

Appendices

Appendix A. Additional tables

Table A1: Sample selection, by cohort

	Overall	Pre-Covid cohort	Covid cohort
Initial sample in Grade 13	879,786	465,774	414,012
Excluding absents from one of the tests in Grade 13	852,862	456,878	395,984
Excluding not-matched observations with sample in Grade 10	618,226	289,197	329,029
Final Sample	618,226	289,197	329,029

Source: own elaboration on INVALSI data.

Table A2: Additional descriptive statistics, by cohort

Variables	Overall mean	Pre-Covid cohort mean	Covid cohort mean
Paternal occupation			
Unemployed	0.027	0.028	0.025
Househusband	0.005	0.006	0.004
Manger/univ. professor/personnel	0.044	0.047	0.043
Entrepreneur	0.068	0.068	0.069
Freelance professional	0.176	0.176	0.176
Self-employed	0.191	0.192	0.191
Employee/teacher	0.120	0.131	0.111
Other occupation	0.238	0.240	0.236
Retired	0.020	0.020	0.020
Maternal occupation			
Unemployed	0.030	0.032	0.028
Housewife	0.278	0.285	0.272
Manger/univ. professor/personnel	0.023	0.025	0.022
Entrepreneur	0.021	0.024	0.018
Freelance professional	0.108	0.108	0.109
Self-employed	0.083	0.085	0.081
Employee/teacher	0.196	0.202	0.191
Other occupation	0.170	0.165	0.174
Retired	0.003	0.003	0.003
Geographic area			
North	0.470	0.484	0.457
Centre	0.202	0.191	0.211
South	0.328	0.325	0.332
Observations	618,226	289,197	329,029

Source: own elaboration on INVALSI data.

Table A3: Variable definition

Variable	Definition
Maths INVALSI test score G10	Score in maths INVALSI test, Grade 10 (standardised at the national level)
Maths INVALSI test score G13	Score in maths INVALSI test, Grade 13 (standardised at the national level and horizontally anchored)
Italian INVALSI test score G10	Score in Italian INVALSI test, Grade 10 (standardised at the national level)
Italian INVALSI test score G13	Score in Italian INVALSI test, Grade 13 (standardised at the national level and horizontally anchored)
Maths teachers' mark first term G10	Teachers' mark in maths, first term Grade 10 (mark that teachers assign to students at the end of the first semester, based on their overall performance during the term; it can range between 0 and 10, and 6 is the pass grade)
Italian teachers' mark first term G10	Teachers' mark in Italian, first term Grade 10 (mark that teachers assign to students at the end of the first semester, based on their overall performance during the term; it can range between 0 and 10, and 6 is the pass grade)
Covid cohort	1 if Covid cohort, 0 if pre-Covid cohort
Female	1 if female, 0 if male
Age	Age of the student
Native	1 if the student is born in Italy with at least one parent born in Italy, 0 otherwise
Migrant first generation	1 if the student is born outside Italy from non-Italian parents, 0 otherwise
Migrant second generation	1 if the student is born in Italy from non-Italian parents, 0 otherwise
Low-educated parents	1 if no parent has a tertiary degree, 0 otherwise
High-educated parents	1 if at least one parent has a tertiary degree, 0 otherwise
<i>Mother/father's occupation</i>	
Unemployed	1 if the parent is unemployed, 0 otherwise
Housewife/Househusband	1 if the parent manages the home and often raises children instead of earning money from a job, 0 otherwise
Manger/univ. professor/personnel	1 if the parent is a manager, a university professor or a university staff member, 0 otherwise
Entrepreneur	1 if the parent is an entrepreneur, 0 otherwise
Freelance professional	1 if the parent is a freelance professional, 0 otherwise
Self-employed	1 if the parent is self-employed, 0 otherwise
Employee/teacher	1 if the parent is an employee or a teacher, 0 otherwise
Other occupation	1 if the parent works in none of the mentioned occupational categories, 0 otherwise
Retired	1 if the parent is retired, 0 otherwise
<i>Geographic area</i>	
North	1 if the student lives in the North of Italy, 0 otherwise
Centre	1 if the student lives in the Centre of Italy, 0 otherwise
South	1 if the student lives in the South of Italy or in an Italian island, 0 otherwise
<i>School track</i>	
Lyceum Scientific	1 if the student is in a scientific lyceum, 0 otherwise
Lyceum Other	1 if the student is in a classical, linguistic or other lyceums, 0 otherwise
Technical	1 if the student is in a technical school, 0 otherwise
Vocational	1 if the student is in a vocational school, 0 otherwise

