Regional gap in students’ performance at the quantiles

Marilena Furno

Received: 28 February 2022 / Accepted: 25 July 2022 / Published online: 5 September 2022
© The Author(s) 2022

Abstract
The average math score in Italy sizably improved in 2009. The quantile regression analysis shows that the determinants of the improvement are related to school resources. The decomposition at the quantiles shows that school characteristics are relevant in the north-center regions. In the south are relevant both covariates and coefficients. The coefficient component is unexplained by the model and is linked to changes in students’ attitude, to a more favorable disciplinary climate. The southern regions gap decreases but does not disappear. The improved southern students’ attitude toward school, if supported by improvements in school characteristics, could further reduce the regional gap.

Keywords Regional discrepancy · Quantile regression · Quantile decomposition

JEL Classification A2 · C2

1 Introduction
This study focuses on the scores at the international OECD-PISA math test for 15 years old Italian students. The test scores are here related to school characteristics and the link between student performance and school environment is analyzed at the quantiles. Most analyses rely on average values whereas it might be more effective for policy purposes to focus on the tails rather than on the ordinary least squares (OLS) estimated regression. OLS computes the conditional mean, and averages are not an appropriate indicator if the goal is to improve the lower scores and increase the general performance. Therefore, the selected model is estimated at the center and in the tails of the conditional distribution, by means of quantile regressions. We estimate the equation at various points of the conditional distribution of the dependent variable, such as for instance the first and third quartiles—representing respectively
the lower and the higher scoring students—and other quantile of interest, like the deciles.

The math performance of Italian students has significantly improved in year 2009 to remain stable afterward.\(^1\) We focus on this improvement and look for its source not only nationwide but also at regional level, by sorting out two macro-regions, the leading north-center and the economically as well as educationally lagging south. The nature of the improvements, school or student driven, provide insights for policy makers to further advance the students’ performance having low/median/high math scores.\(^2\) The results show improvements in school resources, like greater availability of computers and textbooks and labs, hindered by the widening of the gender gap and the worsening of the academic track, mostly attended by girls.

The quantile regressions analysis for low, average, and high scoring students is complemented by a quantile regression decomposition to analyze changes over time. The changes in math scores are decomposed at various quantiles in terms of the impacts of the covariates (the school characteristics) and the coefficients (mirroring effects outside school control) to check whether the respective changes are due to variations in the explanatory variables over time, or in their impact on the math scores. Machado and Mata (2005) introduce the quantile regression decomposition, and Chernozhukov et al. (2013) provide inference tools.

Decomposition analyses are usually computed on average using the Oaxaca (1973) Blinder (1973) approach, as in Gigena et al. (2011) or Barrera-Osorio et al. (2011) or Furno (2021). However, a decomposition based on averages does not allow us to investigate the tails of the student score distribution. Indeed, the impact of the coefficients and/or the covariates may change across quantiles and average values will not signal this behavior. Or the impact may increase at some quantiles and decrease in others. The result would balance on average thus hiding the tail behavior.

Here, we implement the quantile regression decomposition not only with respect to time but also looking at regional characteristics, thus considering the impact of the Italian regional divide as well. This analysis provides results that would not emerge otherwise, and that can be summarized as follows.

1) The average Italian math score has been below the OECD average up to 2006 and raised above the OECD average in 2009. To analyze the source of the improvement we implement a decomposition analysis over time, comparing year 2009 with the previous years. The analysis shows that school covariates steadily improved at all quantiles while the effect of the coefficients declined across quantiles changing its sign at and above the median, thus curbing the total change. The

---

\(^1\) In 2003 the score is high and shows a good students’ performance. However, the 2003 questionnaires were collected mostly in the northern regions and the result is less representative (see footnote VI for further details).

\(^2\) Growth analysis emphasizes school attainment and links it to cross countries differences in economic growth (Barro 1997). Mathematical skills are considered a good proxy for labor quality, and Hanushek (2006) relates a one standard deviation difference in students’ math performance to a 1 percent difference in annual per capita GDP growth rates.
increase in math scores is explained by improvements in the school environment while the effect of the coefficients, which gathers the components unexplained by the variables of the model, improves in 2009 only for the lower scoring students.

2) Next the student performance is decomposed at various quantiles separately in the north-center and in the southern regions, since the economic lag of the Italian southern regions is reflected in the education system (Seta et al., 2014).3 In the north-center the increase in scores over time is linked to improvements in the explanatory variables, i.e., is due to improvements in school characteristics. The improvement is partially or totally counterbalanced by the opposite sign of the coefficients effect at all quantiles that becomes quite sizably at the top decile. In the southern schools the advances are due to favorable changes in both covariates and coefficients, the latter particularly ample at the lower deciles. The coefficients effects signal changes in the conditional distribution of students’ scores that are unrelated to school characteristics. They may be due, for instance, to an improved students’ attitude toward education.

3) The combination of regional and temporal changes shows that the southern regions improvements have reduced the regional gap in 2009 but the gap is far to be closed. The regional decomposition within each period shows that northern school coefficients effects in 2009 are greater than in the south, although the increase/gap is smaller than in the previous period.

In sum, the increased average performance in math can be related to a reduced regional gap and to an improved southern coefficients effect, which is possibly due to a more favorable disciplinary climate in southern schools. These effects could be further increased by rising southern school covariates, although the discouraging impact of a wide unemployment rate in these regions could curb the results.

The novelty of the present work is in the quantile regression decomposition, which shows the dissimilar patterns of covariates and coefficients across regions and over time. These patterns cannot be evaluated by the Oaxaca-Blinder average decomposition analysis, and to the best of our knowledge such quantile-based decomposition of students’ attainment has not been analyzed elsewhere. The decomposition at the quantiles allows us to discriminate between coefficients and covariates effect for low/medium/high scoring students. The covariates effect shows an improvement in school endowment while the coefficients effect shows the presence of changes differing from school characteristics.

The source of a discrepancy in the covariates—i.e., explained by changes in the covariates over time, and in the coefficients—i.e., not explained by the selected model, may differ across quantiles, as occurs in this analysis. The two effects may reinforce one another or may cancel out at some/all quantiles. It may be the case that an increasing effect of the covariates across quantiles counterbalances a decreasing coefficients effect, thus providing a stable overall discrepancy.

3 Southern regions unemployment rate is particularly high for young generations and has a discouraging impact on school attainment. In 2009 the southern unemployment rate is 28.5% for young workers aged between 15 and 29 years, versus a nationwide rate of 18.5% (source: ISTAT at http://dati.istat.it/Index.aspx?DataSetCode=DCCV_TAXDISOCCU).
In our model, the regional decomposition shows how the improved performance is linked to better educational resources in the north curbed at the top quantiles by worsened students’ attitude effect. In the south better resources are instead coupled with improved students’ attitude at most quantiles, particularly at the lower ones. These effects result in an overall reduction in the regional gap in students’ performance both at the top quantiles, due to the curbing coefficient effect in the center-north, and at the lower tail, due to the favorable coefficient effect in the south.

