doi.org/10.26398/asaproc.00109

School socioeconomic status and students' achievements: Exploring territorial heterogeneity through a multilevel regression discontinuity approach

Rosa Fabbricatore[a], Pasquale Sannino[b], Cristina Davino[a]
[a] Department of Economics and Statistics, University of Naples Federico II, Naples, Italy.
[b] Department of Economics and Law, University of Macerata, Macerata, Italy.

1. Introduction

The debate on how socioeconomic status affects students' performance and educational opportunities dates to the 1960s, a period when this issue was first brought to light by what we can consider the pioneering work, the Coleman Report (Coleman et al., 1966). Among the findings of the report, one of the most important is that a student's level of education is correlated with both their family background and the background of other students in the school. Another line of literature has shown that, regardless of the family's socioeconomic status, a significant factor affecting students' academic performance is the average socioeconomic status of the school, highlighting that schools with a higher average socioeconomic status tend to have higher test scores (Rumberger & Palardy, 2005; OECD, 2004). Thus, in a world increasingly leaning towards equality, it is in the interest of institutions to monitor this trend between socioeconomic status and school performance to ensure equal opportunities for students regardless of their socioeconomic conditions, both family and school, aiming to improve education. Indeed, higher education can enable a better standard of living by providing broader access to various job opportunities. Individuals with a stronger cultural background can more easily access to qualified and better-paid jobs, enjoying concrete prospects for professional advancement. On a larger scale, this phenomenon contributes to reducing the income gap between different segments of the population (Abdullah et al., 2015).

In this vein, assessing students' achievements has become of paramount importance worldwide to highlight any disparities between schools, regions, and geographical areas. In Italy, this work is carried out by the INVALSI (National Institute for the Evaluation of the Educational System of Education and Training) through the annual administration of the so-called "INVALSI Tests," which are periodic and systematic assessments of students' skills and knowledge at the end of the second and fifth grades of elementary school, the third grade of middle school, and the second and fifth years of high school. The tests are prepared and provided by INVALSI and are intended by the Ministry of Education to assess students' proficiency in Italian, Mathematics, and English. The purpose of the INVALSI tests is multiple: to verify the level of student learning, monitor school performance, and analyse if there are differences between different schools, regions, or geographical areas of the country.

The present work aims to investigate the impact of the socioeconomic status of the schools attended by Italian students on the mathematics scores obtained in the INVALSI tests, with a particular focus on regional differences. Indeed, Italy is a country characterized by significant regional differences that manifest in many aspects, from the economic aspect to the quality of life up to education. These disparities are particularly evident when comparing the northern regions with those of the south. The study proposes a multilevel regression discontinuity based on the integration of Regression Discontinuity (RD) and Multilevel regression (Hox et al., 2017). On the one hand, the RD design represents a non-experimental strategy for analysing causal effects, which allows us to estimate the causal effect of attending a school with a low socioeconomic status on the results of the INVALSI Mathematics tests. On the other hand, the multilevel specification provides new insights into Italian territorial heterogeneity by examining regional variations in the effect of school socioeconomic status.

The remainder of the paper is organized as follows: Section 2 focuses on the methodology employed, specifically detailing how the multilevel approach was integrated with the regression discontinuity technique to achieve the research aims. The data used in the analysis are described in detail in Section 3, while Sections 4 and 5 present the empirical results and some conclusions, respectively.

2. Methods

The RD design, introduced by Thistlethwaite and Campbell in 1960, is a highly credible non-experimental technique for analysing causal effects (Cattaneo et al., 2019; Cattaneo et al., 2024; Imbens & Lemieux, 2008). Typically, experimental designs, which are studies aiming to evaluate the effectiveness of a treatment, require a treatment group and a control group that differ only in their treatment status. RD is considered a quasi-experiment because receiving the treatment depends wholly or partly on whether an observable variable (called running variable, forcing variable, or index) exceeds a certain threshold value (called cutoff). Thus, in RD, there are three fundamental components: *i)* X_i, which is the variable on which the assignment mechanism to the treatment or control group is based (assuming that there are n units indexed with $i = 1,2, \dots, n$), *ii)* the cutoff c, which is an externally determined threshold beyond which units are assigned to the treatment group, *iii)* the treatment D_i, which is a binary variable equal to 1 if the unit i is part of the treatment group and 0 otherwise (the treatment can consist, for example, of implementing an economic policy, an intervention in education, etc.). Each unit is assigned a specific value of the running variable, and treatment status is guaranteed to all units whose running variable value exceeds the known cutoff point and not assigned to those whose running variable value does not.

