



The Gender Grading Gap in Italy

Giovanni Abbiati¹, Carlos J. Gil-Hernández², Ilaria Lievore³

¹University of Brescia, ²University of Florence, ³Fondazione per la Scuola

X SEMINAR

“DATA FROM AND FOR THE EDUCATIONAL SYSTEM: TOOLS FOR RESEARCH AND TEACHING”

ROME, 19 – 20 – 21 NOVEMBER 2025

Motivation & Puzzle



- **Gender Gap Reversal** in educational achievement ([DiPrete and Buchmann, 2013](#))
 - o Still, horizontal inequalities (STEM fields: advice + choice) and labour market gender penalties
 - o Girls underperform (outperform) boys in math (language) test scores
- **Gender Grading Gap:** Net of test scores (objective ability), girls get higher teacher grades
 - o Girls' achievement advantage: GPA > test scores ([Lehti and Laaninen, 2024](#))
 - o "Boys problem": drop-out, (mis)behaviour ([Triventi, Lievore, and Dian, 2023](#)), anti-school (masculinity) norms ([Reeves, 2024](#))
 - o Teacher gender stereotypes (bias; discrimination) ([Carlana, 2019](#)) and classroom-specific grading criteria (grading on a curve)
- **Fair Assessments** for equal educational opportunity

Research Questions

- 1.** Is there a Gender Grading Gap (GGG) favoring girls within classrooms?
- 2.** Does the GGG differ between subjects (language vs math)?
- 3.** Does the GGG change across educational/developmental stages?
- 4.** Do classroom characteristics moderate the GGG?

Empirical Expectations & Mechanisms

1. Overall: *Positive GGG (girls advantage), net of school/classroom characteristics*

GGG Heterogeneity By:

2. Subject: *Italian > Math*

- Female (dis) advantage in Italian (math) test scores
- Gender subject-ability stereotypes (e.g., better in language vs. math)

3. Grade (2-8) / Development (Age 7-14): *lower-secondary GGG > primary GGG*

- Increasing grading variation
- Adolescence / puberty: Boys misbehaviour

4. Classroom Characteristics ([Lievore and Triventi, 2022](#)):

- 4.1. *More girls, smaller GGG:* Less conflictive behaviour / peer pressure for masculinity norms
- 4.2. *Larger gender test score gap, smaller GGG:* Compensatory grading ([Schuessler and Sønderskov, 2024](#))

Data



- **Administrative Population Data:** *Italian National Institute for the Evaluation of the Education System*
- **N** $\approx$ 450,000 classrooms (unit of analysis) / $\approx$ 9 million students
 - o 2012/2013 – 2018/2019 (Grades 2, 5): No numeric grading from 2019/2020
 - o 2012/2013 – 2023/2024 (Grade 8)
- **Standardized Test Scores** (low-stakes): Nationwide curriculum-based (mean = 200, SD = 40)
 - o 2nd Semester – May
 - o Low-Stakes Tests: Girls' Higher Persistence $\rightarrow$ Smaller Math Gaps; High-Stakes Tests: Girls' Higher Test Anxiety $\rightarrow$ Larger Math Gaps
- **Teacher Grades** (1-10 scale; mean $\approx$ 7/8; SD $\approx$ 1/1.5)
 - o 1st Semester - January
- **Subjects:** Math, Italian

Methods & Variables

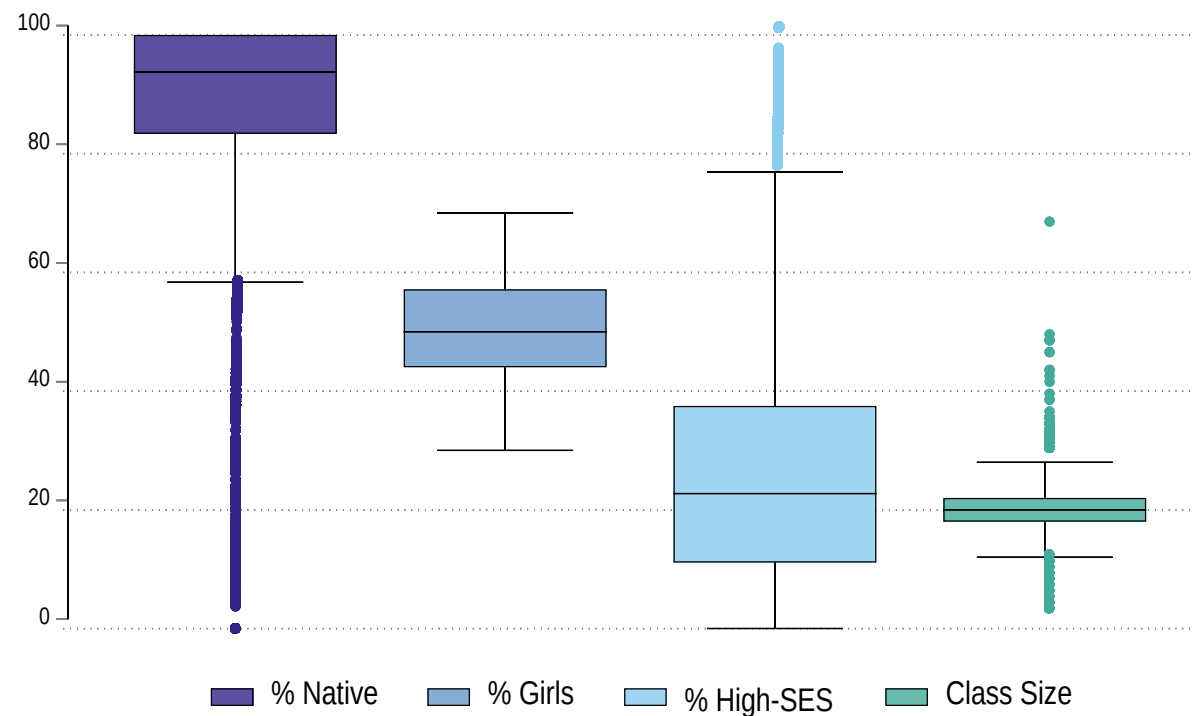
$$\text{Grade}_{ict}^{\text{subject}} = \alpha_c + B^{\text{subject}} \cdot \text{Girl}_i + \gamma \cdot \text{TestScore}_{ict}^{\text{subject}} + \epsilon_{ict}, \quad \text{subject} \in \{\text{Italian, Math}\}$$