The analysis provides clear policy implications. If the goal is the improvement of math performance of the low scoring students, at the low quantiles, the focus is on girls, on school size, on private schools, on math teaching in academic track, as shown by the simple quantile regression analysis. If the goal is to reduce the regional gap, the focus is on increasing resources to southern schools, as revealed by the quantile regression decomposition, coupled with policies quickening the job placement in the south. However, there is a wide-ranging debate on the role of school variables on student proficiency and the sole increase of school resources may not attain the goal.4

2 The quantile regression approach

The linear regression model \( y = x\beta + e \) can be estimated at different points of the conditional distribution, beyond the OLS/conditional mean. If the goal is to estimate the model in the tails, as in the quantile regression estimator, an asymmetric weighting system is introduced to drive the estimated regression above or below the OLS/conditional mean. The quantile regression objective function at the quantile \( \theta \) is given by

\[
\sum_{y_i \leq x_i \beta} (1 - \theta) |y_i - x_i \beta| + \sum_{y_i > x_i \beta} \theta |y_i - x_i \beta| = \sum_i \{ \theta - 1(y_i \leq x_i \beta) \} (|y_i - x_i \beta|)
\]

For instance, to estimate the 75th quantile regression, i.e., the regression passing through the 3rd quartile of the dependent variable, the estimated equation will be characterized by 75% of the residuals below the 3rd quartile, and 25% of the residuals above it. It represents the regression passing through the third quartile of the conditional distribution of the dependent variable, given the selected covariates. This result is achieved by introducing asymmetric weights which assign the value \( \theta = 0.75 \) to the larger observations to attract the estimated equation upward, and \( 1 - \theta = 0.25 \) to the remaining data. If the conditional distribution

\[
\text{OECD Country Note (2012, pg 3) states: “After a certain threshold of cumulative expenditure is reached (roughly USD 50 000), the relationship between spending per student and performance is no longer apparent. For example, Italy and Singapore both spend roughly USD 85 000 on education per student between the ages of 6 and 15, but while Italy scored 485 points in mathematics in PISA 2012, Singapore scored 573 points. Italy and Norway on the other hand have similar levels of performance (485 and 489 points respectively) but very different levels of expenditures (the expenditure in Norway was roughly USD 124 000).” (OECD Country Note: http://www.oecd.org/pisa).}
\]
of the dependent variable shows constant variability in the sample, the regression coefficients do not change across quantiles—with the sole exception of the intercept—for any pair of quantiles \( \theta_j, \theta_k \) it is \( \beta(\theta_j) = \beta(\theta_k) \).

The intercept computes the chosen quantile of the dependent variable when all the other coefficients are set to zero like in the OLS case where the intercept computes the sample mean of the dependent variable. When the dispersion of the dependent variable conditional to the covariates is not constant in the sample, the errors are heteroskedastic. This causes the regression coefficients estimated at a given quantile, \( \beta(\theta) \), to change from one quantile to another: looking at two different quantiles \( \theta_j, \theta_k \) the estimated coefficients will differ, \( \beta(\theta_j) \neq \beta(\theta_k) \) (Koenker, 2005). This applies to student scores: their dispersion changes and this results in regression coefficients that change across quantiles.

In addition, we are interested in checking if the model changes over time. Consider two time periods, past and recent, indexed respectively as 0 and 1. Having estimated the model in the two different subsets, 0 and 1, the changes from one group to the other can be computed by

$$\hat{y}_0(\theta) - \hat{y}_1(\theta) = x_0 \hat{\beta}_0(\theta) - x_1 \hat{\beta}_1(\theta)$$

Next, by adding and subtracting the term \( x_1 \hat{\beta}_0(\theta) \), which multiplies the group 1 covariates by the group 0 estimated coefficients, the difference can be decomposed into

$$x_0 \beta_0(\theta) - x_1 \beta_1(\theta) = x_0 \beta_0(\theta) - x_1 \beta_0(\theta) + x_1 \beta_0(\theta) - x_1 \beta_1(\theta)$$

$$= (x_0 - x_1) \beta_0(\theta) + x_1 (\beta_0(\theta) - \beta_1(\theta))$$

which shows how the difference between groups can be split into changes in the covariates \((x_0 - x_1)\)—i.e., changes explained by the independent variables—and changes in the coefficients \( \beta_0(\theta) - \beta_1(\theta) \) that are not explained by the variables but are caused by a change in the conditional distribution of \( y \). In the above decomposition, the term \( x_1 \hat{\beta}_0(\theta) \) is the counterfactual: it measures the value of \( y_1 \) if the regression coefficients remain unchanged from group 0 to group 1. It is usually estimated on average using the OLS estimates of \( \hat{\beta}_0 \) multiplied by the sample means of the covariates in group 1. In the quantile regression decomposition, the counterfactual is computed at different quantiles \( \theta \) to identify changes in the covariates and the coefficients in the center and/or the tails of the distribution.

Chernozhukov et al. (2013) discuss the conditions required for valid first stage quantile regression estimates to compute the counterfactuals. They show that bootstrap is a valid approach to estimate standard errors and perform inference about the counterfactuals; this allows to assess the statistical relevance of the quantile regression decomposition. The quantile decomposition states, at each decile, whether a discrepancy between the actual and counterfactual values is statistically relevant, and if such discrepancy is stable or changes across quantiles.

In a sample divided into groups 0 and 1, the quantile decomposition approach can be summarized as follows:
1) \( m = 100 \) values are drawn from a uniform distribution \((0,1)\); for each draw from this distribution the corresponding quantile regression is computed at the quantile \( \theta_j \) for that particular draw, \( j = 1,..,m \); 100 vectors of estimated coefficients \( \hat{\beta}_0(\theta_j) \) are provided by minimizing the quantile regression objective function \( \sum_{n_0} \{ \theta_j - 1(y_{0i} \leq x_{0i}\beta) \} (y_{0i} - x_{0i}\beta) \) in group 0; 100 estimated coefficients \( \hat{\beta}_1(\theta_j) \) are separately computed in group 1, by minimizing \( \sum_{n_1} \{ \theta_j - 1(y_{1i} \leq x_{1i}\beta) \} (y_{1i} - x_{1i}\beta) \) (Machado and Mata (2005));

2) \( m = 100 \) random samples with replacement, each of size \( n \), are drawn from the group 0 subset, yielding \( m \) random samples of \( x_0 \), from now on \( \hat{x}_0 \), to compute the distribution of the variable \( \hat{y}_{0/0} = \hat{x}_0\hat{\beta}_0(\theta_j)_{00} \) via the unconditional distribution function of \( \hat{y}_{0/0}, \hat{F}_{0/0}(y) = \sum_j (\theta_j - \theta_{j-1}) 1(\hat{y}_{0/0} \leq y) \), where \( \theta_j \) and \( \theta_{j-1} \) are two adjacent quantiles; \( m = 100 \) random samples with replacement are drawn from group 1 covariates, from now on \( \hat{x}_1 \), to compute \( \hat{y}_{1/1} = \hat{x}_1\hat{\beta}_1(\theta_j) \) and its unconditional distribution \( \hat{F}_{1/1}(y) \) (Melly 2006);

3) the counterfactual can be computed by \( \hat{y}_{0/1} = \hat{x}_1\hat{\beta}_0(\theta_j) \) and its unconditional distribution function by \( \hat{F}_{0/1}(y) = \sum_j (\theta_j - \theta_{j-1}) 1(\hat{y}_{0/1} \leq y) \)

4) the difference in the covariates \( \hat{F}_{0/0}(y) - \hat{F}_{0/1}(y) \), and the difference in the coefficients \( \hat{F}_{0/1}(y) - \hat{F}_{1/1}(y) \), can now be estimated at various quantiles.