Several approaches can be employed to estimate the treatment effect formally. When the research aim involves the detection of cross-site heterogeneity of the treatment, random effects models (Hox et al., 2017) proved to be valuable. Accordingly, let Y_{ij} be the potential outcome of the unit i belonging to the site j, the two-level regression model with random effects for the intercept and slope can be expressed as:

$$Y_{ij} = \beta_{0j} + \beta_1 (X_{ij} - c) + \beta_{2j} D_{ij} + r_{ij}; \tag{1}$$

Where:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{2j} = \gamma_{20} + u_{2j}.$$

In the above model, γ_{00} represents the mean outcome score, across sites, of units with a running variable equal to the cutoff; γ_{20} defines the treatment effect across sites (the parameter we are interested in); u_{0j} and u_{2j} reflect the variation of γ_{00} and γ_{20} between sites, respectively. Finally, β_1 indicates the general effect of the running variable on the outcome.

Notably, in the approach herein proposed, we tried to improve further what is typically done in multilevel RD implementation in education (e.g., Luyten, 2006; Steinmann & Olsen, 2022) by addressing two main issues of the RD method regarding bandwidth selection and unit weighting. Indeed, the underlying idea of RD is that there exists a certain window around the cutoff point in which belonging to the treatment or control group is almost random, which resembles experimental designs. Moreover, considering a certain weighting function ensures that units closer to the cutoff have a higher weight during model estimation. Specifically, our step-wise approach can be summarized as follows: Step 1) estimating a RD model to select the optimal bandwidth; Step 2) assigning a weight to each observation within the bandwidth according to a kernel function; Step

3) estimating a multilevel model as specified in Eq. (1) considering the sub-sample of units within the selected bandwidth and the corresponding weights. The R statistical software can be used for data analysis, exploiting the packages `rdrobust` (Calonico et al., 2014) and `lme4` (Bates et al., 2015) for steps 1 and 3, respectively.

3. Case study on INVALSI Mathematics tests

The study considered the data from the INVALSI Tests[1] ("Prove INVALSI") administered to 13th-grade students, corresponding to the fifth year of high school, between March and May of the 2022/2023 academic year. After a careful preprocessing phase, we have data on 424,400 students. The outcome variable is represented by the score on the INVALSI Mathematics tests, measured using a Rasch model to account for the difficulty level of the questions and to make the performance of students, classes, and schools comparable. The running variable is the socioeconomic status of the school attended by each student, herein defined according to the ESCS (Economic, Social and Cultural Status) index. The ESCS is a standardized index derived from a principal component analysis of three indicators proposed by INVALSI (Campodifiori et al., 2010): *i)* HISEI – the parents' occupational status, *ii)* PARED – the parents' level of education, *iii)* HOMEPOS – the possession of certain material goods. Thanks to the school offices, the ESCS is measured at the individual level and then aggregated into the class and school ESCS by taking the average. The cutoff is represented by the value set by the Ministry of Education and Merit in Ministerial Decree No. 90 of May 19, 2023[2], which is -0.31243 for secondary schools (from this value, the ESCS is considered low). The treatment group consists of all students enrolled in a school with low ESCS, and consequently, the treatment corresponds to attending a school with low ESCS. As mentioned earlier, treatment status is guaranteed for all units whose running variable value exceeds the known cutoff point. In RD designs, it is customary for the treatment group to be on the right and the control group on the left of the cutoff. To achieve this, we multiplied the running variable by -1, ensuring that all students enrolled in a school with low ESCS are on the right side of the cutoff.

In Figure 1, a set of violin plots represents the school ESCS distribution according to the Italian regions that are placed in ascending order based on the median ESCS value. The trend is quite clear: the southern regions are almost all positioned at the lower part of the graph, indicating a well-known geographical disparity.

