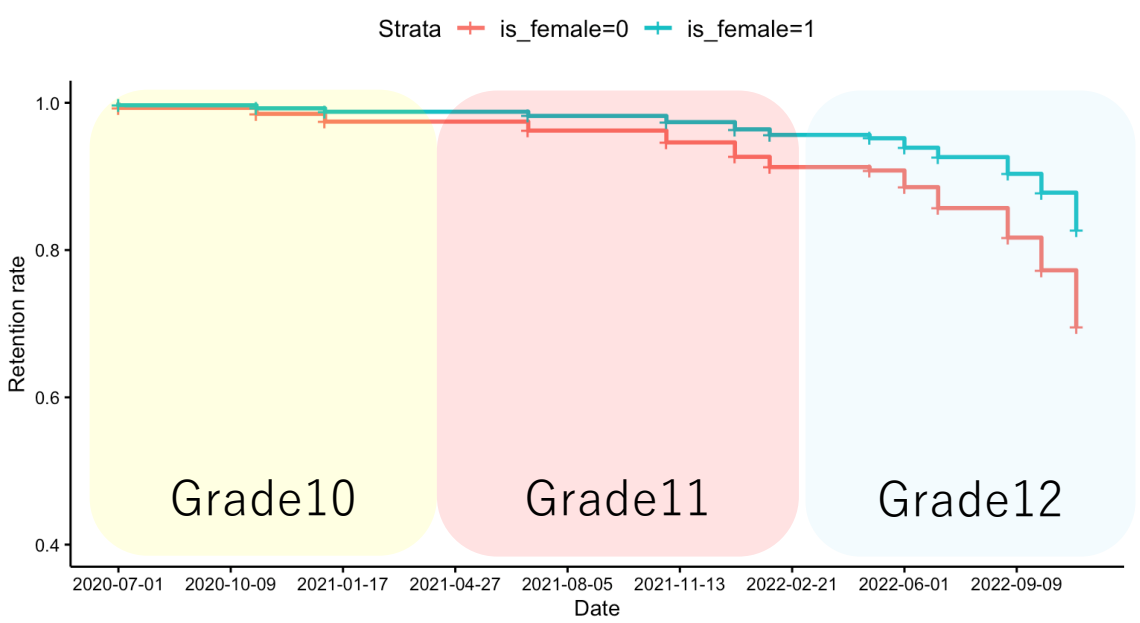
Women

Valuables	N	Mean	Min	Max	Sd
standard_deviation	1712	48.319	27.600	78.900	7.832
female_ratio	1914	0.562	0.200	1.000	0.249
math_sd	1592	49.426	32.900	78.900	7.836
relative_math_sd	1592	1.023	0.679	1.394	0.138
df_dev	1183	9.009	-30.100	34.500	7.591
local_ave	1908	0.699	0.167	0.978	0.201
local_math_ave	1908	0.570	0.106	0.926	0.209
local_percentile_rank_math	1592	0.625	0.019	1.000	0.263
local_relative_math_rank	1592	0.532	0.000	0.994	0.269
stem_exam	1914	0.601	0.000	1.000	0.490
stem_pass	1914	0.592	0.000	1.000	0.492
stem_enter	1914	0.397	0.000	1.000	0.489
drop_stem_all	1845	0.590	0.000	1.000	0.492
drop_stem_after	1845	0.297	0.000	1.000	0.457
to_bunkei_all	1845	0.719	0.000	1.000	0.450
to_bunkei_after	1845	0.195	0.000	1.000	0.396
negative_local	1914	0.296	0.000	1.000	0.456
negative_relative_math_sd	1914	0.307	0.000	1.000	0.461
negative_local_relative_math_rank	1914	0.331	0.000	1.000	0.471