5) a bootstrap approach allows to compute standard errors and to implement inference.

To estimate the decomposition at several quantiles, the comparison between \( y_0 \) and \( y_1 \) becomes a comparison between their unconditional distributions:

\[
\hat{F}_{0/0}(y) - \hat{F}_{1/1}(y) = [\hat{F}_{0/0}(y) - \hat{F}_{0/1}(y)] + [\hat{F}_{0/1}(y) - \hat{F}_{1/1}(y)]
\]

The results of the quantile regression decomposition, which analyzes the difference between the observed and counterfactual distributions at various quantiles, are not revealed by a standard average decomposition analysis, the Oaxaca-Blinder decomposition. For instance, quantile regression decomposition may disclose the case of discrepancies at low quantiles attributable to the covariates, i.e., explained by changes in the covariates, while at the median and top quantiles a divergence is linked to the coefficients, i.e., is not explained by the analyzed regression model. In sum, the source of discrepancy in the covariates and/or the coefficients may differ across quantiles. It may also be the case that an increasing effect of the covariates across quantiles counterbalances a decreasing coefficients effect, thus providing a stable overall discrepancy. The two effects may cancel out at some/all quantiles and the analysis on average would not uncover their pattern.

3 The estimated model

The average Italian score was below average until 2006, and raised above the OECD average in 2009, scoring 495 versus an OECD average of 491 (Fig. 5 in the Appendix reports the OECD graphs).
According to the OECD country note, ‘Italy’s mean performance improved between 2003 and 2012 by an average of 20 score points, moving substantially closer to the OECD average. Most of the improvement in mathematics performance was observed between 2006 and 2009. Italy is one of the fastest improving countries in mathematics performance among those countries that participated in every PISA assessment since 2003 (OECD Country Note 2012, pg.1). Italy shows above-OECD-average equity in education outcomes, with 10% of the variation in student performance in mathematics attributable to differences in student socioeconomic status (ibidem, pg. 5) While performance improved, equity remained stable. The improvement in mathematics performance is observed among all socio-economic groups: disadvantaged students improved by 27 scorepoints and advantaged students by 17 score points (ibidem, pg. 6).’

In our model the math score of Italian students, \( y \), is the dependent variable and is explained by variables describing school characteristics: field, school size, funding, student–teacher ratio, library facilities, number of computers, lab equipment, textbooks, and so forth.

The model is estimated at the center and in the tails of the conditional distribution, and analyzes median, low, and high proficiency students.

The math test scores, \( y \), is related to: funding, from now on private, that is a dummy variable assuming unit value for privately funded schools; school track, split into academic and technical, two dummy variables to single out vocational fields; a gender dummy which takes a unit value for boys; school size, reporting the number of enrolled students; number of computers in the school; student–teacher ratio; proportion of certified teachers; shortage of teachers; four categorical variables for poor library facilities, shortage of computers, shortage of textbooks, shortage of lab equipment. These four variables assume the following categories: a lot, some, a little, none at all. Numerical values ranging from 4 (a lot) to 1 (none at all) are assigned to the categories of this group of variables, as in Likert (1932). Finally, our model specification is completed by adding students’ absenteeism, to proxy school environment, and father and mother education to control for student’s family socio-economic variables.

The sample comprises \( n=52,482 \) observations and in Table 1 summary statistics for the dependent variable in each year are reported.5 In the table the mean reaches a peak in 2003,6 and then in 2009, the median is larger than the mean, signaling skewness. Over time, comparing the scores in 2000 and 2009, the 10th quantile increases sizably more than the 90th: \( \gamma(0.10)_{2009} - \gamma(0.10)_{2000} = \Delta y(0.10)_{2009–00} = 39.88 \), \( \gamma(0.50)_{2009} - \gamma(0.50)_{2000} = \Delta y(0.50)_{2009–00} = 22.78 \), \( \gamma(0.90)_{2009} - \gamma(0.90)_{2000} = \Delta y(0.90)_{2009–00} = 23.0 \)

---

5 The sample size and summary statistics of the analysis, reported in Table 1, do not necessarily coincide with their OECD analogues due to missing values in some of the variables in the selected model.
6 In 2003 the score is large, but the sample size is less than half of 2009 sample. Furthermore, the sample across regions in 2003 is less balanced than in the following waves. In a total of 11639 questionnaire the northern regions were represented by more than half of the questionnaire, \( n_{north} = 6942 \), about one fourth for the south, \( n_{south} = 3188 \), and even less for the central regions, \( n_{center} = 1509 \). The northern regions generally yield better scores, and this may have caused an overrated nationwide average. Indeed, a southern dummy in any year of this model has always a sizable negative sign, as can be seen in Table 8 of the appendix. In year 2009 the regional distribution is decidedly more balanced, with \( n_{south} = 10551 \) and \( n_{north, center} = 20354 \).
These values state that lower scoring students show a greater improvement over time, and this is in line with the OECD Country Note stating an above average equity in education in Italy.

Figure 1 plots the nationwide estimated kernel densities of the frequencies of the math scores, for the period 2000–03-06, \( y_0 \), and for year 2009, \( y_1 \). The graph shows a clear right shift in \( y_1 \), the 2009 kernel density, with respect to the density in the previous years. The average math score is 488 in 2000–03-06, and 495 in 2009.

