ONLINE FIRST ARTICLE
doi.org/10.362653/IJAS.17023

Statistica Applicata - Italian Journal of Applied Statistics

Received: 30.09.25/ **Accepted:** 15.05.25
Published online: 30.12.25

EQUAL OPPORTUNITIES TO "COUNT"?

A LONGITUDINAL POPULATION STUDY ABOUT CHOICE OF STEM AT UNIVERSITY BY GENDER

Patrizia Giannantoni[1], Patrizia Falzetti
National Institute for the Evaluation of the Educational System (INVALSI)
Rome, Italy

Abstract.

Extensive research has reported that Science Technology Engineering and Mathematics (STEM) disciplines are perceived as "male fields". This erratic belief is largely responsible for gender gap observed in school and to the underrepresentation of women in STEM majors at university and in the labour market. This paper examines the factors that influence boys' and girls' decisions to enrol in STEM, making use of a longitudinal population dataset obtained by joining together Italian data from Ministry of education, University Register and INVALSI census test. We investigate the role of math skills, as well as individual and family characteristics, as predictors of STEM academic choices. Key findings highlight the central influence of math ability on the choice of STEM studies, particularly when assessed through teachers' grades, and reveal an opposite effect of mother's education on STEM enrolment by gender: increasing the likelihood of STEM selection for girls while decreasing it for boys.

Keywords: *STEM, gender gap, University, math skills, population, longitudinal data*

[1] Corresponding author (patrizia.giannantoni@invalsi.it)

1. INTRODUCTION

The construction of gender roles, the gender gap in STEM disciplines throughout academic and professional trajectories and occupational segregation are topics widely investigated across multiple research domains. The European Institute for Gender Equality (EIGE, 2021) highlights significant horizontal segregation in higher STEM education, particularly in Southern European countries. This segregation in tertiary education and the labor market represents the culmination of a process that begins in early childhood education and extends throughout the schooling years.

Empirical studies demonstrate that gender disparities in mathematical achievement emerge by adolescence, with girls, on average, scoring lower than boys at age 15 (Borgonovi et al., 2018). This disparity is most pronounced among high-achieving students (Ellison & Swanson, 2010) and substantially influences university major selection, with women being significantly less likely to pursue STEM fields (Turner & Bowen, 1999; Card & Payne, 2021). Cross-national analyses confirm that while gender differences in mathematics performance are negligible in primary school, they expand significantly during adolescence (Contini et al., 2017; Borgonovi et al., 2018; Ellison & Swanson, 2010).

A broad consensus exists that these disparities are not rooted in biological differences but are instead shaped by an array of interrelated social factors (Osborne et al., 2003). Gender stereotypes, differential teacher expectations, familial attitudes and socio-economic constraints contribute to variations in self-confidence, aspirations, and academic choices among boys and girls. Eccles (2007) suggests that girls' lower inclination toward STEM careers is partially attributable to societal expectations that subtly direct them toward non-scientific fields. Similarly, research underscores the role of cultural norms, self-esteem, and perceptions of competence in shaping students' academic trajectories (Eccles & Wang, 2016).

Social representations and stereotypes regarding gender and science influence academic interests and outcomes, as demonstrated by Master and Meltzoff (2011). Their meta-analysis of international datasets, including the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA), confirms that gender equity in education is crucial not only for female students' mathematical

performance but also for their self-confidence and appreciation of mathematics (Else-Quest et al., 2010).

Understanding the mechanisms through which these stereotypes operate is vital for developing interventions aimed at mitigating gender disparities in STEM education and employment. Crucial moments in educational trajectories, such as transitions between schooling levels, significantly impact students' decisions to pursue STEM pathways. Research employing specialized survey instruments has examined the factors influencing university STEM enrollment decisions (Allegretti et al., 2014).

While numerous studies have examined these issues internationally, relatively few have focused on the Italian context. Priulla et al. (2024) applied machine learning techniques to predict university enrollment patterns based on Italian students' educational backgrounds, uncovering significant gender disparities in STEM participation. Additionally, Priulla and Attanasio (2023) used discrete-time multi-state Markov models to analyze gender differences in university completion rates, demonstrating the impact of prior academic performance on tertiary education outcomes. Barone and Assirelli (2019) further investigated gender segregation in higher education, concluding that school type significantly mediates gender disparities in STEM enrollment.

Given the substantial influence of high school type on university major selection, this study focuses on graduates from Scientific Lyceums—a subpopulation most likely to pursue STEM disciplines at university. Understanding the educational choices and academic performance of these students is essential for informing policies aimed at fostering gender equity in STEM fields.

2. RESEARCH QUESTIONS

In the existing literature, there are no comprehensive studies on the factors that influence the choice of STEM in tertiary education, especially at the population level in a country with a strong gender gap, such as Italy. The present study aims to gain a better understanding of these factors, focusing on students attending

Scientific Lyceum, who represent the elective subpopulation targeted to pursue a career in highly scientific.