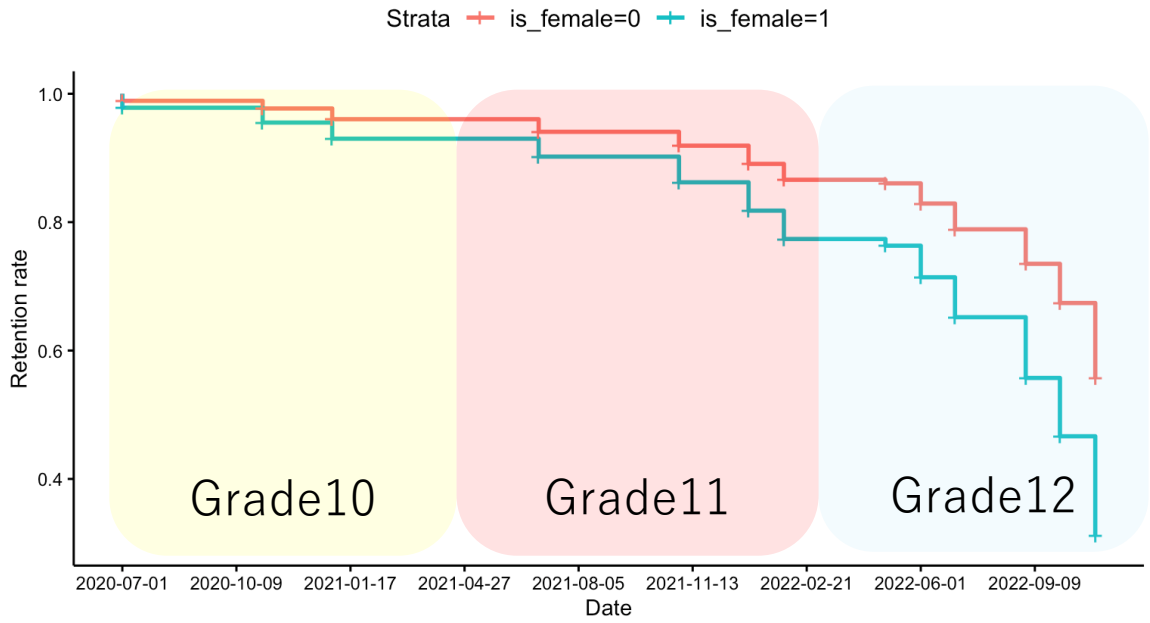
Definition of “STEM”: Students who firmly decided to choose science track at G10 and wanted to choose the science track, but still undecided at G10

Kaplan-Meier Survival Curve

Whole



STEM



Model

$$Y_{isct} = \alpha X_{isct} + Z'_{isct}\beta + \gamma_c + \delta_s + \zeta_t + \varepsilon_{isct}$$

i : $Individual$

s : $school$

c : $class$

t : $time$

Y_{isct} : $Outcome$ （Drop STEM）

X_{isct} : Female Ratio

γ_c : Class Fixed Effect

δ_i : Indivitual Fixed Effect

ζ_t : Mock Exam Fixed Effect

Z'_{isct} : Vector of Control Valuables

（local_percentile_rank_math, math_sd, df_dev, pass_non_general,female_ratio,local_relative_math_rank, relative_math_sd）

ε_{isct} : Error Terms

Results: Whole sample

- Results are consistent with previous literature.
- A **female ratio within class** (female_ratio) is negatively correlated with the likelihood of choosing STEM majors and switching to Liberal Arts (although it is unclear whether this effect differs by gender).
- A **local rank in math** (local_percentile_rank_math) is correlated with with a lower probability of dropping from STEM and switching to Liberal Arts.
- A higher **absolute ability** in math(math_sd) is negatively correlated with the likelihood of choosing STEM majors and switching to Liberal Arts.

ALL

	Drop STEM			To Liberal Arts		
	All	Men	Women	All	Men	Women
female_ratio	0.358*	0.296**	0.570**	0.354**	0.425**	0.459**
	(0.133)	(0.099)	(0.152)	(0.112)	(0.138)	(0.140)
local_percentile_rank_math	-0.296**	-0.368*	-0.152+	-0.308**	-0.442*	-0.024
	(0.079)	(0.137)	(0.085)	(0.102)	(0.176)	(0.091)
math_sd	0.016**	0.016*	0.014*	0.020***	0.022**	0.011*
	(0.005)	(0.006)	(0.006)	(0.005)	(0.007)	(0.005)
df_dev	0.001	0.001	-0.000	0.003**	0.002+	0.001+
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Relative_math_sd	0.056	-0.059	0.294	0.266	0.293	0.181
	(0.207)	(0.220)	(0.412)	(0.291)	(0.355)	(0.383)
local_relative_math_rank	-0.146	-0.102	-0.215	-0.271+	-0.289+	-0.190
	(0.100)	(0.106)	(0.166)	(0.133)	(0.160)	(0.154)
Num.Obs.	14140	9963	4144	10616	6993	3590

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Results: STEM

- A **female ratio** within class (female_ratio) is negatively correlated with the likelihood of choosing STEM majors and switching to Liberal Arts (although it is unclear whether this effect differs by gender).
- A **local rank in math** (local_percentile_rank_math) and **absolute STEM advantage** (Relative_math_sd) are correlated with with a lower probability of dropping from STEM and switching to Liberal Arts .
- A higher **absolute ability** in math(math_sd) is negatively correlated with the likelihood of switching to Liberal Arts.

