Gender Disparity in STEM Enrolment Across Italian Provinces: A Multilevel Analysis

Giuliana La Mantia · Giovanni Boscaino

University of Palermo, Palermo, Italy

`{giuliana.lamantia, giovanni.boscaino}@unipa.it`

Abstract

This paper examines the gender gap in STEM enrolment in first-level degree courses at Italian universities, with a focus on the role played by students' provinces of origin. STEM enrolment is modelled as a function of both individual and territorial characteristics. A multilevel logistic regression model is employed, where students (Level 1) are nested within provinces (Level 2). Random intercepts capture provincial variation in the baseline propensity to enrol in STEM, while random slopes for gender allow the gender effect to vary across provinces. Our findings indicate that provincial differences in STEM enrolment and the gender gap are modest but remain statistically significant even after controlling for individual and territorial characteristics. Provinces with higher female labour market participation and higher female university enrolment rates are associated with a greater likelihood of STEM enrolment, whereas overall economic prosperity does not necessarily translate into higher STEM participation.

Keywords: gender gap · STEM enrolment · multilevel analysis · Italian provinces · higher education

1. Introduction and Motivation

Gender disparities in STEM (Science, Technology, Engineering, and Mathematics) education and careers remain a persistent and well-documented issue in academic research and public policy. Despite the steady increase in women's participation in higher education over recent decades, female students continue to be substantially underrepresented in STEM fields globally. According to OECD (2022), women accounted for only 31% of new enrolments in STEM bachelor's programmes across OECD countries in 2020, while representing 79% of new entrants in health and welfare disciplines. Although the number of female STEM graduates in the European Union has grown since the mid-2000s, significant disparities persist, particularly in fields such as mathematics, engineering, and computer science.

The drivers of this gender gap are multifaceted and operate at different levels. At the individual level, studies have linked the gap to academic preferences, learning orientations, self-perceived mathematical ability, and external influences including teacher expectations, peer effects, and family background. Maltese and Cooper (2017) found that males tend to develop an interest in STEM independently, whereas females are more influenced by teacher encouragement, peer support, and prior academic performance. Even among women who do pursue STEM degrees, field-specific disparities remain evident: female students disproportionately gravitate towards biological and natural sciences, while men continue to dominate in technical fields such as engineering, computer science, and physics.

While the individual-level determinants of the gender gap in STEM have been extensively investigated, considerably less attention has been devoted to understanding how territorial and economic conditions shape STEM enrolment patterns. The Italian context is particularly instructive in this regard, given the country's pronounced North–South divide. Southern Italy faces persistently higher unemployment rates, lower female labour market participation, and greater poverty levels relative to northern regions. These structural disparities create an uneven landscape within which individual educational decisions are made.

Contini et al. (2018) examined the impact of provincial economic conditions on higher education outcomes in Italy, finding that unemployment rates exhibit an inverted U-shaped relationship with enrolment, a negative association with timely degree completion, and a positive correlation with dropout rates. More specifically, higher youth unemployment — particularly concentrated in the South — discourages university enrolment, increases dropout rates, and delays degree attainment. At the same time, Galvin et al. (2024) demonstrated, in the United States context, that gender disparities in STEM vary meaningfully across geographic regions, with institutional initiatives and local economic conditions playing a significant moderating role.

The present study integrates these two strands of research — territorial disparities in university enrolment and gender gaps in STEM education — to investigate whether provincial characteristics, and particularly those related to women's education and employment, influence female students' likelihood of enrolling in STEM degree programmes. This is achieved through a multilevel logistic regression framework that simultaneously accounts for individual student characteristics and province-level contextual factors.

The remainder of the paper is structured as follows: Section 2 describes the data sources and the methodological approach; Section 3 presents and discusses the empirical results; Section 4 concludes with a summary of the main findings and an outline of future research directions.

2. Data and Methodology

2.1 Data Sources

The empirical analysis draws on two main data sources. At the student level, data were obtained from the MOBYSU.IT micro-data provided by the Italian Ministry of University and Research (MUR). This database contains detailed information on university students' sociodemographic characteristics, academic careers, and educational backgrounds in Italy, covering the period from 2008 to 2022. For this study, we use data from the 2021/2022 cohort of first-year university students who graduated from Italian high schools in 2021. Students enrolled in online universities, single-cycle degree programmes, and health and sanitary fields were excluded from the analysis, as were those lacking information on their province of residence. The final sample consists of **87,914 first-year university students**.