This research aims at answering the following questions: to what extent is maths ability important/predictive of STEM enrolment? Higher scores in math at standardized test, or alternatively higher teachers' evaluation in math, are associated with increased probability of choice of STEM subjects at university? How influential is, in this framework, the social and family context?

Trying to answer these questions, the study sheds light on how all these factors operate jointly according to gender, focusing on similarities and differences in the effects that skills, familial and social factors have o boys and girls in the selection of STEM at university.

2. METHODS

2.1 Sample

An original dataset was created for this research project by combining three different administrative data sources: Italian Ministry of Education, National University Register and National Institute for the Evaluation of the Education and Training System (INVALSI) data[2]. INVALSI's mission is to evaluate the national school system as a whole and to provide schools with guidance for self-evaluation and improvement. The INVALSI surveys consist of standardised tests administered to students in Italian schools at 5 different grade levels. The second and fifth classes of primary school (grades 2 and 5), the third class of lower secondary school (grade 8) and the second and final classes (grades 10 and 12-13) of upper secondary school. The INVALSI tests are designed to measure students' basic skills in subjects such as mathematics, reading comprehension and English. The tests are administered on paper in grades 2 and 5, on paper in grades

[2] This study was conducted within the framework of the Agreement "From High School to Job Placement: An Analysis of University Careers and Student Mobility from Southern to Northern Italy" in collaboration with the Italian Ministry of Education and Merit, the Italian Ministry of University and Research and the University of Palermo.

8 and 10 until 2017, and on computer in lower and upper secondary schools from 2018 (computer-based tests). The idea behind the construction of the integrated dataset is to make it possible, for the first time with Italian data, to follow a student's academic career from entry into the education system to university enrolment and graduation. A crucial point of the analysis was the process of matching students in the three data sources: INVALSI has a "student individual code (SIDI)" as a unique identifier for each individual student, while the University Register uses the tax code to identify students, and finally the Ministry of Education has both the SIDI and tax code variables in its dataset. Through this two-way correspondence, it was possible to construct a combined, fully anonymised data source that, at the population level, matched the cohort of students enrolled in all Italian universities (public and private) in the 2019/2020 school year with their characteristics and skills in grade 13 of upper secondary school in 2018/19, in a retrospective longitudinal perspective.

2.2 Variables

After the matching process, we obtained a joint dataset with all the information about the students coming from the different sources. INVALSI provided the following variables: students' performance in all subjects (Italian, mathematics and English) from primary school until the last year of upper secondary school (grade 13), information on students' family background (e.g. parents' level of education, family socio-economic status, country of origin, language spoken at home), but also information on the geographical location of the school. The Ministry of Education provided: basic demographic information on students, year and grade obtained in the final examination at grade 13. University Register provided information related to university's first enrolment: academic year, university course of study, with the specification of the code of the course chosen and the "class of discipline" associated with it. For this research, we selected a subsample of university students composed by all students enrolled in Italian universities (public and private) in the academic year 2019/2020, who attended the last year of scientific lyceum in the previous year (2018/2019). We selected variables related to the following areas:

- Demographic characteristics
- Math Proficiency
- Socio-cultural status of student's family
- School characteristics
- STEM Enrolment at University

As individual characteristics, we considered, together with gender, the variables: regularity of study (regular/repeating) and migration background (native, I-generation, II-generation), as they have been widely shown to influence school careers and choices (OECD, 2015; Ismu 2020). For each student, demographic information from the Ministry of Education was used as a cross-validation for the same data from the INVALSI dataset.

For the assessment of ability in mathematics, both the teacher's marks in the first quarter of the year and the INVALSI test score (WLE score) in mathematics were considered. Both variables are numerical and were categorised into classes for the analysis. Specifically, the teacher's mark in mathematics is collected with a range from 1 to 10 and was classified as follows: 1 to 5 'insufficient', 6 'sufficient', 7 to 8 'good', 9 and 10 'excellent'. The INVALSI WLE score, elaborated on the principles of Rasch analysis, was classified into levels from 1 to 5 according to cut-off scores based on the descriptive proficiency level approach (Desimoni, 2018).

In relation to socio-economic status, two variables were considered: The Index of Economic, Social and Cultural Status (ESCS) and the mother's level of education. The ESCS is a synthetic measure of the cultural and economic background of students, constructed through multiple component analysis of three main factors related to the student's family: school attainment of both parents, profession of parents, and possession of some goods from the family (Campodifiori et al., 2010). Mother's educational attainment was prioritised over that of the father to account for gender-related disparities in tertiary education (Carneiro et al., 2013). This variable was obtained directly from the INVALSI student questionnaire. As this variable was not collected in 2018/2019 for students in grade 13 of schooling, we matched from the student's questionnaire administered in grade 10 in the year 2015/2016 to the same students. This

procedure led to approximately 20% of missing data, and a similar percentage is the proportion of missing data on ESCS.