STEM

	Drop STEM			To Liberal Arts		
	All	Men	Women	All	Men	Women
female_ratio	0.648*	0.522*	1.086*	1.097***	1.062***	1.192***
	(0.243)	(0.202)	(0.445)	(0.132)	(0.228)	(0.180)
local_percentile_rank_math	-0.509*	-0.523*	-0.486	-1.062***	-1.094**	-0.935*
	(0.185)	(0.227)	(0.303)	(0.251)	(0.290)	(0.290)
math_sd	0.019*	0.020+	0.021	-0.057***	-0.057***	-0.061**
	(0.009)	(0.009)	(0.015)	(0.010)	(0.010)	(0.016)
df_dev	0.002	0.004+	-0.008*	0.004*	0.004*	0.006*
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Relative_math_sd	-0.540*	-0.736*	-3.548+	-0.895*	-0.911*	-0.552
	(0.245)	(0.260)	(1.882)	(0.359)	(0.393)	(0.879)
local_relative_math_rank	0.092	0.198	-1.767*	0.099	0.136	-0.669
	(0.110)	(0.157)	(0.618)	(0.140)	(0.151)	(0.393)
Num.Obs.	7059	6038	1021	3535	3068	467

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Results: Whole Sample * after Grade11

- For female, **local rank in math** (local_percentile_rank_math) and **comprative STEM advantage** (local_relative_math_rank) are correlated with with a lower probability of dropping from STEM and switching to Liberal Arts .
- A higher **absolute ability in math**(math_sd) is negatively correlated with the likelihood of switching to Liberal Arts.

ALL

	Drop STEM			To Liberal Arts		
	All	Men	Women	All	Men	Women
female_ratio * After_Grade11	0.373	0.495	0.416	-0.131	-0.025	0.511
	(0.433)	(0.469)	(0.586)	(0.466)	(0.447)	(0.742)
local_percentile_rank_math * After_Grade11	-0.618+	-0.400	-1.380**	-0.320	0.036	-1.322*
	(0.304)	(0.369)	(0.449)	(0.357)	(0.423)	(0.564)
math_sd * After_Grade11	0.038**	0.030*	0.074***	0.022+	0.010	0.069**
	(0.010)	(0.011)	(0.017)	(0.011)	(0.012)	(0.018)
df_dev * After_Grade11	0.005	0.006	0.006	-0.002	-0.003	0.006+
	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)
Relative_math_sd * After_Grade11	-0.288	-0.281	-0.322	-0.657	-1.058	-0.904
	(0.623)	(0.836)	(0.985)	(0.712)	(0.905)	(1.027)
local_relative_math_rank * After_Grade11	-0.345	-0.268	-0.941*	-0.204	0.118	-1.389+
	(0.380)	(0.468)	(0.236)	(0.356)	(0.479)	(0.659)
After_Grade11	-0.662	-0.537	-2.262**	0.487	1.061	-2.529*
	(0.423)	(0.512)	(0.741)	(0.515)	(0.650)	(0.977)
Num.Obs.	13715	9680	4002	10312	6758	3521

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Results: STEM * after Grade11

- For female, **absolute STEM advantage**(relative_math_sd) and **compative STEM advantage** (local_relative_math_rank) are correlated with with a lower probability of dropping from STEM.
- A **female ratio** within class (female_ratio) is correlated with with a lower probability of switching to Liberal Arts.