At the provincial level, contextual indicators were drawn from the Benessere Equo e Sostenibile (BES) database compiled by the Italian National Institute of Statistics (ISTAT) for the year 2021. The BES framework provides a rich set of indicators capturing multiple dimensions of well-being and sustainable development at the sub-national level. Due to missing data, four provinces — Monza and Brianza, Fermo, Barletta-Andria-Trani, and Sud-Sardegna — could not be included in the analysis, resulting in a final set of **104 provinces**.

2.2 Variables

The binary outcome variable, **STEM choice**, takes the value 1 if a student enrols in a STEM degree course and 0 otherwise.

Student-level covariates include:

- **Gender** (female as baseline category)
- **Type of high school attended** (classical, scientific, humanities and social sciences, technical-commercial, technical-professional, vocational, other)
- **Mathematics and Italian scores** in the final year of high school (Maturità examination)
- **High school final grade** (overall grade at graduation)
- **Socioeconomic status (SES)** of the student's family
- **Mother's education level** (high school as baseline; degree/postgraduate; middle school or lower; professional qualification)
- **Mother's occupation** (employee as baseline; manager; self-employed; worker; unemployed)

Province-level covariates include, from the Education and Training dimension:

- Percentage of female students with insufficient numerical competence (Grade III of secondary school)
- Overall percentage of students with insufficient numerical competence
- Rate of female students enrolling in university

From the Wealth and Employment dimension:

- Per capita income (as a proxy for the overall socioeconomic context)
- Female employment rate
- Total employment rate
- NEET rate (young people aged 15–29 not in education, employment, or training)

2.3 Multilevel Logistic Regression Model

To examine the variation in the gender gap across provinces while accounting for the hierarchical structure of the data — students nested within provinces — we employ a **Multilevel Logistic Regression Model (MLM)**. The model explicitly treats the 104 provinces as Level 2 units and the 87,914 students as Level 1 units.

The model is specified with two sources of random variation. First, **random intercepts** are introduced to capture provincial differences in the baseline propensity for STEM enrolment, allowing the overall likelihood of STEM enrolment to vary across provinces independently of the observed covariates. Second, **random slopes for gender** are included to allow the gender effect on STEM enrolment to differ across provinces, thereby modelling heterogeneity in the gender gap itself.

Formally, since female students serve as the baseline category for gender, the random slope coefficient associated with being male indicates how much more (or less) likely male students are to enrol in STEM relative to females in each province. A positive random slope identifies provinces where the propensity for males to enrol in STEM is comparatively higher; a negative random slope indicates provinces where females are relatively more likely to enrol in STEM. After including the covariates, the residual variances of the random intercept and slope capture the unexplained provincial variation in STEM enrolment and in the gender gap, respectively.

3. Results

3.1 Fixed Effects: Student-Level Factors

Table 1 reports the coefficient estimates from the multilevel logistic regression model. The results confirm and extend several well-established findings from the literature on STEM enrolment.

Table 1. Regression results — response variable: propensity to enrol in STEM.

Predictor	Log-Odds	p-value
(Intercept)	−2.00	< 0.001
Gender: Female	Reference	—
Gender: Male	0.42	< 0.001
Student SES	0.01	0.370
Math score (grade 13)	0.82	< 0.001
Italian score (grade 13)	−0.27	< 0.001
Gender Male × SES	−0.18	< 0.001
Gender Male × Math Score	0.07	0.003
Gender Male × Italian Score	−0.08	< 0.001
High school final grade	0.18	< 0.001
High School: Classical	Reference	—
Scientific	0.86	< 0.001
Human and Social Sciences	−0.49	< 0.001
Technical-Commercial	−0.89	< 0.001

Predictor	Log-Odds	p-value
Technical-Professional	1.03	< 0.001
Vocational	0.30	< 0.001
Other	0.18	0.004
Mother's qualification: High school	Reference	—
Degree/Postgrad	0.06	0.003
Middle School or Lower	–0.04	0.143
Professional Qualification	–0.01	0.714
Mother's occupation: Employee	Reference	—
Manager	–0.06	0.018
Self-Employed	–0.14	< 0.001
Worker	0.05	0.076
Unemployed	–0.01	0.627

As expected, **male students are significantly more likely to enrol in STEM** than their female counterparts (log-odds = 0.42, $p < 0.001$). **Higher mathematics scores** strongly and positively predict STEM enrolment (log-odds = 0.82), while **higher Italian scores** are negatively associated with STEM choice (log-odds = –0.27). These effects are further modulated by gender: as mathematics scores increase, the advantage for male students in STEM enrolment grows; conversely, male students with high Italian scores are even less likely to enrol in STEM than females with equivalent scores.

Regarding **high school type**, students from scientific high schools show the highest propensity for STEM enrolment relative to those from classical high schools, while students from technical-commercial institutes and humanities and social science programmes are significantly less likely to enrol in STEM. Interestingly, students from technical-professional institutes show the largest positive effect, exceeding even scientific high school graduates, reflecting the vocational preparation specific to engineering and applied fields.