To ascertain the relationship between missing data and outcome (Stem/Not Stem choice) for both ESCS and mother educational attainment, the non-significance of the missing category in the regression model was tested. Given that the missing data, arising from the discordance between grades 13 and 10, were randomly distributed, only students with valid answers on mother's education and SES items were retained in the sample.

Concerning schools, the geographical location within the macro region (North-East, North-West, Centre, South and Islands) and the management of the school (public/private) were considered.

The outcome variable for this study, i.e. enrolment in a STEM course at university, was created based on the official classification of university courses into five "Area of discipline": These are Art and Literature, Economics, Law and Social Sciences, Health Studies and STEM. The Ministry of University and Research provides this data in an open data format on their website: https://dati-ustat.mur.gov.it/dataset/metadati/resource/adb4d6bd-0ad9-4004-8e23-b7baa2e45495).

The finale sample was composed of 61,574 students, whose characteristics are reported in Table 1.

Table 1 – Characteristics of the sample

	Number	Percentage
Gender		
Female	27,626	44.9
Male	33,948	55.1
Teacher Marks in Math		
Insufficient	12,602	20.5
Sufficient	16,823	27.3
Good	25,485	41.4
Excellent	6,664	10.8
Level in Math score at INVALSI		
Level 1	1,119	1.8
Level 2	2,820	4.6
Level 3	7,419	12.0
Level 4	12,774	20.7
Level 5	37,442	60.8
ESCS-Classes		
Low	5,619	9.1
Medium-low	11,657	18.9
Medium-high	14,904	24.2
High	29,394	47.7
Mother Education		
Lower Secondary school or lower	7,719	12.5
Professional Qualification	2,910	4.7
High school diploma	24,609	40.0
University degree or higher	26,336	42.8
Migration background		
Native	59,050	95.9
I generation migrant	828	1.3
II generation migrant	1,696	2.8
Regularity		
Regular attendant	59,945	97.4
Late attendant	1,629	2.6

	Number	Percentage
Geographical Area of the school		
North-West	16,526	26.8
North-East	10,931	17.8
Centre	12,778	20.8
South	16,965	27.6
South and Islands	4,374	7.1
Management of the school		
Public	58,592	95.2
Private	2,982	4.8
Total	61574	100.0

Source: our elaboration on linked dataset INVALSI-Ministry of Education – University Register

2.3 Data analysis

Descriptive analyses were initially performed on the entire population to illustrate the relationship between typology of upper secondary school and choice of STEM major at university, according to gender. We then compared the proportion of enrolment in STEM according to typology of high school previously attended (Scientific Lyceum, Other Lyceum, Technical and Professional Schools) distinctly, for male and female students.

Given that previous school track is the primary predictor of university course selection, a secondary focus was directed exclusively on students from scientific lyceum, as they constitute the "eligible population" for pursuing a STEM career at university. This subpopulation was subjected to inferential analyses designed to estimate the major drivers for the choice of STEM at university. Given the expectation of differential mechanisms operating on boys and girls, two logistic models stratified by gender were performed to illustrate the gendered effect that each variable has on STEM university choices for the two groups. In instances of multicollinearity (i.e. socio-economic status and mother's education; teacher's mark and INVALSI score in math), a single variable was selected to represent that dimension, based on the higher parsimony of the model, evaluated by the Akaike Information Criterion (AIC) (Akaike, 1974) and the analysis of VIF values. Eventually, a visual representation was created illustrating the probabilities of selecting STEM courses for boys and girls, according to teachers' evaluations in mathematics. This representation was created by means of the following formula for predicted probabilities:

$$P(Y = 1 \mid X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}$$

where Y is the event “selecting STEM”, and included covariates are gender and teacher’s mark in mathematics.

3. RESULTS

3.1 Preliminary analyses

The type of high school previously attended is undoubtedly one of the primary factors influencing the selection of STEM or non-STEM majors at university. As illustrated in Table 2, scientific lyceums and technical schools represent the previous academic background of over 80% of students who have selected STEM at university (55.8% of whom attended scientific lyceums and 25.5% of whom attended technical schools). The chi-squared test for the association between type of high school and STEM field at university has a significance lower than 0.001.

