



A lost generation? The impact of COVID-19 on high school students' achievements

Dalit Contini¹ · Maria Laura Di Tommaso^{1,2} · Caterina Muratori³ · Daniela Piazzalunga^{4,5} · Lucia Schiavon⁶

Received: 17 April 2024 / Accepted: 30 July 2025
© The Author(s) 2025

Abstract

This paper estimates the effect of a full year of the Covid-19 pandemic on school performance in Italy, focusing on students at the end of upper secondary school who are about to enter the labour market or start university without having had the opportunity to recover. Using longitudinal data from standardised tests for the student population nationwide, we use difference-in-differences models to analyse the performance of two cohorts of students: a cohort that has never been exposed to the pandemic and the cohort that graduated in 2021. We find that the pandemic had a substantial negative impact on students' performance in mathematics and reading (approximately 0.4 s.d. in both domains). A similar loss is observed in 2022, suggesting no recovery post-pandemic. Low-achieving students suffered the most, widening the gap between strong and poor performers. The relative position of girls compared to boys improved. Contrary to the findings of the existing international literature, inequalities by parental education remained largely unchanged.

Keywords COVID-19 · School closure · Learning loss · Standardised tests · Inequality

1 Introduction

The Covid-19 pandemic took a toll on the lives of many children and adolescents in both poor and rich countries. Children and adolescents experienced intensified poverty, increased malnutrition and mortality, worse health outcomes (stemming from strained health systems), mounting risks of violence, exploitation, and abuse (as a result of heightened household tensions) and learning loss (UNICEF 2022).

✉ Daniela Piazzalunga
daniela.piazzalunga@unitn.it

¹ University of Torino, Turin, Italy

² Collegio Carlo Alberto, Turin, Italy

³ University of Barcelona & IEB, Barcelona, Spain

⁴ University of Trento, Trento, Italy

⁵ Institute of Labor Economics (IZA), Bonn, Germany

⁶ Ca' Foscari University of Venice, Venice, Italy

Several empirical studies have aimed to quantify the effect of the pandemic on school learning, mostly focusing on children in primary and lower secondary school (see, among others, Engzell et al. 2021, Maldonado and De Witte 2022, Haelermans et al. 2022, Blaskó et al. 2022, Schult et al. 2022) and only a very small number of empirical studies, from Latin America, have focused on the learning loss in late adolescence (Lichand et al. 2022; Vegas 2022). Our study aims to fill this gap in the literature by estimating the impact of Covid-19 on the learning loss of high school students in Italy.

Since the beginning of the pandemic, in spring 2020, many countries imposed total school closures for weeks, sometimes months (UNESCO 2023). In this context, loss of learning could have occurred through several channels. First, the pandemic could have impacted learning achievements by weakening the relationships and cooperation with classmates, causing concentration difficulties, socio-emotional loss and mental health problems, due to isolation and social distancing. The severe restrictions imposed during lockdowns and school closures led to an enormous change in youngsters' social environment, resulting in feelings of social isolation that affected mental health and socio-emotional development (Banko-Ferran et al. 2023; Mazrekaj and De Witte 2024). Medical research has shown that the prevalence of clinically elevated symptoms of depression and anxiety increased as the pandemic progressed and was higher in older children (Racine et al. 2021; Sandner et al. 2023). Second, the pandemic could have increased absenteeism of children and teachers due to contagion. Third, distance learning has been found to be less effective than school-based teaching (Kofoed et al. 2024; Jack et al. 2023). Additionally, students may have faced difficulties in accessing distance learning and/or received insufficient parental support (Blaskó et al. 2022). This last channel is more likely to occur among disadvantaged social groups, thereby exacerbating inequalities. Disadvantaged households have less access to technology and internet (Almeida et al. 2022), and parents from these families may have been less able to help their children with their schoolwork because of limited educational knowledge, constraints in availability of time off work, and limited financial resources to provide private tutoring (Maldonado et al. 2022; Mazrekaj and De Witte 2024). On the contrary, the negative effect of the pandemic on learning achievements might have been attenuated by the mental health benefits of remote learning for some students (Björkegren et al. 2024) partly due to reduced in-person bullying victimisation (Bacher-Hicks et al. 2022).

The existing literature highlights a sharp decline in school learning due to the pandemic, with substantial variation in the size of the loss across countries, age groups, and containment measures (Bethhäuser et al. 2023; Patrinos et al. 2023; Jack and Oster 2023). Nevertheless, there is a general consensus that learning losses were greatest where schools remained closed the longest and that the pandemic exacerbated educational inequalities by socio-economic background. Additionally, most studies find that low-performing students suffered the most significant losses.