Note: G10 stands for Grade 10; G13 stands for Grade 13.

Table A4: Medium-term impact of Covid-19 on maths and Italian test scores, in Grade 13, overall and by school track, 2022

	Grade 13 (1)	Lyceum Scientific (2)	Lyceum Other (3)	Technical (4)	Vocational (5)
Maths					
Covid cohort	−0.366*** (0.003)	−0.406*** (0.007)	−0.334*** (0.006)	−0.391*** (0.005)	−0.288*** (0.007)
Italian					
Covid cohort	−0.424*** (0.003)	−0.441*** (0.006)	−0.403*** (0.006)	−0.465*** (0.005)	−0.332*** (0.007)
Observations	671,408	180,601	204,525	195,206	91,076
Initial Abilities	Yes	Yes	Yes	Yes	Yes
Socio-Demogr.	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes

Note: Initial abilities include maths and Italian INVALSI test scores, and teacher-assigned marks in the subject related to the assessment test (either maths or Italian) in Grade 10. Socio-demographic controls include gender, first and second generations migrant status, age, parental occupations, and high-educated parents (at least one parent has a tertiary degree). Standard errors in parentheses are clustered at the class level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Differential attrition from G10 to G13 by cohort

Region	Covid cohort	Pre-Covid cohort	Region	Covid cohort	Pre-Covid cohort
Abruzzo	0.21	0.18	Piemonte	0.28	0.21
Basilicata	0.24	0.18	PA Bolzano	0.25	0.25
Calabria	0.36	0.17	PA Trento	0.16	0.21
Campania	0.36	0.14	Puglia	0.59	0.18
Emilia-Romagna	0.24	0.22	Sardegna	0.34	0.30
Friuli-Venezia Giulia	0.24	0.21	Sicilia	0.22	0.21
Lazio	0.22	0.19	Toscana	0.24	0.23
Liguria	0.24	0.24	Umbria	0.17	0.17
Lombardia	0.24	0.22	Valle D'Aosta	0.26	0.33
Marche	0.20	0.19	Veneto	0.18	0.19
Molise	0.21	0.18			
Italia ¹	0.28	0.21			

Note: Proportion of students who participated in Grade 10 assessments and not in Grade 13 ones, separately for the Covid and the pre-Covid cohort, across Italian Regions. ¹Average at the national level.

Table A6: Probability of participating in Grade 13 assessments given Grade 10 participation

	Pre-Covid cohort (1)	Covid cohort (2)	Pre-Covid cohort (3)	Covid cohort (4)
Italian INVALSI test score G10	0.057*** (0.001)	0.069*** (0.001)	0.039*** (0.001)	0.051*** (0.001)
Maths INVALSI test score G10	0.037*** (0.001)	0.052*** (0.001)	0.045*** (0.001)	0.060*** (0.001)
Female			0.085*** (0.001)	0.074*** (0.001)
High-educated parents			0.029*** (0.001)	0.029*** (0.001)
Migrant first generation			−0.206*** (0.003)	−0.119*** (0.003)
Migrant second generation			−0.113*** (0.003)	−0.074*** (0.003)
Constant	0.819*** (0.001)	0.743*** (0.001)	0.786*** (0.001)	0.708*** (0.001)
Obs.	363,025	458,506	363,025	458,506