$$\text{GGG}_{c,g}^{\text{SD, subject}} = \frac{B_c^{\text{subject}}}{\text{SD}(\text{Grade}_g^{\text{subject}})}$$

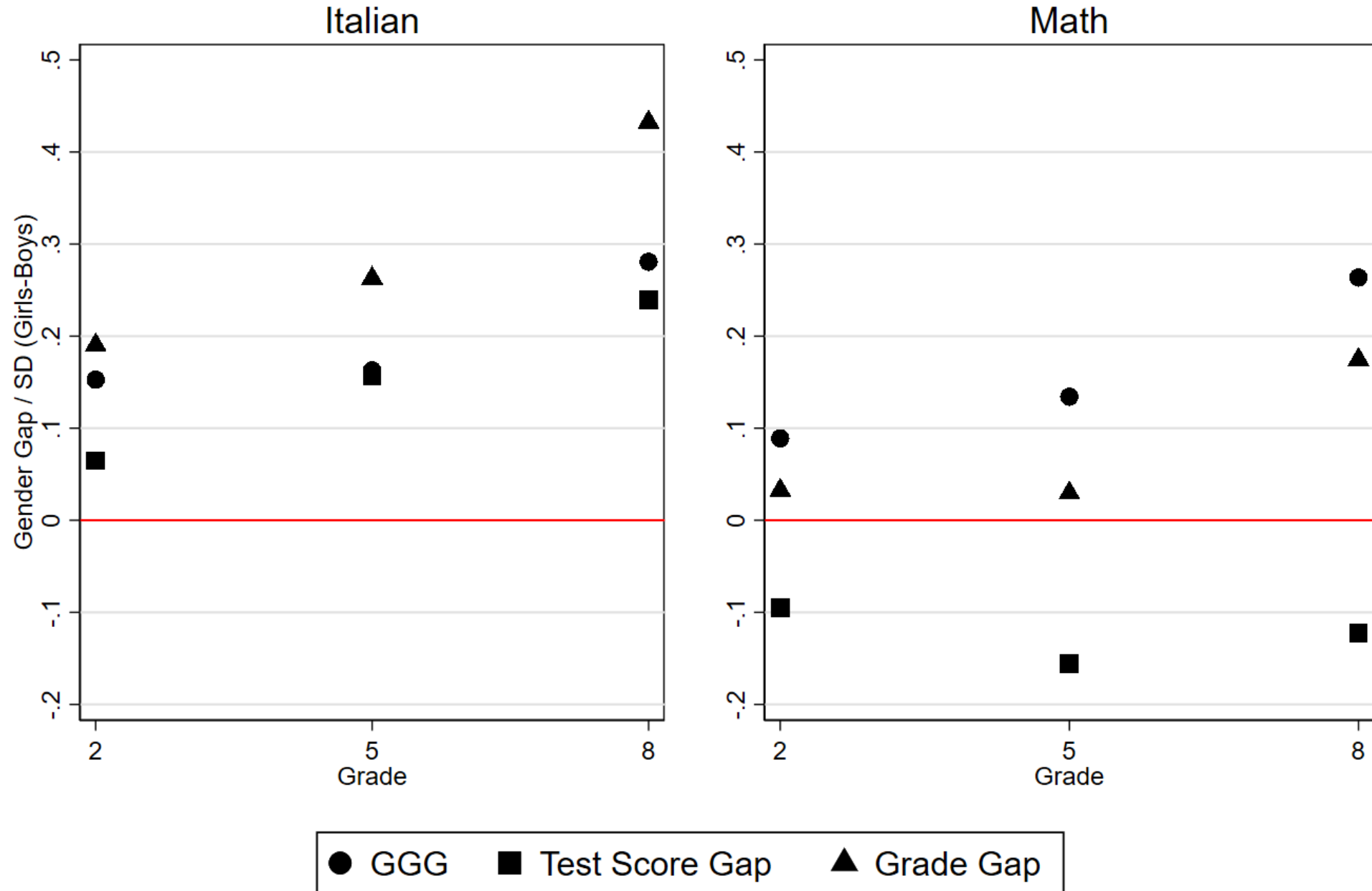
$$\begin{aligned} \text{GGG}_{c,g}^{\text{SD, subject}} = & \alpha \\ & + \beta_1 \cdot \text{ScoreGap}_c^{\text{subject}} \\ & + \beta_2 \cdot \text{AvgScoreBoys}_c^{\text{subject}} \\ & + \beta_3 \cdot \text{PctGirls}_c \\ & + \beta_4 \cdot \text{AvgSES}_c \\ & + \beta_5 \cdot \text{PctMigrants}_c \\ & + \beta_6 \cdot \text{ClassSize}_c \\ & + \gamma_t \text{ (Year FEs)} \\ & + \delta_g \text{ (Grade FEs)} \\ & + \epsilon_c \end{aligned}$$