STEM

	Drop STEM			To Liberal Arts		
	All	Men	Women	All	Men	Women
female_ratio * After_Grade11	-0.136	-0.047	-0.698	-0.787*	-0.718+	-1.172*
	(0.385)	(0.399)	(0.541)	(0.369)	(0.367)	(0.378)
local_percentile_rank_math * After_Grade11	-0.283	-0.177	-0.285	0.522+	0.608+	-0.151
	(0.232)	(0.273)	(0.337)	(0.276)	(0.315)	(0.293)
math_sd * After_Grade11	0.027*	0.026*	0.006	-0.025	-0.023	-0.006
	(0.010)	(0.009)	(0.016)	(0.015)	(0.016)	(0.015)
df_dev * After_Grade11	0.005	0.003	0.010	-0.003	-0.002	-0.002
	(0.004)	(0.003)	(0.006)	(0.003)	(0.003)	(0.002)
Relative_math_sd * After_Grade11	0.035	-0.077	-0.725**	0.043	-0.466	-1.590
	(0.559)	(0.587)	(0.012)	(0.660)	(0.705)	(1.338)
local_relative_math_rank * After_Grade11	-0.301	-0.315	-0.209*	-0.144	-0.042	-0.638
	(0.235)	(0.229)	(0.050)	(0.197)	(0.159)	(0.674)
After_Grade11	-0.458	-0.359	1.412	1.621***	1.865***	2.812*
	(0.315)	(0.282)	(1.114)	(0.355)	(0.370)	(1.014)
Num.Obs.	6928	5907	1021	3415	2948	467

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Results: Mechanism on why and when “STEM” gave up

Female students are affected by local rank in math . If a student’s local rank in math falls at G11 relative to G10, the probability of choosing STEM fields becomes lower.

(Note) Local_top: a dummy variable equal to 1 if a student’s local_percentile_rank_math was in the top 50 percent in Grade 10.
relative_math_sd_top: a dummy variable equal to 1 if a student’s Relative_math_sd was in the top 50 percent in Grade 10.

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Nagtive_local	-0.052**	-0.039	0.052**	-0.037	0.026	0.137*
	(0.016)	(0.030)	(0.015)	(0.079)	(0.072)	(0.061)
Num.Obs.	49541	23932	25576	6934	5913	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Local_Top * Nagtive_local	0.021	0.052	0.127*	-0.088	-0.036	0.313*
	(0.028)	(0.040)	(0.048)	(0.136)	(0.140)	(0.106)
Local_Top	-0.002	-0.057+	0.045+	-0.024	-0.036	-0.370
	(0.020)	(0.029)	(0.023)	(0.061)	(0.060)	(0.271)
Nagtive_local	-0.061**	-0.083**	-0.026	0.002	0.031	-0.364+
	(0.019)	(0.024)	(0.023)	(0.106)	(0.097)	(0.200)
Num.Obs.	48686	23426	25260	6724	5706	1018

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral_Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Nagtive_local	-0.043**	-0.040**	-0.035+	0.027	0.069+	0.095
	(0.014)	(0.011)	(0.017)	(0.059)	(0.038)	(0.152)
Num.Obs.	49436	23877	25526	6930	5909	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral_Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Local_Top * Nagtive_local	0.027	0.065*	-0.017	0.017	0.056	0.201*
	(0.026)	(0.031)	(0.039)	(0.096)	(0.081)	(0.040)
Local_Top	-0.015	-0.057*	0.019	0.002	0.027	-0.599**
	(0.015)	(0.022)	(0.017)	(0.047)	(0.048)	(0.171)
Nagtive_local	-0.059***	-0.089***	-0.022	0.017	0.043	0.283
	(0.013)	(0.016)	(0.019)	(0.087)	(0.057)	(0.312)
Num.Obs.	48578	23368	25210	6717	5699	1018

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Results: Mechanism on why and when "STEM" gave up

Female students are affected by absolute STEM advantage. If a student's Relative_math_sd falls at G11 relative to G10, the probability of choosing STEM fields becomes lower. (For men, the probability of choosing STEM increase.)

(Note) relative_math_sd_top: a dummy variable equal to 1 if a student's Relative_math_sd was in the top 50 percent in Grade 10.

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Nagtive_relative_math_sd	0.009	0.025+	0.006	0.012	0.028	0.129***
	(0.009)	(0.014)	(0.012)	(0.045)	(0.050)	(0.024)
Num.Obs.	49541	23932	25576	6934	5913	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Relative_math_sd_Top * Nagtive_relative_math_sd	-0.027	-0.077*	0.000	-0.168	-0.099	0.153**
	(0.022)	(0.033)	(0.032)	(0.148)	(0.170)	(0.018)
Relative_math_sd_Top	0.055***	0.054*	0.050**	0.023	-0.074	0.120
	(0.014)	(0.025)	(0.014)	(0.095)	(0.115)	(0.173)
Nagtive_relative_math_sd	0.038*	0.078**	0.028*	0.121	0.072	0.123
	(0.015)	(0.025)	(0.013)	(0.129)	(0.159)	(0.163)
Num.Obs.	48686	23426	25260	6724	5706	1018