4. Results

Before implementing the RD, we ensured that our study was suitable by verifying the assumptions. Specifically, we checked the "no sorting around the cutoff" assumption (McCrary, 2008), which states that individuals cannot sort themselves around the cutoff; they cannot decide whether to be to the right or the left of it. Since the running variable is discrete, we tested this assumption using the test proposed by Frandsen (2017) and found no evidence of possible sorting. To select the optimal bandwidth in Step 1, we used the algorithm that minimizes the Mean-squared error (MSE), which accounts for the so-called "bias-variance" tradeoff. In Step 2, observations within the bandwidth were weighted with a triangular kernel function to ensure that observations closer to the cutoff had a greater weight in the analysis. The results obtained from the RD implementation show that the selected bandwidth is 0.175 to the right and 0.175 to the left of the cutoff, reducing the analysis to an effective number of observations of 54,470 on the left and 42,205 on the right.

[1]https://www.invalsiopen.it/

[2]https://www.miur.gov.it/documents/20182/7414469/m_pi.AOOGABMI.Registro+Decreti%28R%29.0000090.19-05-2023.pdf/44e239ff-2151-8c6a-3d45-851a38834ded?version=1.0&t=1690386661238

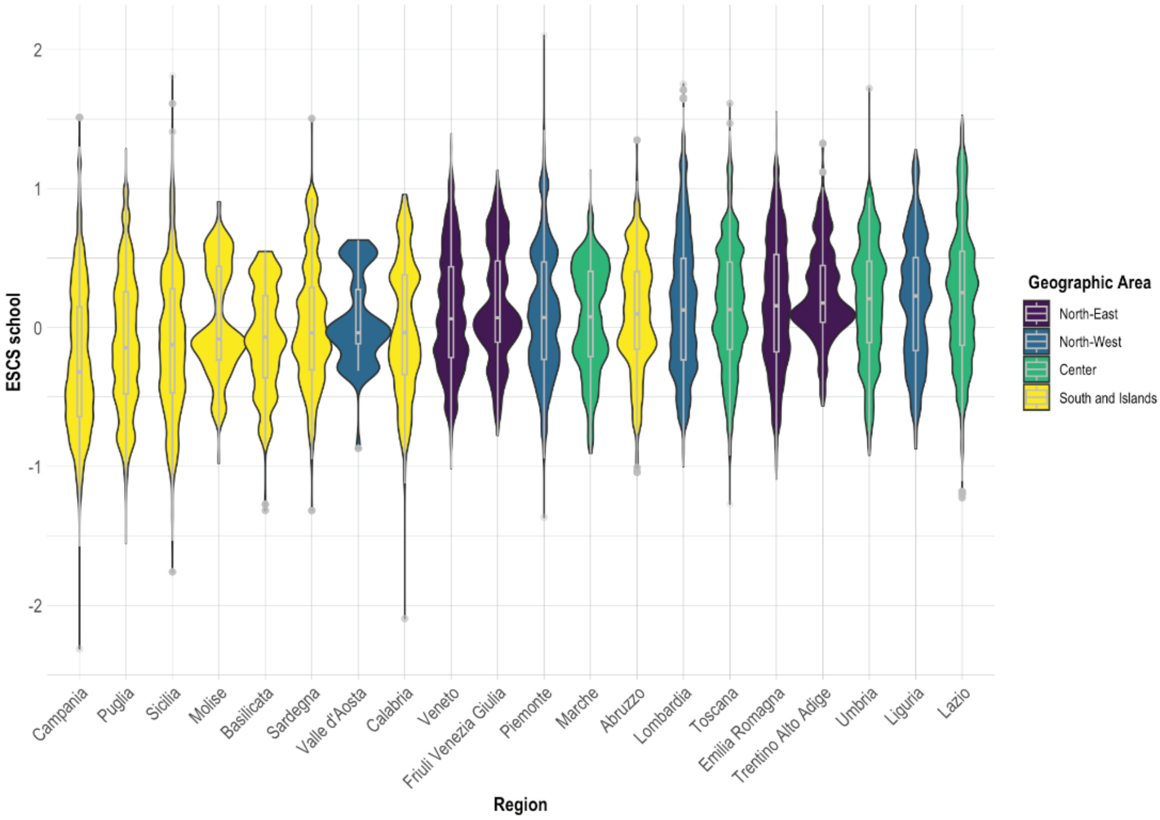

Figure 1 – Violin plots for school ESCS distribution according to Italian regions. Colors indicate the geographical macro-area.