$$w_c^{\text{subject}} = \frac{1}{\text{SE}(B_c^{\text{subject}})^2}$$

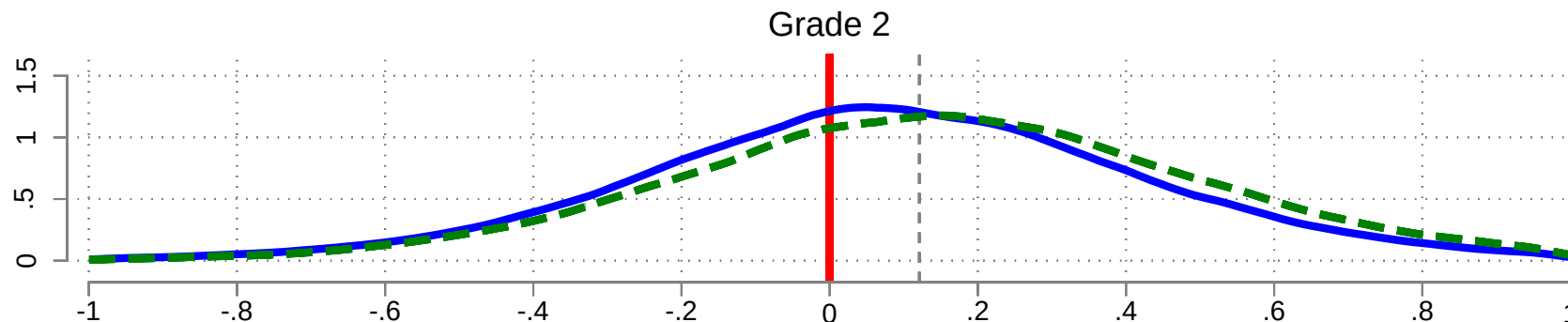
Classroom-Level Covariates



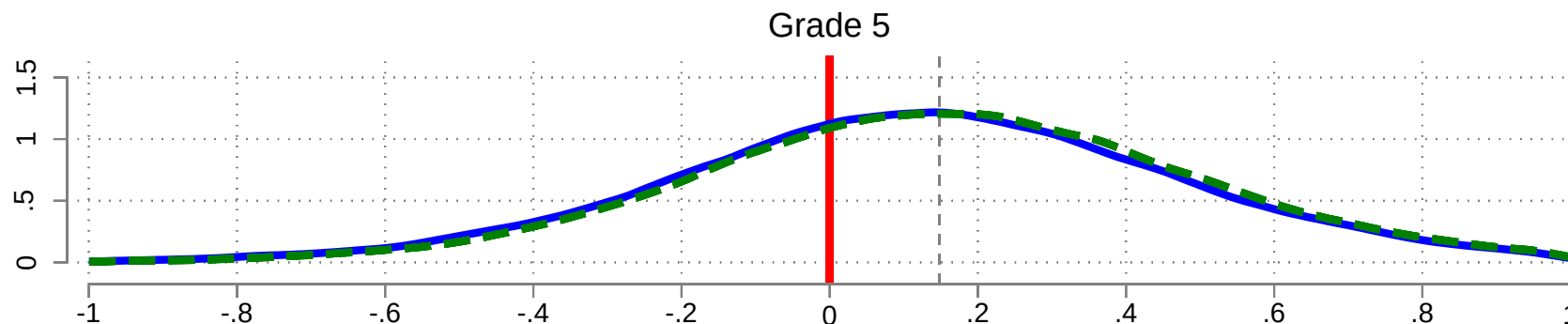
Gender Performance Gaps by Subject/Grade