Socioeconomic status does not significantly influence the probability of STEM enrolment for female students. However, the interaction between SES and gender reveals a significant and negative effect: as SES increases, the probability of STEM enrolment actually *decreases* for male students. This counterintuitive result may reflect the broader range of educational opportunities available to students from higher-SES families, reducing the relative attractiveness of STEM pathways.

Maternal background emerges as a meaningful predictor. Students with self-employed mothers are less likely to enrol in STEM than those whose mothers are employees, while students with working-class mothers show a slightly higher propensity for STEM enrolment. Students whose mothers hold a university or postgraduate degree are more likely to choose STEM than those whose mothers have only a high school diploma, suggesting an intergenerational transmission of educational aspirations.

3.2 Fixed Effects: Province-Level Factors

At the provincial level, the results reveal important contextual influences on STEM enrolment patterns.

Table 2. Provincial-level predictors.

Predictor	Log-Odds	p-value
Female University Enrolment Rate	0.01	< 0.001
Female students with insufficient numerical competence (INC)	0.34	0.039
Total students with INC	−0.25	0.137
Female Employment Rate	0.23	0.035
Total Employment Rate	−0.33	0.005
NEET rate	−0.08	0.025
Per Capita Income	−0.10	0.005

A higher **female university enrolment rate** at the provincial level is positively associated with the probability of STEM enrolment, suggesting that environments where women are more actively engaged in higher education overall create more favourable conditions for female participation in STEM specifically. Similarly, a higher **female employment rate** is associated with a greater likelihood of STEM enrolment, indicating that the visibility of women in the labour market may positively shape aspirations and perceived opportunities for young women considering technical fields.

Conversely, higher **overall employment rates** and higher **per capita income** at the provincial level are negatively associated with STEM enrolment. This result is counterintuitive from the perspective of human capital theory and suggests a more complex mechanism: in wealthier provinces, the opportunity cost of investing in demanding STEM programmes may be perceived as higher relative to the broader and immediately accessible labour market. Additionally, provinces with higher **NEET rates** show a lower propensity for STEM enrolment, reflecting the general discouraging effect of precarious economic conditions on higher education investment.

3.3 Random Effects and Provincial Heterogeneity

The random effects results provide critical insight into the geographic dimension of the gender gap in STEM.

Random Effect Parameter	Estimate
Residual variance σ^2	3.25
Provincial variance (intercept) τ_{00}	0.03
Provincial variance (slope for male) τ_{11}	0.05
Intercept–slope correlation ρ_{01}	−0.79

Random Effect Parameter	Estimate
Number of provinces	104
Observations	87,914
Marginal R^2 / Conditional R^2	0.331 / 0.336

Although the estimated provincial variances are small in absolute terms — indicating that individual-level characteristics explain the lion's share of variation in STEM enrolment — they remain **statistically significant**, confirming that the province of residence exerts a genuine independent effect on STEM enrolment propensity and on the gender gap.

Strikingly, the estimated **correlation between the random intercept and the random slope for gender is –0.79.** This strong negative correlation implies that provinces where the overall propensity for STEM enrolment is higher also tend to exhibit a *smaller* gender gap, while provinces with lower overall STEM enrolment show a wider gender gap. In other words, policies that raise overall STEM participation may simultaneously help narrow the gender divide.

The posterior predictions of the random effects allow us to rank provinces according to both the overall propensity for STEM enrolment and the strength of the gender gap. The provinces with the most pronounced gender gaps in favour of males are **Bergamo** (random slope estimate: 0.483), **Vicenza** (0.267), **Verona** (0.260), and **Treviso** (0.247) — all located in Northern Italy's highly industrialised regions. By contrast, the provinces with the smallest — or even reversed — gender gaps include **Potenza** (–0.457), **Cagliari** (–0.306), **Lecce** (–0.235), **Naples** (–0.182), and **Rome** (–0.178), predominantly in southern or central Italy. This geographic pattern is consistent with the broader labour market structure of these regions and with the association between female employment rates and STEM enrolment documented in the fixed effects analysis.

4. Conclusion and Future Developments

This study provides new empirical evidence on the determinants of gender disparities in STEM enrolment in Italian universities, combining individual-level and province-level factors within a unified multilevel modelling framework. The results confirm well-established individual-level predictors — including gender, mathematics scores, high school type, and maternal background — while also revealing the significant moderating role of territorial context.