Then, we estimated the multilevel model in Step 3. First, we compared the fully random model specified in Eq (1) (Mod0) with simpler versions, which are the random intercepts model (Mod1) and the null random intercepts model (Mod2). Results reported in Table 1 suggest that Mod0 is preferred for our data, indicating that both intercept and slope (the treatment) significantly vary across regions. Specifically, in Mod0, the fixed effects are equal to $\gamma_{00} = 181.97$ for the intercept, $\beta_1 = -52.29$ for the centered running variable $(X_{ij} - c)$ and $\gamma_{20} = -1.84$ for the treatment effect. Hence, overall, students enrolled in a school with a low ESCS experience a negative causal effect, resulting in a loss of 1.84 points in the INVALSI Mathematics tests. Concerning random effects, the variances of level 2 (regions) residuals are equal to $\sigma^2_{u_0} = 87.11$ and $\sigma^2_{u_2} = 22.19$, with a negative correlation $\rho = -0.65$. Finally, the variance of level 1 (individuals) residuals is equal to $\sigma^2_r = 453.58$.

Table 1 – Fit statistics for multilevel models

Model	#par	AIC	BIC	LogLik	χ^2	p-value
Mod2	3	963,354	963,382	-481,674		
Mod1	5	961,751	961,798	-480,870	1607.00	< 0.001
Mod0	7	961,309	961,376	-480,648	445.49	< 0.001

The caterpillar plots in Figure 2 provide more details on regional differences, showing the estimates of second-level random effects and the related 95% confidence interval. Results reveal that northern regions generally report significantly better performance in the INVALSI Mathematics tests than the overall mean, whereas southern regions present the opposite trend. Moreover, given the negative sign of the treatment effect γ_{20}, the residuals u_{2j} at Level 2 highlight that being enrolled in a school with a low ESCS has a stronger negative effect on students' performance in the northern regions than in the southern ones. An exception is the Veneto region, which reports a positive residual unlike the other northern regions.