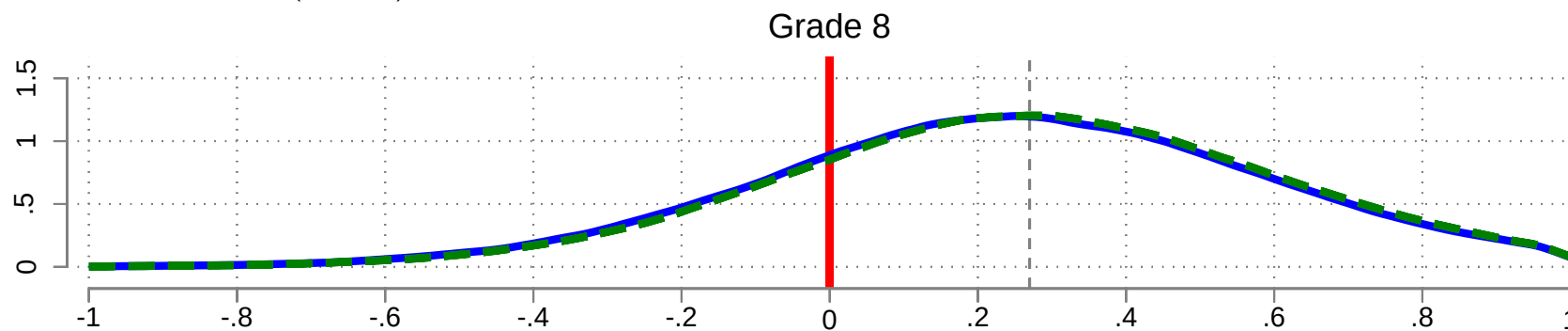
Standardized Gender Grading Gap (GGG) Distribution by Classrooms, Subject and Grade



n = 116,000 classrooms (2012/2019)



n = 116,000 classrooms (2012/2019)