Next, we implement quantile regression analysis at the center of the conditional distribution as computed at the median, and in the tails, at the first and third quartile. Estimating the quantile regression in the tails is particularly useful in case of asymmetric behavior of low and high scoring students since it allows to analyze the impact of school resources at the various proficiency levels. In Table 2, the columns

| Year | Sample size | Mean | Std. dev | Median | 10th | 25th | 75th | 90th | OECD mean |
|------|-------------|------|----------|--------|------|------|------|------|-----------|
| 2000 | 1940        | 470.85 | 89.54   | 474.34 | 344.30 | 418.34 | 531.99 | 581.15 | 500       |
| 2003 | 10,001      | 499.51 | 87.29   | 502.29 | 384.90 | 440.13 | 561.09 | 611.80 | 500       |
| 2006 | 16,973      | 478.84 | 90.71   | 479.58 | 361.49 | 416.64 | 541.89 | 595.10 | 497       |
| 2009 | 23,568      | 495.09 | 84.42   | 497.12 | 384.18 | 436.95 | 555.39 | 604.23 | 491       |
| All  | 52,482      | 488.64 | 88.08   | 490.74 | 373.57 | 428.72 | 550.54 | 601.65 |           |
| \( \Delta_{2009-00} \) |          | 24.24  | 22.78   | 39.88  | 18.61  | 23.40  | 23.08  | −9     |           |

**Fig. 1** Kernel density of \( y_0 \) and \( y_1 \). The frequencies of math scores are depicted by the dotted line for the period 2000–2006, and by the thin dash-dotted line for year 2009. The latter curve presents a right shift and higher peak.
Table 2 Quartile regression estimates, $\theta = .25, .50, .75$

|                      | Coef | St.err | Coef | St.err | Coef | St.err | Coef | St.err |
|----------------------|------|--------|------|--------|------|--------|------|--------|
|                      | 2000–03–06 | 2009 | 2000–03–06 | 2009 | 2000–03–06 | 2009 | 2000–03–06 | 2009 |
| Academic             | 82.34 | .69    | 68.35 | .62    | 86.16 | .57    | 66.30 | .51    | 85.92 | .54    | 71.23 | .60    |
| Technical            | 40.49 | .56    | 50.36 | .46    | 46.95 | .50    | 48.73 | .44    | 47.10 | .53    | 48.76 | .51    |
| Stud. teach. ratio   | 1.61  | .04    | 5.03  | .23    | 1.25  | .04    | 5.91  | .08    | 1.16  | .03    | 4.22  | .09    |
| Short. lab equip     | −4.68 | .27    | −1.31 | .21    | −5.13 | .21    | −1.56 | .18    | −7.85 | .23    | −3.1  | .18    |
| Shortage library     | 1.99  | .25    | −3.23 | .23    | 1.42  | .22    | −2.53 | .19    | 1.86  | .26    | −2.16 | .22    |
| Boys                 | 23.86 | .40    | 31.02 | .35    | 24.69 | .35    | 30.18 | .29    | 24.34 | .39    | 29.24 | .34    |
| Teacher shortage     | 3.08  | .24    | 5.03  | .23    | 3.79  | .17    | 5.98  | .18    | 4.18  | .19    | 6.07  | .21    |
| Shortage computers   | 8.73  | .31    | .88   | .26    | 9.16  | .21    | −31   | .21    | 8.08  | .27    | −56   | .24    |
| Private              | −48.65| .65    | −70.66| .81    | −28.23| .65    | −60.09| 1.40   | −29.12| .58    | −47.83| 1.11   |
| Prop. cert. teachers | −14.39| .86    | −2.35 | 1.00   | 1.78  | .76    | −8.50 | .84    | 5.31  | 1.01   | −7.86 | .85    |
| School size          | −.008 | .0006  | −.066 | .001   | −.009 | .0006  | −.058 | .001   | −.006 | .0006  | −.048 | .001   |
| Computer             | .26   | .004   | .31   | .007   | .25   | .004   | .299  | .005   | .23   | .005   | .27   | .005   |
| Shortage textbook    | −15.04| .28    | −6.41 | .29    | −13.74| .26    | −5.51 | .27    | −12.27| .29    | −4.87 | .29    |
| Students’ absenteeism| −8.37 | .26    | −18.02| .25    | −10.43| .20    | −16.13| .25    | −9.73 | .21    | −15.88| .24    |
| Father educ          | 1.24  | .14    | −.49  | .13    | 1.03  | .12    | .24   | .08    | 2.47  | .14    | 1.80  | .10    |
| Mother educ          | .78   | .15    | 1.11  | .13    | 1.71  | .12    | 1.26  | .08    | .81   | .13    | .67   | .10    |
| Sample size          | 28,914| 23,568 | 28,914| 23,568 | 28,914| 23,568 | 28,914| 23,568 | 28,914| 23,568 | 28,914| 23,568 |

The standard errors here reported are adjusted for intra-cluster correlation, to account for correlation among students in the same school using the Parente and Santos Silva (2016) approach, and they are robust to heteroskedasticity.
1, 3 and 5 report the quartile regression estimates at $\theta=0.25$, $\theta=0.50$, $\theta=0.75$, for the pooled years 2000–03–06, and columns 2, 4 and 6 report the estimates for year 2009. All the quartile regressions standard errors are corrected for within school cluster effects, using the Parente and Santos Silva (2016) approach, and are robust to heteroskedasticity. The comparison over time shows in 2009 an improvement of technical track students, and of many school facility variables like shortage of lab equipment, computer shortage, textbook shortage. Vice versa, 2009 shows a wider