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral_Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Nagtive_relative_math_sd	-0.008	-0.006	-0.002	-0.035	-0.045	0.148
	(0.009)	(0.016)	(0.011)	(0.031)	(0.040)	(0.113)
Num.Obs.	49436	23877	25526	6930	5909	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral_Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Relative_math_sd_Top * Nagtive_relative_math_sd	-0.030+	-0.082*	0.011	-0.210*	-0.264*	0.417
	(0.016)	(0.030)	(0.023)	(0.085)	(0.102)	(0.259)
Relative_math_sd_Top	0.052***	0.066**	0.037*	0.131+	0.132+	0.159
	(0.010)	(0.023)	(0.015)	(0.065)	(0.065)	(0.104)
Nagtive_relative_math_sd	0.021+	0.055*	0.012	0.136+	0.168	0.052
	(0.012)	(0.025)	(0.012)	(0.069)	(0.097)	(0.116)
Num.Obs.	48578	23368	25210	6717	5699	1018

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Results: Mechanism on why and when “STEM” gave up

Female students are affected by comparative STEM advantage . If a student’s local_relative_math_rank falls at G11 relative to G10, the probability of choosing STEM fields becomes lower.

(Note) relative_local_math_rank_top: a dummy variable equal to 1 if a student’s local_relative_math_rank was in the top 50 percent in Grade 10.

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Nagtive_local_relative_math_rank	0.009	0.027	0.005	0.026	0.044	0.086*
	(0.014)	(0.023)	(0.010)	(0.061)	(0.074)	(0.014)
Num.Obs.	49541	23932	25576	6934	5913	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Relative_local_math_rank_Top * Nagtive_local_relative_math_rank	-0.013	-0.047	0.011	0.115	0.130	0.509*
	(0.027)	(0.047)	(0.037)	(0.176)	(0.189)	(0.133)
Relative_local_math_rank_Top	0.062***	0.063**	0.047*	-0.134	-0.160	-0.257
	(0.012)	(0.019)	(0.017)	(0.108)	(0.117)	(0.203)
Nagtive local relative math rank	0.040*	0.070+	0.026	-0.089	-0.092	-0.235
	(0.018)	(0.035)	(0.015)	(0.164)	(0.163)	(0.219)
Num.Obs.	48686	23426	25260	6724	5706	1018

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral_Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Nagtive_local_relative_math_rank	-0.005	-0.002	0.001	-0.022	-0.024	0.097
	(0.012)	(0.021)	(0.012)	(0.040)	(0.043)	(0.114)
Num.Obs.	49436	23877	25526	6930	5909	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral_Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Relative_local_math_rank_Top * Nagtive_local_relative_math_rank	-0.035*	-0.095**	0.023	-0.090	-0.102	0.325**
	(0.017)	(0.032)	(0.028)	(0.116)	(0.140)	(0.092)
Relative_local_math_rank_Top	0.067***	0.083**	0.043*	0.076	0.036	0.262*
	(0.014)	(0.022)	(0.016)	(0.080)	(0.084)	(0.097)
Nagtive_local_relative_math_rank	0.036**	0.070**	0.017	0.053	0.040	0.006
	(0.010)	(0.023)	(0.010)	(0.065)	(0.078)	(0.095)
Num.Obs.	48578	23368	25210	6717	5699	1018

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Results: Effects of female ratio

When the female ratio is high, the probability of dropping from the stem even with negative_local, negative_relative_math_sd and negative_local_relative_math_rank doesn't increase.