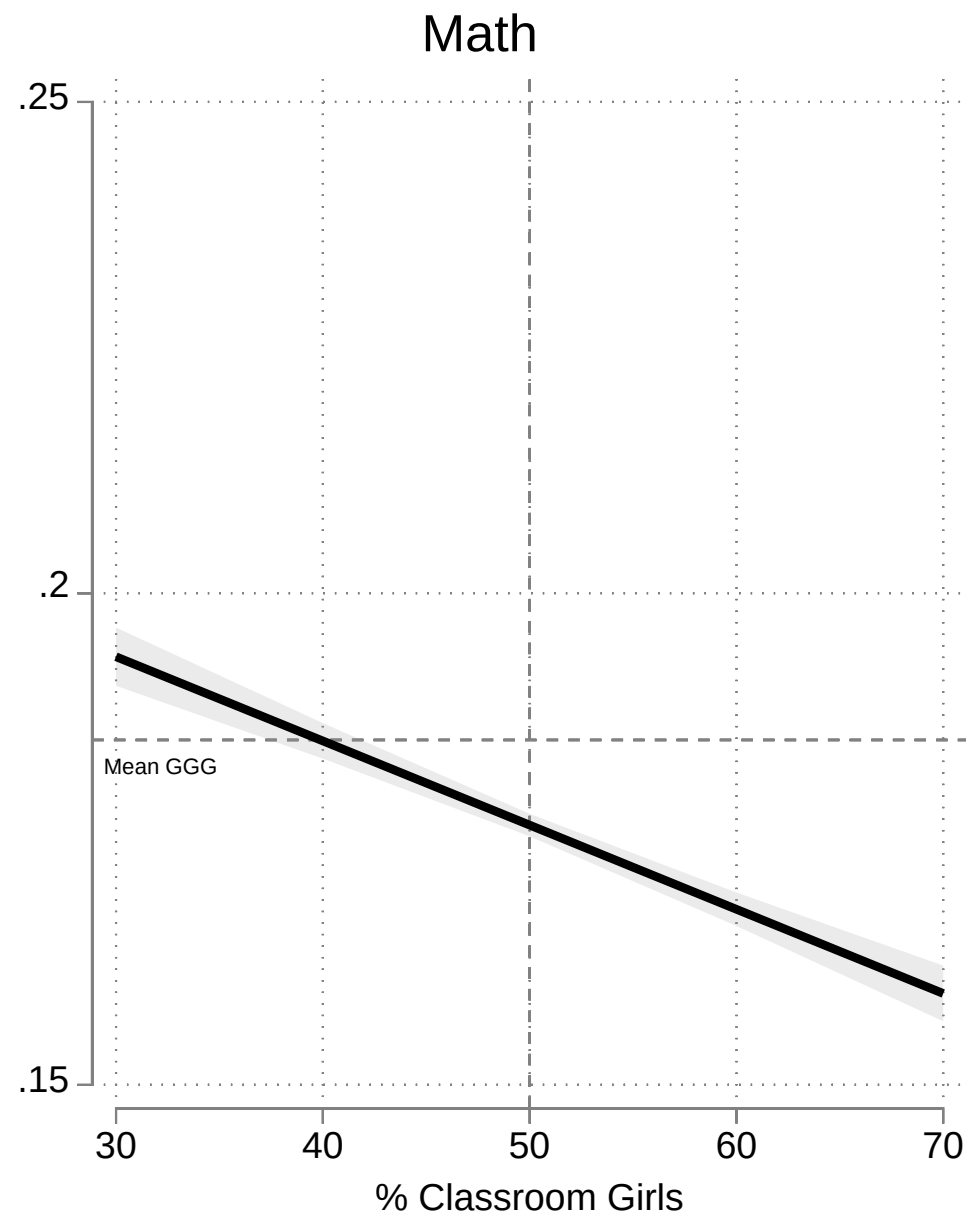
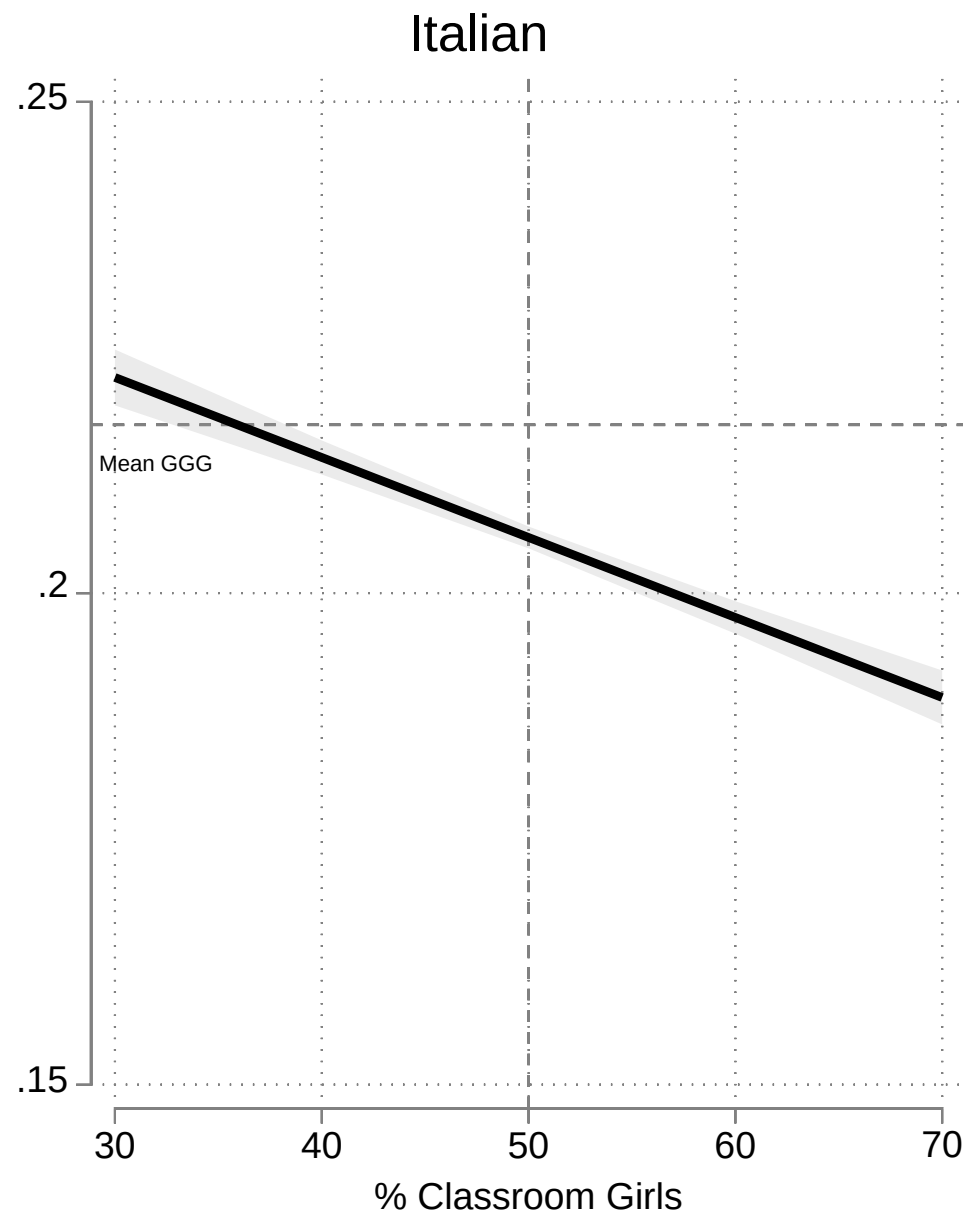
n = 217,000 classrooms (2012/2024)