| Table 3 Quantile regression estimates pooling all years together, 2000–03–06–09, n = 52,482 |
|------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                           | 10th            | 20th            | 30th            | 40th            | 50th            | 60th            | 70th            | 80th            | 90th            |
| Academic                                 | 75.32 (0.52)    | 77.22 (0.43)    | 81.14 (0.46)    | 81.63 (0.38)    | 80.00 (0.45)    | 81.86 (0.41)    | 81.75 (0.38)    | 81.11 (0.39)    | 77.53 (0.42)    |
| Technical                                | 43.94 (0.49)    | 47.50 (0.32)    | 51.17 (0.32)    | 51.75 (0.33)    | 50.67 (0.34)    | 50.56 (0.33)    | 49.84 (0.34)    | 50.68 (0.35)    | 47.78 (0.40)    |
| Student teacher ratio                    | 1.75 (0.04)     | 1.93 (0.06)     | 1.91 (0.08)     | 2.00 (0.04)     | 2.03 (0.06)     | 1.90 (0.05)     | 2.16 (0.03)     | 1.88 (0.04)     | 2.21 (0.07)     |
| Sh. lab equip                            | −1.69 (0.18)    | −2.93 (0.17)    | −2.52 (0.13)    | −3.23 (0.15)    | −2.81 (0.15)    | −2.53 (0.14)    | −2.62 (0.15)    | −2.86 (0.16)    | −1.15 (0.14)    |
| Poor library                             | −0.67 (0.21)    | −2.07 (0.18)    | −1.15 (0.14)    | −1.70 (0.14)    | −1.80 (0.16)    | −0.60 (0.16)    | −1.26 (0.15)    | −1.41 (0.16)    | −2.13 (0.18)    |
| Boys                                     | 25.36 (0.36)    | 26.61 (0.27)    | 27.82 (0.24)    | 27.59 (0.24)    | 27.09 (0.25)    | 26.78 (0.24)    | 26.30 (0.23)    | 26.97 (0.28)    | 28.44 (0.27)    |
| Sh. teachers                             | 4.98 (0.18)     | 4.16 (0.14)     | 4.40 (0.12)     | 5.08 (0.11)     | 5.96 (0.13)     | 5.82 (0.14)     | 6.20 (0.12)     | 6.60 (0.15)     | 5.66 (0.14)     |
| Sh. computers                            | 0.51 (0.21)     | 2.69 (0.16)     | 3.97 (0.15)     | 4.45 (0.15)     | 4.01 (0.15)     | 2.85 (0.17)     | 2.66 (0.15)     | 3.17 (0.18)     | 2.94 (0.20)     |
| Private                                  | −58.68 (0.66)   | −60.07 (0.46)   | −54.22 (0.97)   | −46.13 (0.83)   | −36.96 (0.60)   | −35.97 (0.59)   | −34.57 (0.41)   | −36.47 (0.51)   | −31.26 (0.79)   |
| Proportion certified teachers            | 8.51 (0.94)     | −5.17 (0.57)    | −3.37 (0.54)    | 1.44 (0.58)     | 0.34 (0.66)     | 1.78 (0.59)     | −10 (0.51)      | −10.9 (0.71)    | 1.22 (0.64)     |
| School size                              | −0.015 (0.0007) | −0.017 (0.0007) | −0.018 (0.0005) | −0.018 (0.0005) | −0.015 (0.0005) | −0.015 (0.0005) | −0.013 (0.0005) | −0.013 (0.0005) | −0.011 (0.0007) |
| Computers                                | .15 (.003)      | .15 (.003)      | .16 (.002)      | .16 (.002)      | .16 (.002)      | .16 (.002)      | .15 (.002)      | .14 (.002)      | .11 (.003)      |
| Sh. textbooks                            | −9.89 (0.26)    | −10.13 (0.20)   | −11.32 (0.19)   | −9.91 (0.20)    | −10.20 (0.20)   | −10.33 (0.20)   | −8.87 (0.20)    | −9.26 (0.20)    | −9.60 (0.22)    |
| Students’ absenteeism                    | −14.35 (0.19)   | −13.11 (0.15)   | −13.59 (0.14)   | −12.73 (0.15)   | −12.19 (0.15)   | −12.45 (0.15)   | −12.00 (0.13)   | −11.79 (0.17)   | −9.91 (0.17)    |
| Father education                         | .007 (0.09)     | −.39 (0.09)     | −.19 (.07)      | .21 (.08)       | .67 (.08)       | 1.04 (.09)      | 1.59 (.09)      | 2.06 (.08)      | 2.53 (.10)      |
| Mother education                         | 1.19 (0.11)     | 1.41 (0.09)     | 1.16 (.09)      | 1.45 (.08)      | 1.72 (.07)      | 1.58 (.09)      | 1.08 (.09)      | .47 (.08)       | −.21 (.09)      |

The standard errors within parenthesis are robust to heteroscedasticity and are adjusted for intra–cluster correlation, to account for correlation among students in the same school.
negative impact of school size, poor library facilities, gender gap, students’ absenteeism, and private schools. The latter two variables have a larger negative impact at the lower quantiles, showing their wider detrimental effect for weaker students. The table shows changing coefficients across quantiles and points out the presence of heteroskedasticity.

Table 3 reports the various quantile regression estimates for all the years pooled together, 2000–03–06–09 results, for a larger number of quantiles considering the proficiency levels at the 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and the 90th deciles. The great majority of the estimated coefficients are statistically significant. Shortage of laboratory equipment, shortage of textbook, school size, students’ absenteeism and private schools have a negative effect on student performance throughout. The last two are particularly sizable at the lower deciles, for the weaker students. The gender gap increases with the deciles.

The left section of Table 4 reports the F test on the equality of the coefficients in year 2009 compared to the estimates of the same coefficients computed in the rest of the sample. Most estimated coefficients in year 2009 are statistically different from the estimates of the previous waves – almost all of them at all quartiles. This confirms that year 2009 is a turning point. The right section of this table considers the stability of the estimated coefficients across quantiles, looking separately at year 2009 and at the pooled years 2000–2006. The significantly varying coefficients across quantiles are very few in the entire sample and in the pooled 2000–2006 subset, while are more numerous in year 2009, signaling the presence of heteroscedasticity.

Tables 2, 3, 4 show that the impact of each explanatory variable is not homogeneous across quantiles, and that the variables for school characteristics have a different impact on the scores of high, medium, or low proficiency students.

3.1 Decomposition analysis

The changes over time in the estimated regressions, reported in Table 2, can be decomposed into changes due to coefficients, and changes explained by the covariates

$$\hat{y}_0(\theta) - \hat{y}_1(\theta) = (x_0 - x_1)\hat{\beta}_0(\theta) + x_1(\hat{\beta}_0(\theta) - \hat{\beta}_1(\theta))$$

where the comparison is between the pooled data for years 2000–03–06 which are indexed 0, and year 2009 data which are indexed 1; $x_i$ collects all the explanatory variables at time $i = 0,1$, and $x_i\hat{\beta}_0(\theta)$ is the counterfactual measuring the impact of the recent value of the covariates evaluated at the previous period estimated coefficients. A change in the covariates ($x_0 - x_1$) is a variation in math scores explained by school resources, while a change in the coefficients ($\hat{\beta}_0(\theta) - \hat{\beta}_1(\theta)$) is an unexplained change in student performance, a change unrelated to the covariates of the model. To implement the decomposition at several quantiles, the comparison between $y_0(\theta)$ and $y_1(\theta)$ becomes a comparison between their distributions:
where the index 0 is for the years 2000–03-06 and 1 is for year 2009. The first term after the equality measures the impact of the covariates since the counterfactual distribution $\hat{F}_{0}∕\hat{F}_{1}(y)$ considers the 2009 covariates at the 2000–03-06 coefficients. The second term after the equality measures the effect of the coefficients since both distributions consider the group 1 covariates for year 2009 while the coefficients differ across subsets.

The estimates of the above decomposition, computed at all deciles, are in Table 5 together with their standard errors. This table reports the estimates of the covariate effect and their standard errors at each decile in columns 1 and 2; columns 3 and 4 collect the coefficients effects at each decile together with their standard errors. In

\[
\hat{F}_{0/0}(y) - \hat{F}_{1/1}(y) = \left[\hat{F}_{0/0}(y) - \hat{F}_{0/1}(y)\right] + \left[\hat{F}_{0/1}(y) - \hat{F}_{1/1}(y)\right]
\]

where the index 0 is for the years 2000–03-06 and 1 is for year 2009. The first term after the equality measures the impact of the covariates since the counterfactual distribution $\hat{F}_{0/1}(y)$ considers the 2009 covariates at the 2000–03-06 coefficients. The second term after the equality measures the effect of the coefficients since both distributions consider the group 1 covariates for year 2009 while the coefficients differ across subsets.