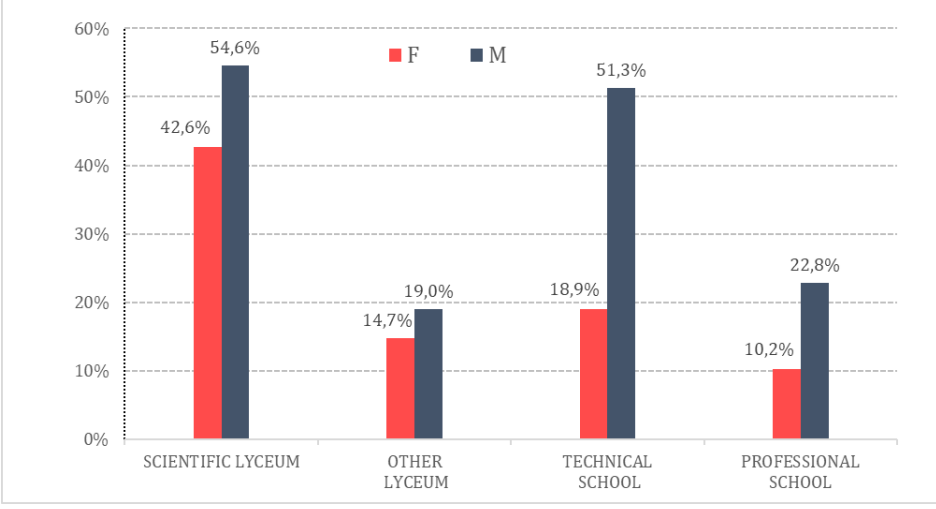
Table 2 – Proportion of STEM choice at university according to previous school track attended – Italy, academic year 2019/2020

	University area of enrolment		
Typology high school	Non Stem	**Stem**	Total
Scientific Lyceum	28.6%	**55.8%**	37.7%
Other Lyceum	45.2%	**16.7%**	35.7%
Technical School	20.7%	**25.5%**	22.3%
Vocational School	5.6%	**1.9%**	4.4%
Total	100.0%	100.0%	100.0%

However, the gender imbalance in STEM choice is evident irrespective of the school track previously attended by the students (Figure 1). The gender disparity in STEM fields persists even when controlled for the inequality in the distribution of students in high-school track by gender (girls more concentrated in humanities, boys in the scientific fields). Notably, the choice of STEM is characterized by a pronounced male predominance among students coming from technical schools,

with a gap of approximately 38 percentage points favouring males (51.3% versus 18.9%). Furthermore, among students enrolled in scientific lyceums, the proportion of male students choosing a STEM course at university significantly exceeds that of their female counterparts by more than 12 percentage points (54.6% versus 42.6%), underscoring a pronounced gender imbalance. For all the subgroups (Scientific Lyceum, Other Lyceum, Technical and Vocational School) the difference in the proportion of STEM by gender is significant with chi squared test with a value lower than 0.001.

Figure 1 – Proportion of males and females choosing STEM at University according to previous school track attended – Italy, academic year 2019/2020



The analyses presented hereinafter are thus developed only on the subpopulation of Italian students coming from Scientific Lyceums, in order to better understand the factors that hinder the propensity of girls in pursuing STEM studies at university, although their possession of scientific academic background.

In a first step of analysis, we focused exclusively on female students at university who attended scientific lyceum, comparing those who selected STEM courses with those who selected non-STEM courses. The analysis encompassed a range of individual characteristics (migration background, regularity during study), family and geographical context (ESCS index, mother educational attainment, geographical area where school is located) and mathematical skills as assessed by teachers and by INVALSI test.

In Table 3 are reported all the variables that exhibited statistical significance at the 0.001 level at the Chi-squared test.
The distribution of girls selecting STEM subjects varies significantly with respect to their mathematical skills, as measured by both teachers' grades and INVALSI test scores. A substantial proportion of girls enrolled in STEM programmes had been assessed as "good/excellent" by their mathematics teachers the previous year (40.8%), compared to only 22.2% of girls in non-STEM programmes. Moreover, nearly two-thirds (63.0%) of "STEM girls" achieved Level 5 (very high competence) on the INVALSI test in grade 13, whereas the corresponding share in the non-STEM group was 48%. Conversely, 10% of girls in the non-STEM group obtained very low/low INVALSI scores, compared to only 4.3% among STEM girls.

With regard to contextual variables, mother's education was the only characteristic showing a significant though numerically modest, difference between the STEM and non-STEM groups, with the expected advantage for girls whose mothers have higher educational attainment. No statistically significant differences emerged for the remaining variables examined.