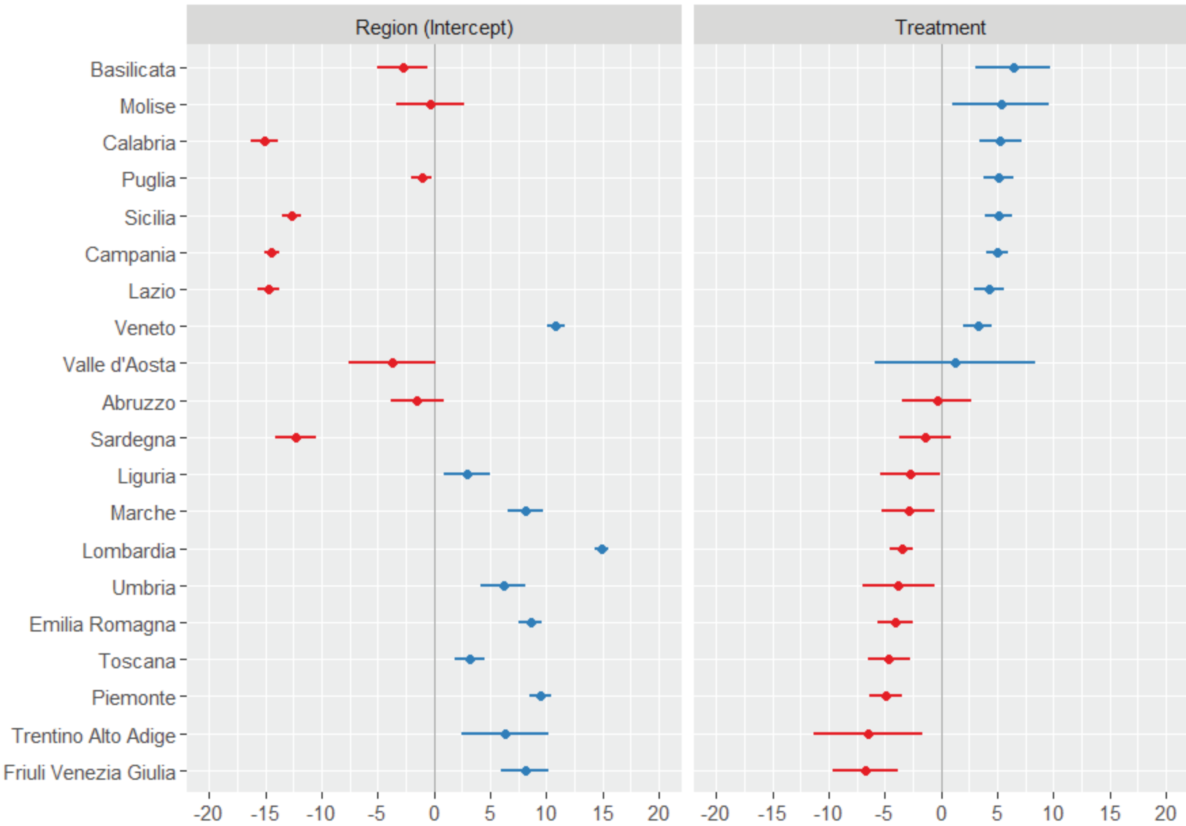

Figure 2 – Caterpillar plots: Estimated second-level random effects with 95% confidence interval for the intercept and the slope (treatment).

5. Conclusion

By integrating multilevel modelling with RD design, this study provides new insights into the interplay between individual, school, and regional factors affecting students' performance in INVALSI Mathematics tests. Results highlight that attending a school with a low ESCS has a negative causal effect on students' performance in mathematics. Moreover, evidence from the territorial heterogeneity in Italy shows that the southern regions generally report the worst mean score in the INVALSI Mathematics tests whereas the northern regions present a stronger negative causal effect of school ESCS on students' achievements. The cause of this more pronounced negative causal effect for regions in the North could be explained by: *i)* a resilience factor: students in the South might have a more resilient outlook towards challenges and disadvantaged contexts; *ii)* a contextual factor: in the South, schools may already operate under generally more unfavourable conditions, and therefore the additional disadvantage associated with a low ESCS index might be less evident, as it is compounded with other pre-existing factors.

This analysis provides significant insights for policymakers, suggesting that a tailored approach to regional needs could be the solution to addressing educational inequalities. Indeed, the emerging regional variations may be influenced by factors such as economic development and educational policies. Future research could focus on either methodological advancement in integrating the three proposed steps into a single-step approach or a more in-depth investigation of the identified regional differences, particularly the uncommon behaviour observed for the Veneto region.

Acknowledgment

Rosa Fabbricatore acknowledges the financial support provided by the European Union - NextGenerationEU, in the framework of the GRINS - Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 - CUP E63C22002140007). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

This work was also supported by the University Research Funding Program (FRA) 2022 of the University of Naples Federico II, with the contribution of the Compagnia San Paolo (Title of the project: "Measuring and Mapping Educational Poverty"; principal investigator: Cristina Davino).

References

Abdullah, A., Doucouliagos, H., Manning, E. (2015). Does education reduce income inequality? A meta-regression analysis. *Journal of Economic Surveys*, **29**(2), pp. 301-316.

Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), pp. 1-48.

Campodifiori, E., Figura, E., Papini, M., Ricci, R. (2010). Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia. *Working paper INVALSI n.2*. Available from: https://www.invalsi.it/download/wp/wp02_Ricci.pdf

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., York, R.L. (1966). *Equality of Educational Opportunity*. Government Printing Office, Washington, DC, (U.S.).

Calonico, S., Cattaneo, M.D., Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, **14**(4), pp. 909-946.

Cattaneo, M.D., Idrobo, N., Titiunik, R. (2019). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge University Press, Cambridge, (UK).

Frandsen, B.R. (2017). Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. *Regression Discontinuity Designs: Theory and Applications*, **78,** pp. 281-315.

Hox, J., Moerbeek, M., Van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications*. Routledge, New York, (U.S.).

Imbens, G.W., Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* **142**(2), pp. 615-635.

Luyten, H. (2006). An empirical assessment of the absolute effect of schooling: regression-discontinuity applied to TIMSS-95. *Oxford Review of Education*, **32**(3), pp. 397-429.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142**(2), pp. 698-714.

OECD (2004). *Education at a Glance 2004: OECD Indicators*, OECD Publishing, Paris.

Perry, L.B., McConney, A. (2010). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teachers College Record*, **112**(4), pp. 1137-1162.

Rumberger, R.W., Palardy, G.J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record*, **107**(9), pp. 1999-2045.

Steinmann, I., Olsen, R.V. (2022). Equal opportunities for all? Analyzing within-country variation in school effectiveness. *Large-Scale Assessments in Education*, **10**(1), pp. 2.

Thistlethwaite, D.L., Campbell, D.T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, **51**(6), pp. 309-317.