The estimates of the above decomposition, computed at all deciles, are in Table 5 together with their standard errors. This table reports the estimates of the covariate effect and their standard errors at each decile in columns 1 and 2; columns 3 and 4 collect the coefficients effects at each decile together with their standard errors. In

\[
\hat{F}_{0/0}(y) - \hat{F}_{1/1}(y) = \left[\hat{F}_{0/0}(y) - \hat{F}_{0/1}(y)\right] + \left[\hat{F}_{0/1}(y) - \hat{F}_{1/1}(y)\right]
\]

where the index 0 is for the years 2000–03-06 and 1 is for year 2009. The first term after the equality measures the impact of the covariates since the counterfactual distribution $\hat{F}_{0/1}(y)$ considers the 2009 covariates at the 2000–03-06 coefficients. The second term after the equality measures the effect of the coefficients since both distributions consider the group 1 covariates for year 2009 while the coefficients differ across subsets.

The estimates of the above decomposition, computed at all deciles, are in Table 5 together with their standard errors. This table reports the estimates of the covariate effect and their standard errors at each decile in columns 1 and 2; columns 3 and 4 collect the coefficients effects at each decile together with their standard errors. In

\[
\hat{F}_{0/0}(y) - \hat{F}_{1/1}(y) = \left[\hat{F}_{0/0}(y) - \hat{F}_{0/1}(y)\right] + \left[\hat{F}_{0/1}(y) - \hat{F}_{1/1}(y)\right]
\]

where the index 0 is for the years 2000–03-06 and 1 is for year 2009. The first term after the equality measures the impact of the covariates since the counterfactual distribution $\hat{F}_{0/1}(y)$ considers the 2009 covariates at the 2000–03-06 coefficients. The second term after the equality measures the effect of the coefficients since both distributions consider the group 1 covariates for year 2009 while the coefficients differ across subsets.

The estimates of the above decomposition, computed at all deciles, are in Table 5 together with their standard errors. This table reports the estimates of the covariate effect and their standard errors at each decile in columns 1 and 2; columns 3 and 4 collect the coefficients effects at each decile together with their standard errors. In

\[
\hat{F}_{0/0}(y) - \hat{F}_{1/1}(y) = \left[\hat{F}_{0/0}(y) - \hat{F}_{0/1}(y)\right] + \left[\hat{F}_{0/1}(y) - \hat{F}_{1/1}(y)\right]
\]

where the index 0 is for the years 2000–03-06 and 1 is for year 2009. The first term after the equality measures the impact of the covariates since the counterfactual distribution $\hat{F}_{0/1}(y)$ considers the 2009 covariates at the 2000–03-06 coefficients. The second term after the equality measures the effect of the coefficients since both distributions consider the group 1 covariates for year 2009 while the coefficients differ across subsets.
3.2 Regional gap

The OECD country notes state that:

‘In Italy more than half (51.7%) of the overall variation in student performance lies between schools: this means that two students who attend different schools can be expected to perform at very different levels. The comparatively large between-school variation in performance to an extent reflects the large regional differences in performance which can be observed in Italy, although large between-school differences can be observed even when regional differences are considered’ (ibidem, pg. 7).

Figure 2 compares the score densities estimated separately for northern and southern schools in each period. The northern regions kernel densities are represented by the thick dotted lines while the thin dashed lines represent the southern regions kernel estimates. In both time periods, the northern school density

---

Table 5  Decomposition of nationwide time effects

| Quantile | Temporal effect | Covariate effect | Coefficient effect |
|----------|-----------------|------------------|-------------------|
|          |                 | Quantile effect  | Pointwise Std. Err | Quantile effect  | Pointwise Std. Err |
| .10      | -29.2733        | 2.516            | -22.5368          | 4.562            |
| .20      | -32.3516        | 1.869            | -4.04085          | 2.234            |
| .30      | -33.8339        | 1.895            | 1.69701           | 1.761            |
| .40      | -34.9772        | 1.906            | 4.31446           | 1.763            |
| .50      | -36.1518        | 1.985            | 6.48963           | 1.799            |
| .60      | -37.5834        | 2.107            | 8.46824           | 1.942            |
| .70      | -38.2257        | 2.346            | 9.7641            | 2.200            |
| .80      | -38.7321        | 2.671            | 11.3997           | 2.530            |
| .90      | -37.9164        | 3.396            | 15.1435           | 3.233            |

In bold are the statistically significant coefficients
is to the right of the southern school density but the lower tail for the southern schools is closer to the lower tail of the northern schools in year 2009, signaling an improvement in southern low scoring students. Figure 3 compares densities over time within each region. The average scores for the north-center are 508 and 505 respectively for years 2000–03-06 and 2009, while the averages drop to 444 and 463 for the same time periods in the south. Thus, the northern distributions remain relatively unchanged while the 2009 distribution for the southern schools moves up to the right, although still centered on a lower mean with respect to the north-center.

Fig. 2 Kernel density of $y_0$ and $y_1$, the math scores densities, separately estimated for the northern and the southern schools in each period. In both past and recent years, the northern regions densities, represented by the thick curves, are to the right of the southern densities, depicted by the thin dashed curves. The southern lower tail is closer to the northern schools’ lower tail in the right graph of year 2009.

Fig. 3 Temporal difference in the score distributions within each region, Kernel densities. In the southern regions the score density for year 2009, the dotted line, has moved to the right, while in the northern regions the two distributions do not change much over time.
### Table 6 Decomposition of temporal effects within each region

| Quantile | North |          | South |          |
|----------|-------|----------|-------|----------|
|          | Covariates | Coefficients | Covariates | Coefficients |
| .10      | −2.207021 0.763 | 2.204882 1.413 | −10.11611 1.343 | −19.29005 1.989 |
| .20      | −4.503515 0.611 | 2.15248 1.010 | −11.47407 1.232 | −13.14981 1.677 |
| .30      | −5.877249 0.566 | 2.672244 0.914 | −12.6342 1.119 | −8.615979 1.560 |
| .40      | −7.045082 0.561 | 3.60467 0.829 | −13.65111 1.157 | −6.062267 1.636 |
| .50      | −8.297261 0.565 | 4.816581 0.844 | −13.51421 1.209 | −4.66078 1.726 |
| .60      | −9.120223 0.609 | 6.197368 0.858 | −13.05568 1.330 | −3.867492 1.733 |
| .70      | −9.681885 0.647 | 7.381402 0.943 | −14.17515 1.420 | −2.631131 1.835 |
| .80      | −10.90264 0.728 | 9.776959 1.198 | −15.52007 1.738 | −0.9995747 2.130 |
| .90      | −11.73765 0.852 | 13.19835 1.392 | −18.71298 2.279 | 3.179051 2.791 |

In bold are the statistically significant coefficients. The four answers of the categorical variables have been grouped in two. Sizable shortage—defined by the replies “a lot” and “some”—receive a unit value, while moderate or absence of shortage—in the replies “a little” and “not at all”—receive a zero value. Fifty bootstrap replicates of the quantile decomposition routine are separately implemented within schools’ strata.
Next, we decompose the temporal changes into covariates and coefficients effects in each region. Thanks to the reduced number of observations in the relevant subsets, the following results are computed by implementing 50 bootstrap replicates of the quantile decomposition routine. In Table 6 columns 1–4 report the results for the northern region, and columns 5–8 report the results for the southern region. The northern covariates, \( \hat{F}_{0/0}(y) - \hat{F}_{0/1}(y) \) _north_, show a steady improvement across deciles for 2009 compared to the previous years but the impact of the northern coefficients has an opposite sign, thus signaling a worsening over time that deteriorates across quantiles. This result is like the country decomposition of Table 5: improvement in student covariates and worsening in the coefficients. However here the coefficients effect halves the impact of covariates at the lower deciles, and offsets it at the top ones.