Table 3 – Percentage distribution of girls from Scientific Lyceum selecting STEM /NON STEM according to the most relevant variables discriminating between the 2 groups

		NOT STEM	STEM
MATH TEACHER MARK	Insufficient	22.0%	10.40%
	Sufficient	29.8%	22.40%
	Good	26.0%	26.40%
	Excellent	22.2%	40.80%
Total		*100.0%*	*100.0%*
MATH INVALSI SCORE	1 - Very Low	2.4%	0.90%
	2 – Low	7.0%	3.40%
	3 – Sufficient	16.8%	11.60%
	4 – High	25.7%	21.00%
	5 - Very High	48.1%	63.00%
Total		*100.0%*	*100.0%*
MOTHER EDUCATIONAL ATTAINMENT	Lower secondary school or lower	14.9%	12.50%
	Professional Qualification	5.4%	4.50%
	High School Diploma	40.7%	42.50%
	University Degree or higher	39.0%	40.50%
Total		*100.0%*	*100.0%*

In this second research line, an inferential perspective was adopted to estimate the effects of all the characteristics previously described on the propensity to select a STEM career, *ceteribus paris*. The objective of this study was to estimate the relative weight of mathematical ability—hypothesised to exert a legitimate and theoretically grounded influence on the selection of academic orientation—in comparison to the impact of social factors, which are posited to attenuate the

propensity of female students to pursue university trajectories in science and technology disciplines.

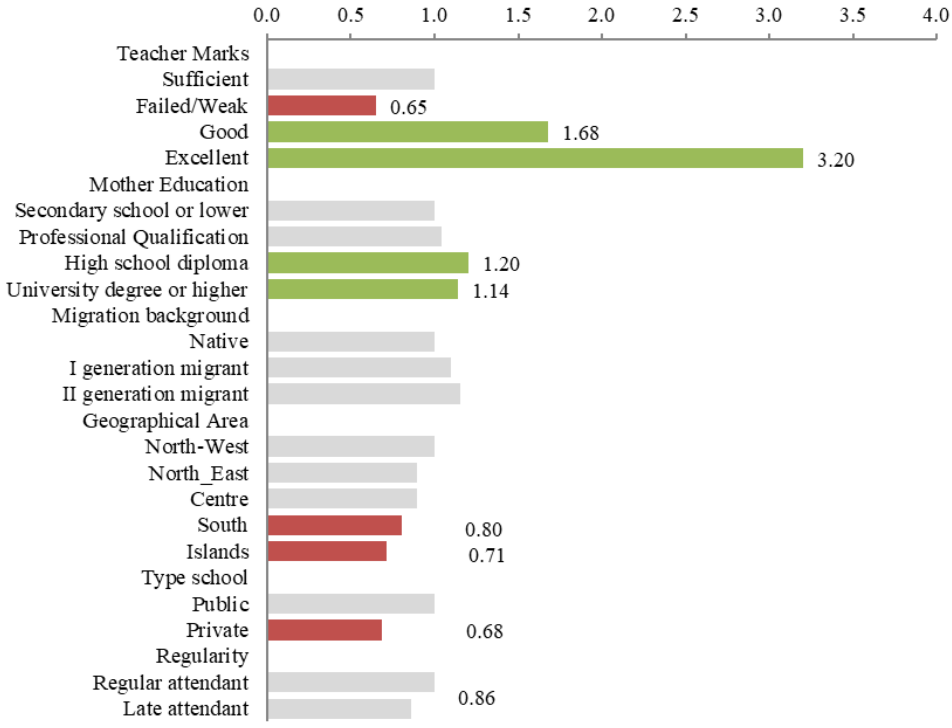
To this end, we have implemented two logistic regression models, stratified by gender. This approach enabled the estimation of the predictive power of distinct characteristics separately for boys and for girls, facilitating the comparison of similar and different patterns of effects in the two groups.

In selecting the variables to be included in the model, it was necessary to choose between pairs of variables capturing the same underlying latent constructs, namely:

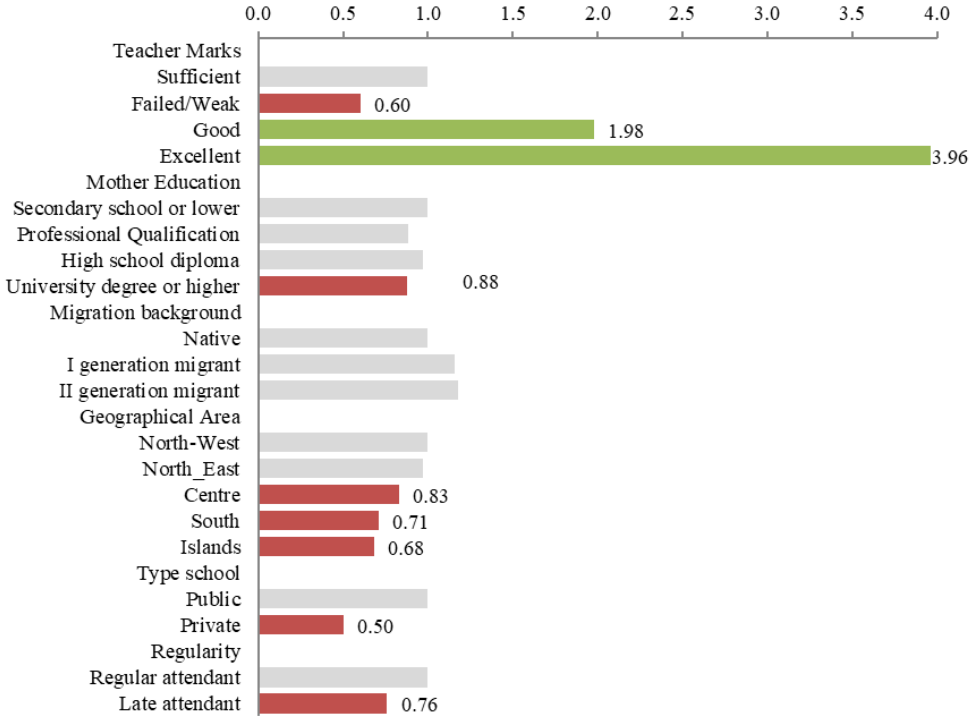
- ESCS and mother's educational attainment, both serving as proxies for the family's socio-cultural background;
- Teacher-assigned grades and INVALSI test scores, both reflecting students' mathematical aptitude.
-