Girls Disadvantage < 0 > Advantage Girls

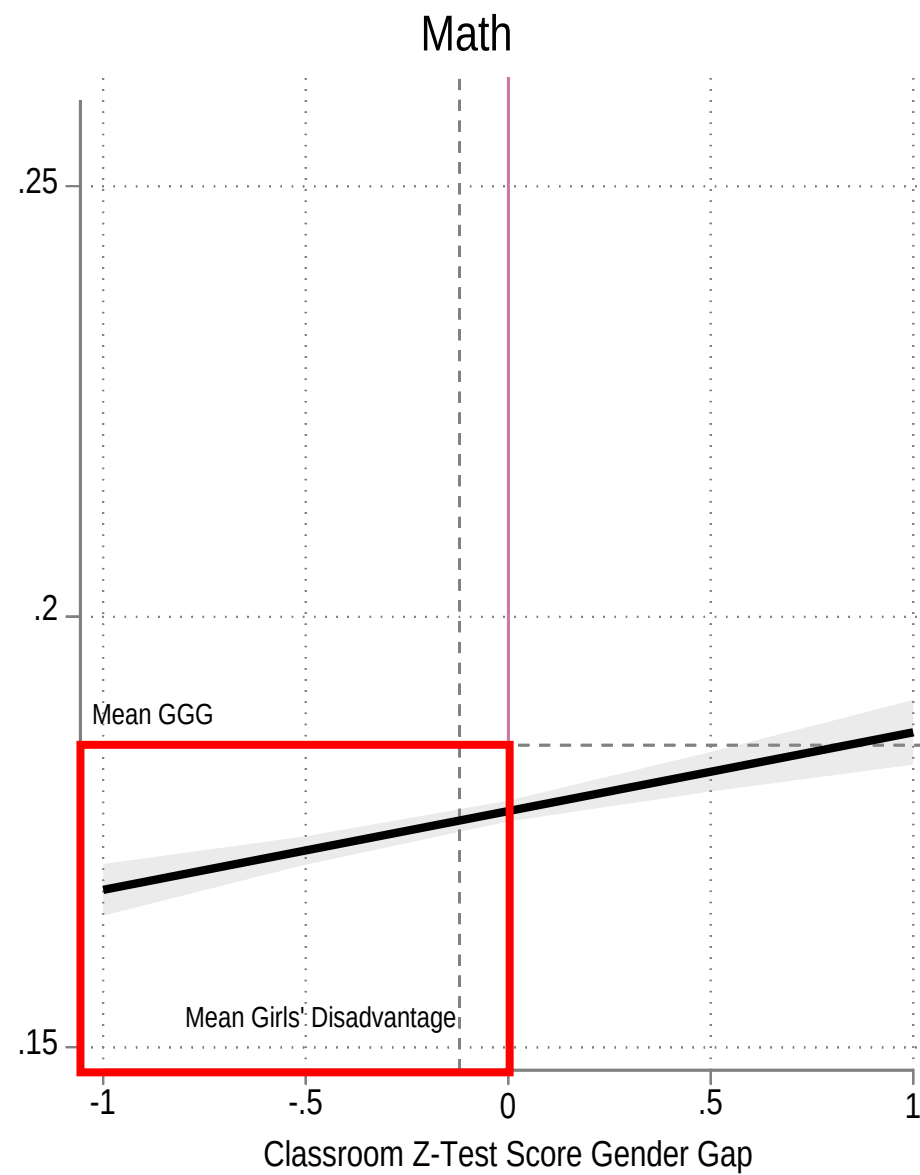
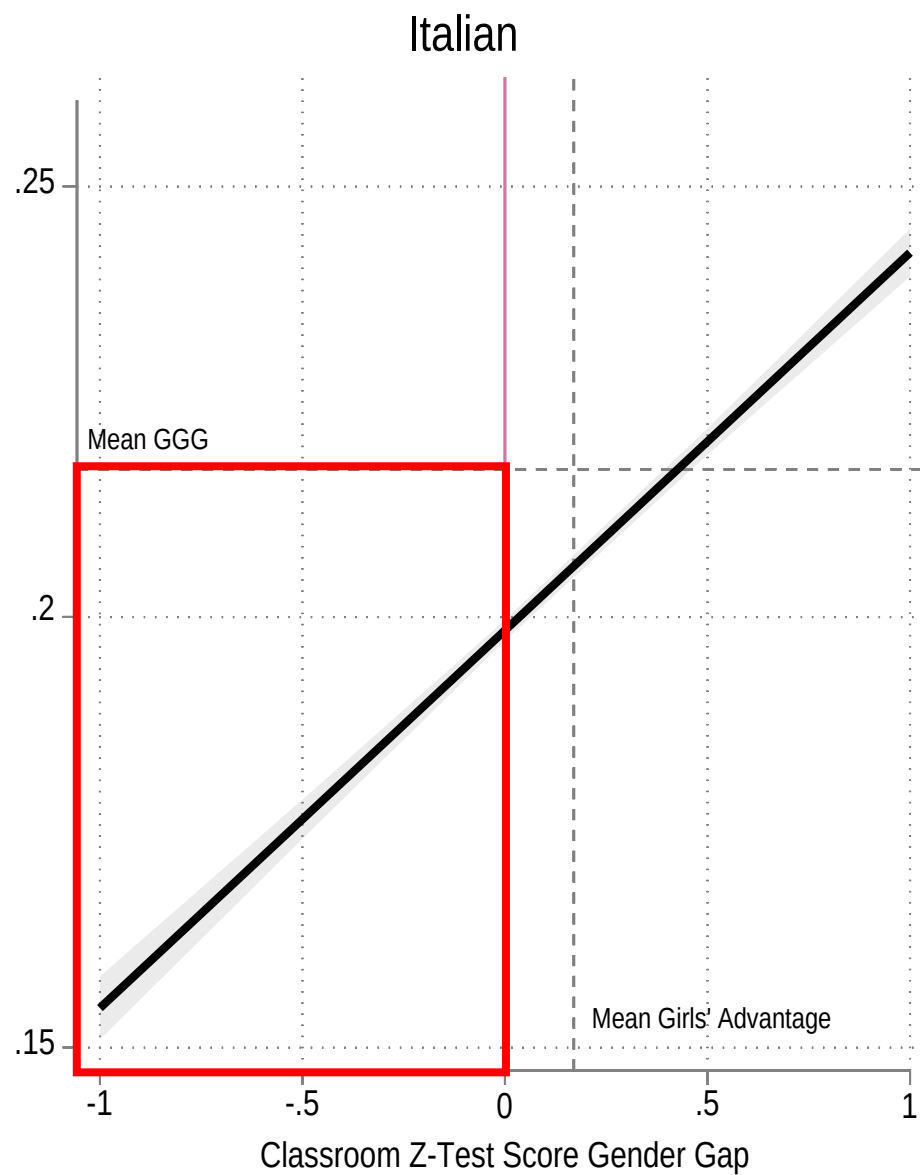