Note: Estimation of the probability of participating in Grade 13 INVALSI assessment tests in maths and in Italian, given participation in Grade 10, using a linear probability model. The sample is composed of the population of students who undertook the national assessment in maths and Italian in Grade 10, and the dependent variable is a dummy variable equal to 1 if the student has participated in the assessment in grade 13, 0 otherwise. High-educated parents: at least one parent has a tertiary degree. Standard errors in parentheses are clustered at the class level. G10 stands for Grade 10. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B

B.1. The extended difference-in-differences

In the following, we show that equation (2) in the main text can be seen as an extended or generalised version of a difference-in-differences model, when γ is not necessary equal to 1. Let us start from the classical difference-in-differences model:

$$Y_{Gikj} = b + \theta C_k + \alpha_0 G + \alpha_1 C_k * G + \lambda_G X_{ikj} + \delta_{Gj} + e_{Gikj} \quad (7)$$

where G takes value 1 if test scores are measured in Grade 13 and value 0 if they are measured in Grade 10 (it corresponds to “Post” in the usual DID formulation); as above C takes value 1 if the child is in the Covid cohort and 0 otherwise (it corresponds to “Treated” in the usual DID formulation). In a common difference-in-differences, the causal parameter of interest is α_1 .

We can derive equation (7) for Grade 13 ($G = 1$) and Grade 10 ($G = 0$). For $G = 1$:

$$Y_{1ikj} = (\alpha_0 + b) + (\alpha_1 + \theta)C_k + \lambda_1 X_{ikj} + \delta_{1j} + e_{1ikj} \quad (8)$$

For $G = 0$:

$$Y_{0ikj} = b + \theta C_k + \lambda_0 X_{ikj} + \delta_{0j} + e_{0ikj} \quad (9)$$

And the difference between the two test scores is:

$$Y_{1ikj} - Y_{0ikj} = \alpha_0 + \alpha_1 C_k + \lambda X_{ikj} + \delta_j + e_{ikj} \quad (10)$$

where $\lambda = \lambda_1 - \lambda_0$, $\delta_j = \delta_{1j} - \delta_{0j}$, and $e_{ikj} = e_{1ikj} - e_{0ikj}$.

Equation (10) is a special case of equation (2) (in the main text) with $\gamma = 1$. As discussed above, given that Y_1 (test scores in Grade 13) and Y_0 (test scores in Grade 10) are not vertically equated, equation (2) is more appropriate because it does not make untestable assumptions on the relation between Y_1 and $Y : 0$.

B.2. Conditional and unconditional changes in inequalities

In the following, we discuss in details the possible sources of changing inequalities due to Covid-19 and how to assess changes in both a conditional (net) and unconditional (gross) perspective.

Consider one single cohort. The average distance between achievements across social groups (assuming only one binary explanatory variable for simplicity) can be decomposed into three components:

$$\begin{aligned} E(Z_{1ij}|X = 1) - E(Z_{1ij}|X = 0) &= \\ &= \lambda' + \gamma'[E(Z_{0ij}|X = 1) - E(Z_{0ij}|X = 0)] + [E(\delta'_j|X = 1) - E(\delta'_j|X = 0)] \end{aligned} \quad (11)$$

The first component captures ‘new’ social inequalities that developed between time 0 and time 1 between children with the same prior abilities and in similar schools (λ'); the second captures carry-over effects of prior achievement gaps ($\gamma'[E(Z_{0ij}|X = 1) - E(Z_{0ij}|X = 0)]$); the third is related to possible differences in the average quality of schools attended by children in different social groups ($[E(\delta'_j|X = 1) - E(\delta'_j|X = 0)]$).

Hence, the coefficients of the interaction terms in equation (6) capture the gross gain (or loss) of different social groups relative to each other that occurred in the pandemic years, which could be attributed to one of the following mechanisms: differences in ‘new’ gaps developed between Grades 10 and 13 given prior abilities and school features; differences in carryover effects of prior ability; and differences in the value-added of the schools attended. Instead, the coefficients of the interaction terms in equation (5) capture the net gain (or loss) of different social groups relative to each other, which occurred in the pandemic years, imputable only to differences in new gaps, conditional on prior abilities and school characteristics.