The selection was made based on the Akaike Information Criterion (AIC) in the comparison models with the alternative variables.

For socio-economic factors, the model with the mother's education resulted more parsimonious than the one with ESCS (AIC 18979.7 versus 18995.8); with regard to math competencies, the teacher marks resulted as a better predictor for STEM choice than standardized test scores (AIC 18255.7 versus 18670.5).


0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Teacher Marks
Sufficient
Failed/Weak
0.65
Good
1.68
Excellent
3.20
Mother Education
Secondary school or lower
Professional Qualification
High school diploma
1.20
University degree or higher
1.14
Migration background
Native
I generation migrant
II generation migrant
Geographical Area
North-West
North_East
Centre
South
0.80
Islands
0.71
Type school
Public
Private
0.68
Regularity
Regular attendant
0.86
Late attendant

(a)



(b)

Figures 2a and 2b –Odds Ratio from Logistic Models predicting choice of STEM at university for girls (a) and boys (b)

In Figure 2 the results from the logistic models are presented for both girls (a) and boys (b). The findings indicate that mathematical skills are the strongest predictor of choosing a STEM degree at university, regardless of gender. For both boys and girls, strong performance in mathematics increases the likelihood of selecting a STEM course by more than threefold. This effect is slightly stronger for boys: their probability of enrolling in a STEM program nearly quadruples when they receive high mathematics evaluations from their teachers, whereas for girls the probability increases by approximately 3.2 times following an excellent mathematics assessment.

Mother's educational attainment appears to exert an opposite effect on girls and boys. Specifically, girls whose mothers hold a university or higher degree show approximately 20% higher odds of enrolling in a STEM degree compared to those whose mothers completed only upper secondary education (OR 1.17 and 1.20). Conversely, for boys, a higher maternal education is associated with lower odds of selecting a STEM course (OR 0.88 for boys whose mothers had a university degree). The geographical location of the student's high school, which is a good proxy of geographical region of life, exhibits a comparable pattern for both genders: Students from northern regions display a higher propensity to pursue STEM fields compared to their peers from central and southern regions.

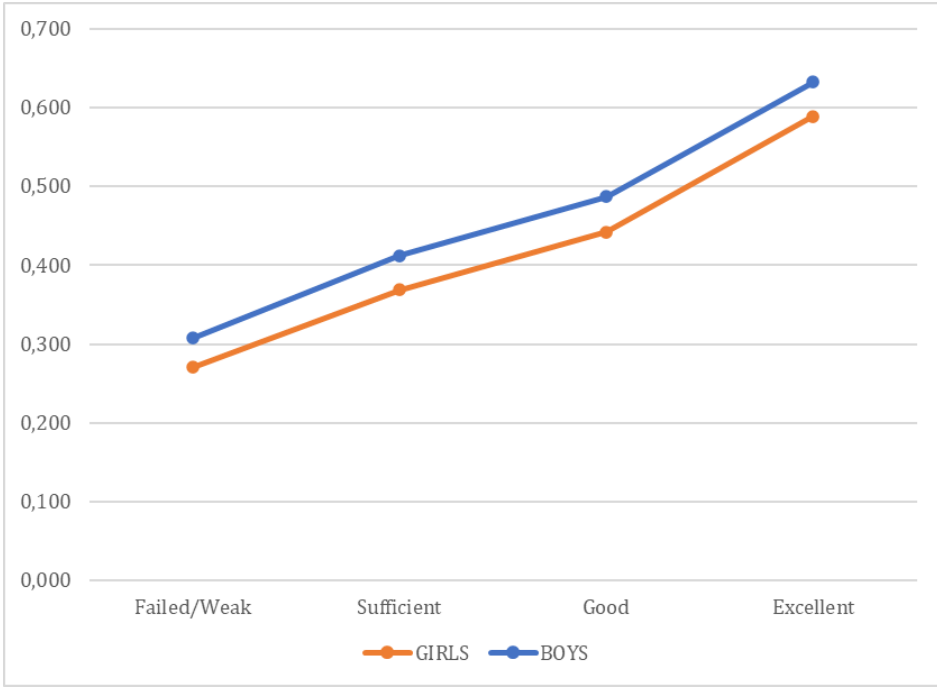


Figure 3–Predictive probability of choosing STEM according to teachers' evaluation in Math by gender -Italy academic year 2019/2020