GGG SD = 0.4

Standardized GGG by % of Girls in the Classroom (SD = 10%)



Standardized GGG by Classroom Gender Gap in Test Scores (SD = 0.4)



Conclusion & Discussion

1. **Gender Grading Gap Favoring Girls ($\approx + 20\%$ SD):** similar effect size to GG in achievement
2. **Subject Heterogeneity:** GGG slightly higher in Italian than Math (test score gaps and stereotypes)
3. **Classroom Heterogeneity (% Girls):** Lower GGG in classrooms with more girls (behaviour; norms)
4. **No Compensatory Grading (?):** Grading in Italian reinforces gender performance gaps, while grading in Math remains largely unresponsive (subject-specific gender stereotypes?)

Limitations

- Measurement error in test scores (random + systematic): overestimation of GGG
- Mechanisms: student socio-emotional traits; teacher characteristics and stereotypes
- Teacher fixed effects (student-teacher matched data)
- Teacher grades are measured before test scores (learning effects)

Next Steps

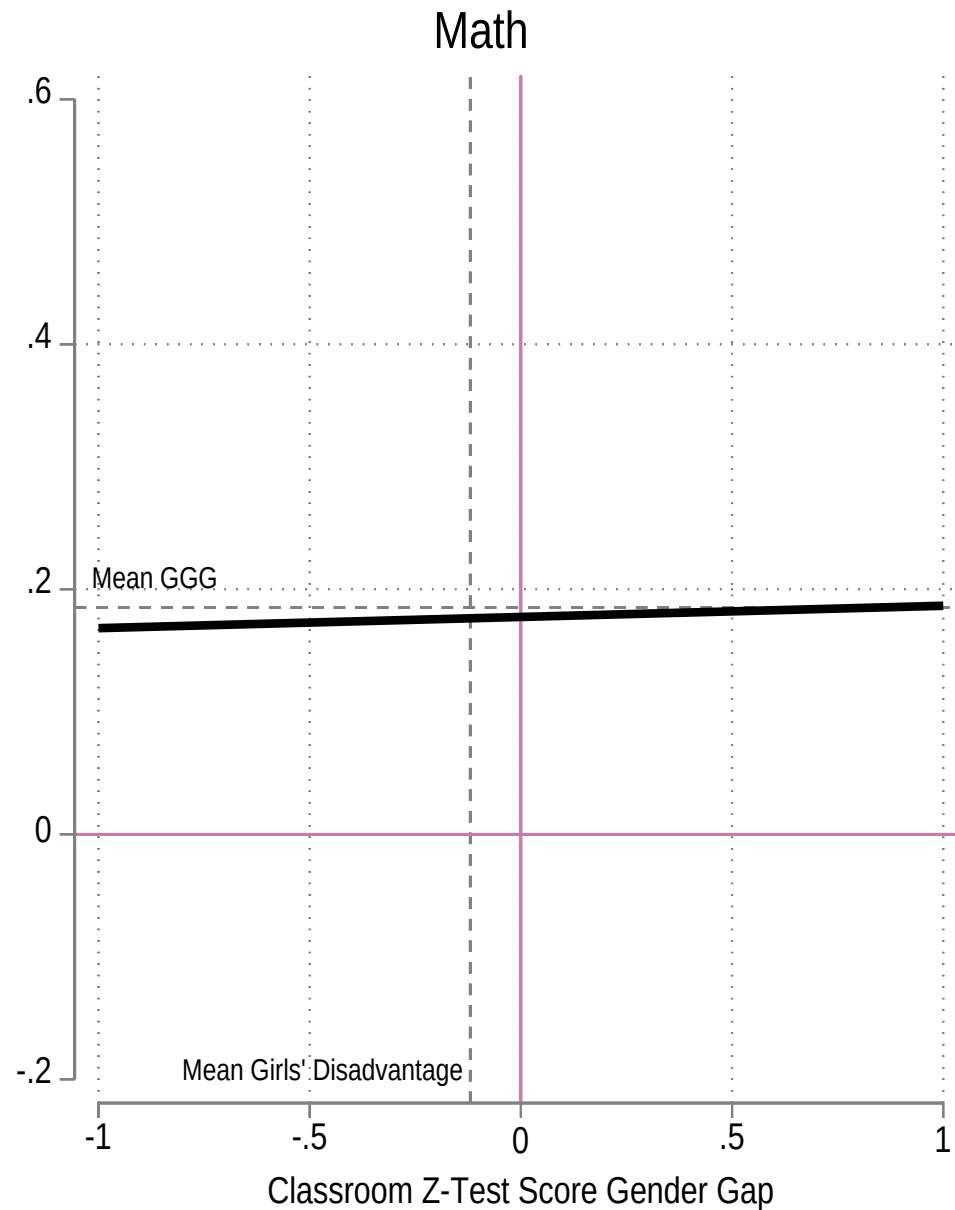
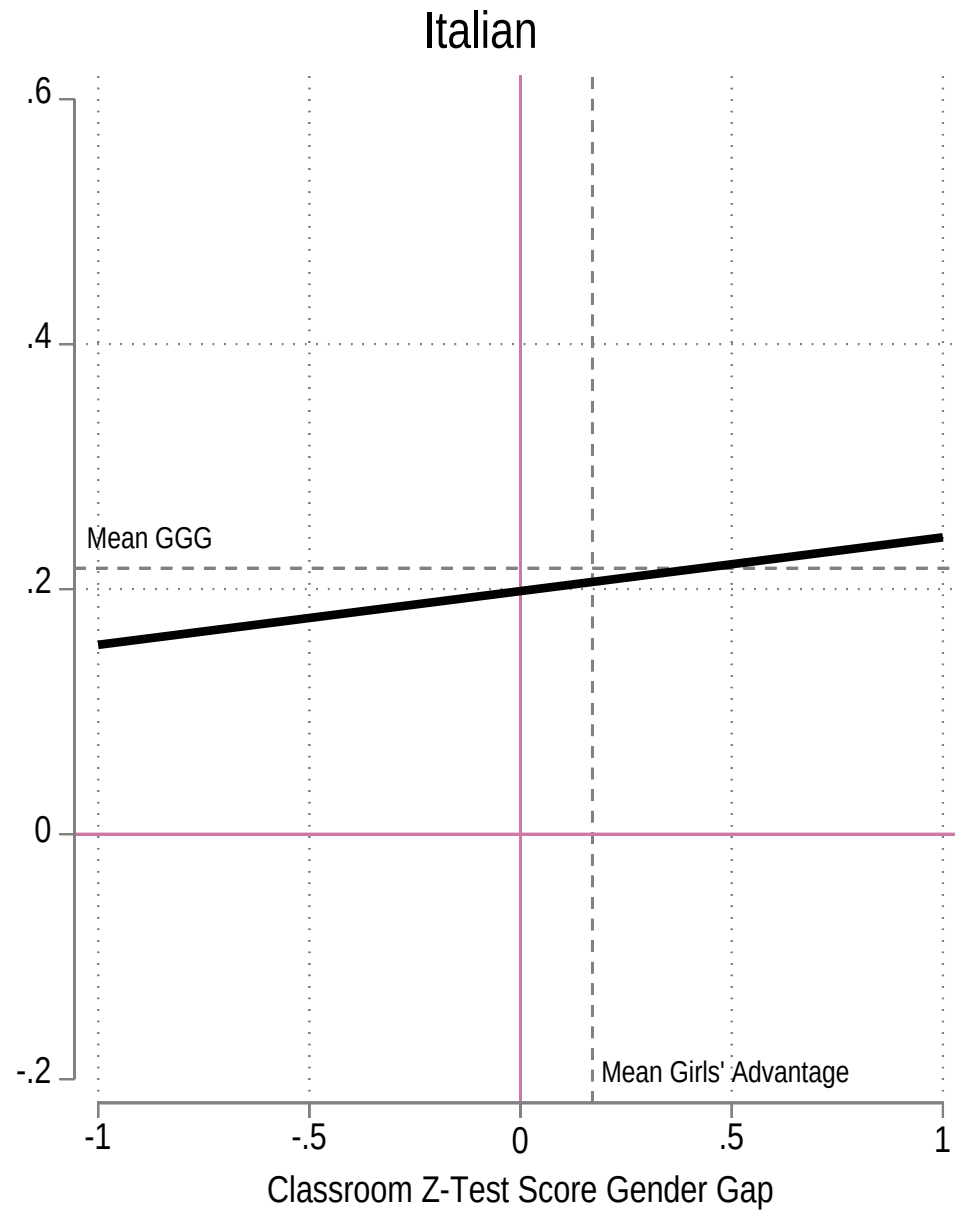
- Net GGG, adjusted by student-level ascribed status (SES, migrant origin)
- GGG heterogeneity by classroom SES and migrant-origin composition
- Period analysis by academic years
- Longitudinal dataset to assess GGG effects on later performance



X SEMINAR
“DATA FROM AND FOR THE EDUCATIONAL SYSTEM: TOOLS FOR RESEARCH AND TEACHING”
ROME, 19 - 20 - 21 NOVEMBER 2025

Thanks for your attention! 😊
carlos.gil@unifi.it

Standardized GGG by Classroom Gender Gap in Test Scores (SD = 0.4)



GGG by % of Girls in the Classroom (SD = 0.4)