Most notably, the analysis demonstrates that **higher female labour market participation and higher female university enrolment rates at the provincial level are associated with a greater likelihood of STEM enrolment**, suggesting that structural features of the local environment shape individual educational trajectories beyond what individual characteristics alone can explain. Conversely, general economic prosperity, as captured by per capita income and total employment rates, does not translate into higher STEM participation, pointing to the need for targeted interventions rather than reliance on broad economic development.

The strong negative correlation between the random intercept and the random slope for gender (–0.79) is a particularly policy-relevant finding: it suggests that raising overall STEM participation at the provincial level is likely to also reduce the gender gap in that province, creating a positive dual return on investment in STEM promotion policies.

The provincial rankings produced by the model provide actionable intelligence for policy design. Provinces in Northern Italy, despite their higher general prosperity, exhibit the largest gender gaps in STEM, while some southern provinces show comparatively smaller or even reversed gaps. This finding challenges simple narratives linking economic development to gender equity in education and calls for more nuanced, geographically targeted policy responses.

Future developments of this analysis will include the examination of multiple student cohorts to assess whether the documented patterns are stable over time or reflect transitional phenomena. The analysis will also explore potential spillover effects across contiguous geographic areas, whereby the educational and labour market conditions of neighbouring provinces may influence enrolment decisions. Further extensions may consider the role of specific university characteristics — including the availability and quality of STEM programmes — as additional Level 2 predictors.

Acknowledgements

Financial support is gratefully acknowledged under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU – Project Title: *STEM in Higher Education & Women INequalitieS* [SHE WINS], CUP I53D23004810006, Grant Assignment Decree No. 1060 adopted on 07/17/2023.

Data Availability Statement. Database MOBYSU.IT (Mobilità degli Studi Universitari in Italia), research protocol MUR–Universities of Cagliari, Palermo, Siena, Torino, Sassari, Firenze, Cattolica and Napoli Federico II, P.I. Massimo Attanasio (UNIPA), Data Source ANS-MUR/CINECA.

References

1. OECD (2022). *Education at a Glance 2022*. OECD Publishing, Paris.
2. Caprile, M., Palmén, R., Sanz, P., Dente, G. (2015). Encouraging STEM studies: labour market situation and comparison of practices targeted at young people in different member states. *Policy Department A*, 12.
3. Crosnoe, R., Riegle-Crumb, C., Field, S., Frank, K., Muller, C. (2008). Peer group contexts of girls' and boys' academic experiences. *Child Development*, 79(1), 139–155.
4. Wang, M.T., Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy–value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, 33(4), 304–340.
5. Weeden, K.A., Gelbgiser, D., Morgan, S.L. (2020). Pipeline dreams: occupational plans and gender differences in STEM major persistence and completion. *Sociology of Education*, 93(4), 297–314.

6. Maltese, A.V., Cooper, C.S. (2017). STEM pathways: do men and women differ in why they enter and exit? *AERA Open*, 3(3), 2332858417727276.
7. Barone, C., Assirelli, G. (2020). Gender segregation in higher education: an empirical test of seven explanations. *Higher Education*, 79(1), 55–78.
8. Enea, M., Attanasio, M. (2020). Gender differences in Italian STEM degree courses: a discrete-time competing-risks model. In *Book of Short Papers – SIS 2020*, pp. 385–390. Pearson.
9. Galvin, D.J., Anderson, S.C., Marolf, C.J., Schneider, N.G., Liebl, A.L. (2024). Comparative analysis of gender disparity in academic positions based on US region and STEM discipline. *PLOS ONE*, 19(3), e0298736.
10. Istituto Nazionale di Statistica (ISTAT). (2024). *Rapporto annuale 2024. La situazione del Paese*.
11. Contini, D., Cugnata, F., Scagni, A. (2018). Social selection in higher education: enrolment, dropout and timely degree attainment in Italy. *Higher Education*, 75(5), 785–808.
12. D'Agostino, A., Ghellini, G., Longobardi, S. (2016). University mobility at enrolment: geographical disparities in Italy. In *Proceedings of the 48th Scientific Meeting of the Italian Statistical Society*.
13. Istituto Nazionale di Statistica (ISTAT). (2021). *Benessere equo e sostenibile (BES)*. https://www.istat.it/en/statistical-themes/focus/well-being-and-sustainability/
14. Goldstein, H. (2011). *Multilevel Statistical Models*. John Wiley & Sons, Hoboken.
15. Sulis, I., Barbieri, B., Salaris, L., Melis, G., Porcu, M. (2024). Gender bias in university student mobility: a cohort analysis in Italy. *International Journal of Manpower*, 45(5), 1076–1092.
16. Leckie, G. (2013). Three-level multilevel models – concepts. *LEMMA VLE Module 11*, 1–41. http://www.bristol.ac.uk/cmm/learning/course.html