A visual result that better illustrates the effects shown in the logistic models is the graph of predicted probabilities of enrol in STEM courses at university according to teacher grades received in high school by gender (Figure 3). The probability of choosing STEM majors at university increases steeply as maths grades improve (students with very high mathematics performance, i.e. excellent in the final year of high school have an estimated probability of 0.6 of enrolling in a STEM-university degree, compared to 0.3 for students who have been evaluated as weak in this discipline). A slightly lower, still parallel, trend is observed for girls.

5. DISCUSSION

Although multidisciplinary research on gender and STEM is extensive and ongoing, many aspects of this phenomenon are still unknown, and we have only a partial understanding of the mechanisms that prevent horizontal gender segregation.

A further complexity arises from the fact that the school system and the university system in Italy constitute two markedly distinct institutional contexts. When a student enters the school system, he or she is assigned a number that identifies him or her throughout his or her school career. At the end of the regular school cycle and upon entering the university world, the student receives a new university identification number, which is different from the one used until then. This procedure means that data from the school system and the university system cannot "talk" to each other, as they belong to two different databases. It also excludes the use of the personal identification number, which would make it possible to identify the student, and this would not be allowed under data protection regulations.

The peculiarity of this study lies in the fact that it has been possible to follow students outside the school system and to link them to their choice of university and career.

It is therefore an innovative study that opens the possibility of various more in-depth studies that will also be useful in identifying indicators of predictability of failure in the university world, indicators of inequalities between groups of male

and female students, etc., which can then be used at the policy level to intervene promptly in the areas/places where these phenomena are reproduced.

We evaluated the decision-making process from the last year of secondary school to university from a gender perspective, paying particular attention to the most relevant aspects that positively influence the choice of a STEM subject at university. The main finding of the study is that mathematics achievement is a highly influential factor in the choice of STEM subjects, but that the ability as assessed by teachers is much more predictive than the scores obtained in a standardised test designed to measure the same ability.
This effect may be due to two different mechanisms: firstly, the teacher's mark may capture a more multidimensional aspect of the student; secondly, the teacher's mark represents not only a strict evaluation of the student's ability, but also a generalised opinion about the person and his/her potential, which shapes his/her self-esteem and ambitions.
Another important finding is the effect of the mother's educational level on boys and girls in terms of encouraging the choice of STEM studies. For girls, positive effects were found, with a higher propensity to choose STEM if they come from a highly educated family (in terms of mother's title), presumably indicating a weakening of social stereotypes where there are more cultural elements. For boys, on the other hand, the model shows the opposite: those who have a highly educated mother are less likely to end up in a STEM faculty.
This result, that can appear to be contradictory, may represent two side effects of the same phenomenon: the intense social conditioning based on stereotypes, which is relaxed when the cultural level of the family is higher than the norms, allows girls to move towards STEM and boys to leave STEM studies in favour of non-STEM paths.
These findings lend support to the hypothesis that gender stereotypes exert a bidirectional influence, constraining male students' freedom to select non-STEM fields (such as the humanities, social work, or education) in higher education, just as they restrict female students' engagement with STEM disciplines for tertiary education.
Our findings reveal notable gender differences in the factors influencing the choice to enrol in STEM disciplines at university. While teacher-assigned grades

play an important role for both male and female students, they appear to have a stronger impact on male students' decisions. Moreover, maternal education, used as a proxy for cultural capital, shows contrasting effects: higher maternal education is associated with a greater likelihood of enrolling in humanities for males, but with an increased likelihood of choosing STEM fields for females. In addition, we observe significant geographical disparities: female students from Southern regions show a lower probability of enrolling in STEM disciplines, while for male students this lower likelihood extends to both Central and Southern regions. These results partially align with Priulla and Attanasio (2023), who emphasized the complex interplay between gender and academic background in shaping university pathways, highlighting how social and cultural resources differently affect male and female students' educational choices. Similarly, Priulla et al. (2022), using a machine learning approach, demonstrated that high school performance indicators are significant predictors of university enrolment choices, though without explicitly addressing the gender-specific dynamics of teacher-assigned grades or territorial variations that we observe. Finally, Barone and Assirelli (2020) directly investigated gender segregation in higher education, testing multiple explanatory mechanisms including the role of cultural capital, social expectations, and regional contexts. Their study acknowledged the importance of geographical inequalities, but our findings refine this understanding by showing how regional effects intersect with gender, particularly disadvantaging female students from the South in accessing STEM fields. Our results thus add a novel perspective to the understanding of gendered and territorial pathways in higher education.