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Female_Ratio * Nagtive_local	-0.000	-0.044	0.023	-0.167	0.000	0.186
	(0.040)	(0.099)	(0.051)	(0.147)	(0.246)	(0.180)
Female_Ratio	-0.098+	-0.102	-0.059	0.011	0.014	-0.298
	(0.051)	(0.166)	(0.035)	(0.101)	(0.154)	(0.329)
Nagtive_local	-0.038	-0.005	-0.057	0.056	0.050	-0.476***
	(0.030)	(0.044)	(0.042)	(0.083)	(0.082)	(0.071)
Num.Obs.	49541	23932	25576	6934	5913	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Female_Ratio * Nagtive_local	0.013	0.032	-0.002	-0.085	0.178	-0.335
	(0.029)	(0.065)	(0.059)	(0.153)	(0.179)	(0.412)
Female_Ratio	-0.080+	-0.081	-0.036	-0.051	-0.054	0.001
	(0.041)	(0.130)	(0.037)	(0.133)	(0.162)	(0.155)
Nagtive_local	-0.040+	-0.042	-0.029	0.076	0.012	0.362
	(0.023)	(0.027)	(0.049)	(0.084)	(0.082)	(0.393)
Num.Obs.	49436	23877	25526	6930	5909	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Female_Ratio * Nagtive_relative_math_sd	-0.007	0.069	-0.010	0.150	0.217	0.071
	(0.043)	(0.085)	(0.056)	(0.167)	(0.296)	(0.174)
Female_Ratio	-0.105*	-0.152	-0.051	-0.113	-0.041	-0.033
	(0.042)	(0.141)	(0.030)	(0.097)	(0.170)	(0.258)
Nagtive_relative_math_sd	0.012	-0.003	0.012	-0.048	-0.051	0.088
	(0.024)	(0.037)	(0.040)	(0.074)	(0.113)	(0.094)
Num.Obs.	49541	23932	25576	6934	5913	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Female_Ratio * Nagtive_relative_math_sd	0.006	0.046	-0.012	0.049	0.103	-0.119
	(0.034)	(0.080)	(0.051)	(0.146)	(0.224)	(0.330)
Female_Ratio	-0.084*	-0.098	-0.035	-0.090	-0.024	-0.144
	(0.040)	(0.129)	(0.024)	(0.112)	(0.189)	(0.323)
Nagtive_relative_math_sd	-0.012	-0.025	0.006	-0.055	-0.082	0.218
	(0.019)	(0.031)	(0.038)	(0.049)	(0.080)	(0.295)
Num.Obs.	49436	23877	25526	6930	5909	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Drop STEM

	ALL			STEM		
	All	Men	Women	All	Men	Women
Female_Ratio * Nagtive_local_relative_math_rank	-0.013	-0.001	0.014	-0.094	0.004	-0.172
	(0.039)	(0.082)	(0.060)	(0.246)	(0.319)	(0.284)
Female_Ratio	-0.102+	-0.120	-0.061	-0.033	0.027	0.066
	(0.051)	(0.137)	(0.036)	(0.119)	(0.169)	(0.276)
Nagtive_local_relative_math_rank	0.015	0.027	-0.003	0.064	0.043	0.189
	(0.029)	(0.044)	(0.043)	(0.113)	(0.112)	(0.134)
Num.Obs.	49541	23932	25576	6934	5913	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

To_Liveral Arts

	ALL			STEM		
	All	Men	Women	All	Men	Women
Female_Ratio * Nagtive_local_relative_math_rank	0.009	0.024	0.007	-0.070	0.046	-0.394
	(0.035)	(0.064)	(0.067)	(0.164)	(0.199)	(0.425)
Female_Ratio	-0.086+	-0.088	-0.044	-0.048	-0.006	-0.017
	(0.048)	(0.122)	(0.034)	(0.129)	(0.187)	(0.329)
Nagtive_local_relative_math_rank	-0.009	-0.012	-0.003	0.006	-0.040	0.332
	(0.027)	(0.030)	(0.051)	(0.061)	(0.073)	(0.323)
Num.Obs.	49436	23877	25526	6930	5909	1021

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Policy Implication

- The increasing number of women-only admissions (comprehensive selection for women only in STEM departments) and other diverse admission systems may increase the likelihood of women choosing STEM fields.
- The necessity of incorporating major guidance that takes into account the different influences of gender and STEM aspirations on major choices.
- As previous studies mentioned, encouraging active contact with role models of female is also a possible effective measure to avoid underestimating comparative advantage.

Reference

Bobba, M., & Frisancho, V. (2022). Self-perceptions about academic achievement: Evidence from Mexico City. Journal of Econometrics, 231(1), 58-73.

Brenøe, A. A., & Zölitz, U. (2020). Exposure to more female peers widens the gender gap in STEM participation. Journal of Labor Economics, 38(4), 1009-1054.

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

Contini, D., Di Tommaso, M., Maccagnan, A. & Mendolia, S. (2025). Is it a Matter of Skills? High School Choices and the Gender Gap in STEM. The B.E. Journal of Economic Analysis & Policy. https://doi.org/10.1515/bejeap-2024-0461

Delaney, J. M., & Devereux, P. J. (2021). High school rank in math and English and the gender gap in STEM. Labour Economics, 69, 101969.