Column 5 of Table 6 reports the impact of the school covariates for the southern regions, \( \hat{F}_{0/0}(y) - \hat{F}_{0/1}(y) \) _south_, which improves in 2009 more than in the northern regions. Furthermore, column 7 shows a sizable and significant improvement in the southern coefficient effect in the recent period, \( \hat{F}_{0/1}(y) - \hat{F}_{1/1}(y) \) _south_, which has a wide impact particularly at the lower deciles. The combination of covariates and coefficients effects in the southern regions shows an overall sizable improvement across deciles which essentially drives the overall nationwide improvement of Table 5.

Summarizing, the decomposition results for the north-center and southern regions are quite different. The covariates in the north-center regions, the school characteristics, show an improvement that is partially or totally curbed by the coefficients effect. The southern changes are driven by both covariates and coefficients, where the latter is quite sizable at the lower deciles and decreases to become irrelevant at the top three deciles. The coefficient effect is particularly relevant since it could mirror students’ attitude toward school. A more favorable disciplinary climate in southern schools is among the many possible explanations for this effect. The OECD country note reports that:

• ‘Between 2003 and 2012, the disciplinary climate in Italian schools improved significantly. In 2003, 39% of students reported that, in most or all lessons, the teacher has to wait a long time for students to quiet down; by 2012 that proportion had decreased to 31%. Similarly, in 2003, 42% of students reported that there is noise and disorder in most or all lessons. By 2012 this percentage had decreased to 36%’ (ibidem, pg. 7).

The different behavior in the decomposition within regions confirms that both temporal and regional discrepancies should be analyzed. This requires the introduction of two terms for changes in the coefficients—one related to time, and the other to region—and the inclusion of two terms for changes in the covariates—one for time, and one for the regional divide. For instance, to confirm whether southern schools are closing the gap, the difference between current southern student performance and past northern student scores can be decomposed as

\[ \text{The North/South gap could be modelled by a dummy, which turns out to be statistically significant, as in table A1 of the appendix. However, the regional gap affects the entire regression, and the regional decomposition is more informative than a single regional dummy.} \]
Regional gap in students’ performance at the quantiles

where $F_{0, \text{north}}$, $F_{1, \text{south}}$, $F_{0, \text{north}}/0, \text{south}$, and $F_{0, \text{south}}/1, \text{south}$ are the respective unobservable counterfactual terms for past northern coefficients at past southern covariates, and current southern covariates at past southern coefficients.

The first line in the above decomposition shows that the actual scores of southern students with respect to the northern student scores in the previous period, $F_{0, \text{north}}$ - $F_{1, \text{south}}$, coincides with the regional difference observed in the previous period, $F_{0, \text{north}}$ - $F_{0, \text{south}}$ (the corresponding densities are reported in the left graph of Fig. 2) and temporal changes in the south, $F_{0, \text{south}}$ - $F_{1, \text{south}}$ (these densities are depicted in the right graph of Fig. 3). In Fig. 4 the left graph depicts the left-hand side of the decomposition, $F_{0, \text{north}}$ - $F_{1, \text{south}}$. It shows that, when considering the combined regional and temporal impact, northern student scores are greater than southern ones: the temporal improvement achieved by southern regions is not large enough to close the regional gap.

Next, we decompose past and recent regional differences at all deciles, into coefficient and covariate effects; the results are presented in Table 7, respectively in columns 1–4, and columns 5–8.

Table 7 columns 1–4 shows that the past regional effect of the covariates is positive and statistically significant above the median. This means that in the 2000–03-06 period the characteristics of northern schools improve upon the southern ones. The past regional difference in coefficients is positive as well, very large, and statistically significant at all quantiles so that on the overall the northern regions students outperform the southern ones, as shown in the left graph of Fig. 2.
A similar decomposition, depicted in the right graph of Fig. 3, can be implemented as follows

\[
F_{1,\text{north}}(y) - F_{0,\text{south}}(y) = \left[ F_{1,\text{north}}(y) - F_{1,\text{south}}(y) \right] + \left[ F_{1,\text{south}}(y) - F_{0,\text{south}}(y) \right]
= \left[ F_{1,\text{north}}(y) - F_{1,\text{north/1, south}}(y) \right] + \left[ F_{1,\text{north/1, south}}(y) - F_{1,\text{south}}(y) \right]
+ \left[ F_{1,\text{south}}(y) - F_{1,\text{south/0, south}}(y) \right] + \left[ F_{1,\text{south/0, south}}(y) - F_{0,\text{south}}(y) \right]
\]

where \( F_{1,\text{north/1, south}}(y) \) and \( F_{1,\text{south/0, south}}(y) \) are the counterfactuals estimating respectively actual southern covariates at actual northern coefficients, and past southern covariates at actual southern coefficients. Again, the regional and temporal discrepancies are decomposed into covariate and coefficient effects. The temporal component is alike the previous decomposition since the terms are the same but are in reverse order, and the results are the same as those in Table 6 columns 5–8 but with the opposite sign.

The regional component differs from the previous decomposition and measures current regional discrepancy (see right graph in Fig. 2). Figure 2 shows that the regional divide decreases over time; indeed, the left tails of the two regional densities get closer in year 2009. However, the northern schools’ densities are always to the right of the southern ones, showing the persistence of an educational divide across the regions. The graphs in Fig. 4 confirm the existence of this gap: when the combined temporal and regional impacts are considered, the northern schools’ kernel densities are again to the right of the southern schools’ kernel densities.

\begin{table}[h]
\centering
\begin{tabular}{lllllll}
\hline
Quantile & Past regional effect & & & 2009 Regional effect & & \\
& Covariates & Pointwise & & Covariates & Pointwise & \\
& Quantile & Effect & Std. Err & Quantile & Effect & Std. Err \\
\hline
.10 & -1.10589 & .709 & & 61.80704 & 1.698 & \\
.20 & -5.80586 & .567 & & 64.05431 & 1.268 & \\
.30 & -4.930598 & .524 & & 64.30647 & 1.070 & \\
.40 & -1.755197 & .477 & & 64.82186 & .946 & \\
.50 & .40145 & .499 & & 64.36278 & .987 & \\
.60 & 1.22676 & .476 & & 63.05086 & 1.031 & \\
.70 & 1.336039 & .501 & & 62.79888 & 1.118 & \\
.80 & 1.610684 & .513 & & 60.53896 & 1.212 & \\
.90 & 1.641533 & .630 & & 57.9825 & 1.618 & \\
\hline
\end{tabular}
\caption{Decomposition of regional gaps into covariates and coefficients changes}
\end{table}
Table 7 columns 5–8 report the results of the analysis of the regional component in year 2009. The impact of the covariates is negative throughout and statistically significant but at the upper deciles. This signals that in 2009 southern schools’ covariates improved up to the 70th decile. The impact of the coefficients is positive, large, and statistically significant—i.e., northern coefficient effect is larger than the southern one. However, this table shows that the regional divide has been reduced over time: the prevalence of northern coefficients over the southern analogues in 2000–03-06 is always larger than in 2009. Thus, the regional gap has improved in year 2009 in terms of both covariates and coefficient effects, since the estimated values of the past coefficients effect are numerically larger than the 2009 values.