Furthermore an innovative aspects is given by the fact that findings come from a population study that jointly considers different sets of variables: geographical, economic and cultural background, individual characteristics and mathematics ability, to give an integrated picture of the mechanisms that influence university choice. An important limitation, however, is that we have only looked at enrolment, with no information on subsequent academic careers. For the success of the academic career, the results of the INVALSI test could play a much more important role than the teacher's assessment. In fact, the latter may be influenced by gender stereotypes and risk directing students towards disciplines not based on their best qualities, but on social expectations linked to gender.

More boys enrol in STEM subjects, but it is still unclear whether they complete their studies as quickly and successfully as their female counterparts. Girls, who are more highly selected at entry, may achieve more brilliant results in less time. Future lines of research should look in these directions, focusing on post-enrolment careers and estimating factors for higher 'success' in STEM from a gender perspective.

REFERENCES

Akaike, H. (1974). "A new look at the statistical model identification". *IEEE Transactions on Automatic Control*, 19(6), 716–723.

Allegrini, A., Pellegrini, G., and Segafredo, C. (2015). Italian Female and Male Students' Choices: STEM Studies and Motivations. In: Henriksen, E., Dillon, J., Ryder, J. (eds) Understanding Student Participation and Choice in Science and Technology Education. Springer, Dordrecht.

Barone, C., and Assirelli, G. (2020) Gender segregation in higher education: an empirical test of seven explanations. *Higher Education: The International Journal of Higher Education Research* 79, 55–78.

Borgonovi, F., Choi A. and Paccagnella M. (2018). The evolution of gender gaps in numeracy and literacy between childhood and adulthood. *OECD Education Working Papers*, 184, OECD Publishing, Paris.

Campodifiori E., Figura E., Papini M. and Ricci R. (2010) Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia. *INVALSI Working Paper*, n. 2.

Carneiro, P., Meghir, C., & Parey, M. (2013). "Maternal Education, Home Environments, and the Development of Children and Adolescents." Journal of the European Economic Association, 11(s1), 123-160.

Card, D. and Payne, A.A. (2021), High School Choices And The Gender Gap In Stem. *Econ Inq*, 59: 9-28.

Contini, D., Di Tommaso, M.L., Mendolia, S. (2017). The Gender Gap in Mathematics Achievement: Evidence from Italian Data. *Economic Education Review* 58: 32-42.

Desimoni, M. (2018). I livelli per la descrizione degli esiti delle Prove INVALSI. Roma: INVALSI.

Eccles, J. S. (2007). Where are all the women? Gender differences in participation in physical science and engineering. In S. J. Ceci & W. M. Williams (Eds.), Why aren't more women in science? Top researchers debate the evidence (pp. 199-210). American Psychological Association.

Eccles, J. S., and Wang, M. T. (2016). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development*, 40(2), 100-106.

Ellison, G. and Swanson, A. (2010). The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions. *Journal of Economic Perspectives*, 24 (2): 109-28.

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bullettin* 136, 103–127.

European Institute for Gender Equality (EIGE). (2021). Browse Gender Statistics | Gender Statistics Database. European Institute for Gender Equality

Hill, C., Corbett, C., & St. Rose, A. (2010). Why So Few? Women in Science, Technology, Engineering, and Mathematics. American Association of University Women.

Ismu. (2020). *Alunni con background migratorio in Italia. Le opportunità oltre gli ostacoli*. Fondazione ISMU.

Makarova E, Aeschlimann B and Herzog W (2019) The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations. *Frontiers in Education*. 4:60.

Master, A., and Meltzoff, A. N. (2020). Cultural Stereotypes and Sense of Belonging Contribute to Gender Gaps in STEM. *International Journal of Gender, Science and Technology*, 12(1), 152–198.

Osborne, J., Shirley, S., and Collins, S. (2003). Attitudes towards Science: A Review of the Literature and its Implications. *International Journal of Science Education*, 25(9): 1049-1079.

OECD (2015), *Immigrant Students at School: Easing the Journey towards Integration*, OECD Reviews of Migrant Education, OECD Publishing, Paris.

Priulla, A., Albano, A., D'Angelo, N., & Attanasio, M. (2024). Investigating the association between high school outcomes and university enrolment choices: A machine learning approach. In *Proceedings of the Statistics and Data Science 2024 Conference*

Priulla, A., and Attanasio, M. (2023). Unveiling gender disparities in university pathways: Insights from Italy's master's level. *European Journal of Higher Education*, 13(4), 345–360

Turner, S.E. and Bowen W.G. (1999). Choice of major: The changing (unchanging) gender gap. *Industrial and Labor Relations Review*, 52(2): 289–313.