Gong, Y., Skura, T., Stinebrickner, R., & Stinebrickner, T. (2025). Fertility Beliefs and Outcomes: the Role of Relationship Status and Attractiveness. The Economic Journal, 135(668), 1410-1431.

Goulas, S., Griselda, S., & Megalokonomou, R. (2024). Comparative advantage and gender gap in STEM. Journal of Human Resources, 59(6), 1937-1980.

Hakimov, R., Schmacker, R., & Terrier, C. (2023). Confidence and college applications: Evidence from a randomized intervention (No. 962). Working Paper.

Makarova, E., Aeschlimann, B., & Herzog, W. (2019, July). The gender gap in STEM fields: The impact of the gender stereotype of math and science on secondary students' career aspirations. In Frontiers in education (Vol. 4, p. 60). Frontiers Media SA.

Ngo, D., & Dustan, A. (2024). Preferences, access, and the stem gender gap in centralized high school assignment. American Economic Journal: Applied Economics, 16(4), 257-287.

Sansone, D. (2019). Teacher characteristics, student beliefs, and the gender gap in STEM fields. Educational Evaluation and Policy Analysis, 41(2), 127-144.

Verdugo-Castro, S., García-Holgado, A., & Sánchez-Gómez, M. C. (2022). The gender gap in higher STEM studies: A systematic literature review. Heliyon, 8(8).

Appendix

Preference

Sharp_Grade1	Men	
	N	Mean
high_useful_math	942	0.709
notgood_math	926	0.060
good_math	930	0.454
local_percentile_rank_math	946	0.647
local_relative_math_rank	946	0.620
math_sd	946	49.049
relative_math_sd	946	1.073

Sharp_Grade1	Women	
	N	Mean
high_useful_math	132	0.727
notgood_math	136	0.059
good_math	136	0.471
local_percentile_rank_math	136	0.666
local_relative_math_rank	136	0.539
math_sd	136	48.815
relative_math_sd	136	1.021

Fuzzy_Grade1	Men	
	N	Mean
high_useful_math	412	0.592
notgood_math	408	0.157
good_math	404	0.267
local_percentile_rank_math	416	0.562
local_relative_math_rank	416	0.525
math_sd	416	46.516
relative_math_sd	416	1.022

Fuzzy_Grade1	Women	
	N	Mean
high_useful_math	108	0.667
notgood_math	108	0.074
good_math	108	0.296
local_percentile_rank_math	108	0.639
local_relative_math_rank	108	0.554
math_sd	108	46.948
relative_math_sd	108	1.014

Sharp_Grade2-3	Men	
	N	Mean
high_useful_math	1222	0.606
notgood_math	1199	0.182
good_math	1190	0.385
local_percentile_rank_math	1161	0.656
local_relative_math_rank	1161	0.591
math_sd	1161	50.140
relative_math_sd	1161	1.081
negative_local	1225	0.266
negative_relative_math_sd	1225	0.348
negative_local_relative_math_rank	1225	0.291

Sharp_Grade2-3	Women	
	N	Mean
high_useful_math	203	0.626
notgood_math	199	0.302
good_math	189	0.317
local_percentile_rank_math	182	0.621
local_relative_math_rank	182	0.495
math_sd	182	47.352
relative_math_sd	182	0.977
negative_local	203	0.557
negative_relative_math_sd	203	0.212
negative_local_relative_math_rank	203	0.374

Fuzzy_Grade2-3	Men	
	N	Mean
high_useful_math	436	0.580
notgood_math	433	0.192
good_math	433	0.305
local_percentile_rank_math	398	0.640
local_relative_math_rank	397	0.555
math_sd	398	49.313
relative_math_sd	397	1.052
negative_local	436	0.342
negative_relative_math_sd	436	0.596
negative_local_relative_math_rank	436	0.507

Fuzzy_Grade2-3	Women	
	N	Mean
high_useful_math	179	0.726
notgood_math	175	0.120
good_math	175	0.194
local_percentile_rank_math	156	0.503
local_relative_math_rank	156	0.519
math_sd	156	43.330
relative_math_sd	156	0.979
negative_local	179	0.318
negative_relative_math_sd	179	0.274
negative_local_relative_math_rank	179	0.274