To summarize, in 2009 we find an improvement in the impact of school covariates on student scores over time—regardless of the quantile—in both regions. While the northern schools’ improvement is curbed by the coefficients effect, the southern schools’ improvements occur in both covariates and coefficients. However, over time the regional gap has been reduced but not at all closed.

4 Discussion and conclusion

In year 2009, the average math score of the OECD-PISA test in Italy has been particularly growing, from below to above the OECD average. The quantile regression estimates clearly identify the improved and the worsened variables in 2009. The improved variables, like technical track students’ performance together with many school facility variables—shortage of lab equipment, computer shortage, textbook shortage—can explain the jump in students’ performance while the worsened ones—school size, gender gap, private schools, absenteeism—clearly point out where intervention could further improve students’ proficiency. The determinants of this growth are then decomposed into covariate and coefficient effects not only on average but also at the quantiles. We focus on student performance linked to school characteristics, controlling for student’s family socio-economic variables like mother and father education. A change in covariates mirrors the change in the impact of school characteristics while a change in coefficients shows that this variation should be ascribed to outside variables. The analysis considers both temporal and regional effects. Indeed, in Italy the economic regional divide is echoed by the educational attainment, with southern students scoring less than students in the rest of the country.

The temporal decomposition shows steady improvements in the covariates at all quantiles, and a coefficients improvement at the lower decile. The regional decomposition shows a differing behavior. Over time, the northern regions present a sizable improvement in the covariates that grows across quantiles counterbalanced by a worsening of the coefficients effect at the top quantiles. The southern schools, instead, show an improvement in covariates reinforced by the coefficients effect, the latter particularly sizable at the lower deciles. Nationwide, the rise in student math scores in 2009 is due to a higher impact of the school covariates. In the southern school the improvements in coefficients, i.e., changes in characteristics outside schools’ endowments, reinforce the covariate effect while in the rest of the country has an opposite sign and mitigates it. The southern coefficient effect
is particularly noticeable at the lower deciles and drives the nationwide improvement for the low scoring students.

Changes in covariates are unambiguously related to school settings. Changes in coefficients are less directly interpretable. Variables such as students’ attitude to education have improved in recent years, and this can provide a possible explanation.

The combined decomposition with respect to both regional and temporal aspects shows that the regional gap is reduced, but not at all closed, in 2009.

The finding that southern students’ improvements in math scores are partly related to school characteristics leaves room for policy interventions: providing additional school resources in these regions would improve the overall performance and reduce the regional divide at least in the educational attainment (Furno, 2021). However, the wide unemployment rate of young generations in the southern regions has a negative impact on school effectiveness, and policy cannot overlook it.

Further research could focus on reading and science performances, each characterized by a different behavior over time, as shown in Figure 5 of the Appendix. The science declining pattern, both in Italy and more generally in all the OECD countries, is quite worrisome. Its decline across countries, although milder than in Italy, shows a general rather than national or local nature and calls for an accurate analysis of its determinants to define effective policies.

Appendix

The PISA test has been implemented up to 2018. The 2020 did not take place due to the Covid19 pandemic. The math scores in the latest waves do not significantly change from 2009 scores, as stated by the OECD report on Italian students’ performance: ‘Mean mathematics performance in Italy improved between 2006 and 2009, then remained stable after 2009. (PISA 2018 Italy Country Note, pg 4)’. The following graphs are from the latest 2018 OECD report of the Italian country notes (Fig. 5).

Table 8 reports the OLS estimates in each of the waves here considered, including variables describing students’ attitude toward school, like noise in class, disciplinary...
Regional gap in students’ performance at the quantiles

clime, bully, not work well. These variables could not be included in the previous analysis since their collection is not homogeneous across different waves, in some years are collected and in some other years are not. The regional dummy, south, is always negative, sizable, and significant, reaching the minimum in 2003 and slowly improving after then.

**Funding** Open access funding provided by Università degli Studi di Napoli Federico II within the CRUI-CARE Agreement. There is no funding, acknowledgment of funding is not applicable.

**Data availability** Data of the case study are available on the OECD website.

**Declarations**

**Conflict of interest** There are no conflict of Interest.

**Ethical approval** Ethical Conduct has been fulfilled.
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Barrera-Osorio F, Garcia-Moreno V, Patrinos H, Porte E, (2011) Using the Oaxaca-Blinder decomposition technique to analyze learning outcomes changes over time: an application to Indonesia’s results in PISA mathematics. The World Bank, Policy Research Working Papers, 5584: 1–23
Barro R (1997) Determinants of economic growth: a cross-country empirical study. MIT Press
Blinder A (1973) Wage discrimination: reduced form and structural estimates. J Human Res 8:436–455
Chernozhukov V, Fernandez-Val I, Melly B (2013) Inference on counterfactual distributions. Econometrica 81:2205–2268
Furno M (2021) Italian students’ performance and regional decomposition. Educ Res Policy Pract. https://doi.org/10.1007/s10671-021-09292-y
Gigena M, Vera M, Giuliodori R, Gertel H (2011) Exploring the gap difference in 2000–2009 Pisa test scores between Argentina. Chile Mexico, Regional Sectoral Eco Studies 11–3:85–96
Hanushek E (2006) Handbook of economics of education. In: Hanushek EA, Welch P (eds) School resources, vol 2., North Holland. Amsterdam, pp 866–903
Koenker R (2005) Quantile regression. Cambridge University Press
Likert R (1932) Technique for the measure of attitudes. Arch Psychol 140:5–55
Machado J, Mata J (2005) Counterfactual decomposition of changes in wage distributions using quantile regression. J Appl Economet 20:445–465
Melly B (2006) Estimation of counterfactual distributions using quantile regression, SIAW. University of St, Gallen
Oaxaca R (1973) Male-female wage differentials in urban labor markets. Int Econ Rev 14:693–709
Parente P, Santos Silva J (2016) Quantile regression with clustered data. J Economet Methods 5(1):1–15
Seta L, Pipitone V, Gentile M, Allegra M (2014) A model to explain Italian regional differences in PISA 2009 outcomes. Procedia Soc Behav Sci 143:185–189

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.