We focus on the case of Italy, one of the countries that experienced the longest school closures (UNESCO 2023), with younger students prioritised in the re-opening of schools in 2020/2021, and characterised by low levels of literacy and numeracy among the population (OECD 2023). So far, however, also for Italy the existing research on the learning loss experienced during the pandemic has focused on primary and lower secondary school (Contini et al. 2022; Borgonovi and Ferrara 2023; Battisti and Maggio 2023; Carlana et al. 2023).¹

We contribute to the existing literature by analysing the learning loss suffered by students affected by the pandemic at the end of high school, a level of schooling for which there is

¹ Battisti and Maggio (2023) also include high school, but estimate models with pooled data from primary and secondary schools together, without distinguishing the effect of the pandemic by school level.

still little research, despite being a crucial time in the life of young people. From a policy perspective, this is of particular interest because these students are about to enter the labour market or embark on a university career without having had the opportunity to recover. Without downplaying the extreme importance of early childhood development and the risk that younger children are more impaired due to the cumulative nature of human capital acquisition (e.g., Fuchs-Schündeln et al. 2022), children in the early grades do have several years of schooling ahead of them to make up for learning deficits if appropriate remedial policies are put in place. Additionally, students who were in their final year of school in 2021 did not benefit from the unprecedented stimulus package implemented by the European Union to support the recovery in the aftermath of the pandemic.² Therefore these students might suffer the long-term effects of learning loss on high school completion (Neidhöfer et al. 2021), at university, and in the labour market (Hampf et al. 2017).

We also make a methodological contribution to the existing literature. We analyse when classical difference-in-differences (DID) models can be used and when extended DID models are needed, depending on the nature of the test scores at different assessments (whether they are measured on different scales or not) and the underlying model of child growth. These issues, typically recognised in the educational literature, are generally not recognised in the economic literature.

In this study, we use difference-in-differences techniques to examine achievements in reading and maths and a rich panel database on students' learning, covering the full population of students at the national level (INVALSI data). In Italy, the national standardised assessment was suspended in 2020 due to the pandemic. Thus, we compare the results in Grade 13 (2021) of the cohort of students hit by the pandemic the previous year with those of the cohort of students attending the same grade two years before (2019), controlling for previous achievements in Grade 10. Controlling for previous achievements is fundamental because initial skills can vary among cohorts for reasons not related to the pandemic (Werner and Woessmann 2023). To analyse whether there was a recovery after the pandemic, we replicate the analyses by focusing on students in Grade 13 in 2022, who attended in-person teaching for the entire final school year, but were exposed to the pandemic and school closure in the previous years.

In addition to estimating the average effect of the pandemic, we also analyse how educational inequalities develop in relation to prior performance, gender, parental education, immigrant background, and geographic area. To address the fact that in Italy, as in many other national and international assessments, tests are not horizontally equated, we propose and estimate an empirical model in which the dependent variable is represented by test scores that are standardised within each assessment, administered to a given cohort and grade. Measuring learning loss and disparities among children from different backgrounds is crucial because a significant reduction in skill acquisition and the widening of social gaps can have major negative repercussions on a country's social and economic development (UNDP 2020; Hanushek and Woessmann 2020; Fuchs-Schündeln et al. 2022).

Our results reveal that students at the end of high school in 2021 suffered huge learning losses during the pandemic, about 0.4 standard deviations in both mathematics and reading. On average, each week of school closure results in a loss of -0.013 s.d. in both mathematics and in Italian (comparable to results derived in the meta-analyses by Betthäuser et al. (2023); Patrinos et al. (2023)). We find a learning loss of a similar magnitude when we focus on

² The package is called Next Generation EU and includes, among others, a budget for school renovation and dedicated projects.

students who were in Grade 13 in 2022, suggesting no recovery even after the end of the pandemic.

The analysis also shows that low-achieving students suffered the most. Boys lost ground to girls both in Italian (where girls were already doing better, meaning that the gap widened) and, to some extent, in mathematics (where girls typically do worse, narrowing the gap in favour of boys). When comparing students with similar performance at Grade 10, the disadvantage between migrant and native students and between southern and northern students decreased. However, because of the pre-existing gap in favour of native and northern students, and the fact that low-achieving students lost the most, overall inequalities between these groups (i.e., not controlling for previous achievement) increased. In contrast, there is no evidence of a widening of achievement gaps related to parental education, a surprising result given that many other countries have experienced growing inequalities along this dimension.

The structure of the paper is as follows. Section 2 presents the Italian schooling system and details of the Italian school closure during the pandemic. Section 3 describes the data, sample and cohorts utilised in the analyses. Section 4 deals with the empirical strategy and addresses the issue of test scores not measured on the same scale. Section 5 illustrates the results. Section 6 concludes.

2 The Italian context

2.1 The schooling system

In Italy, the school year starts in early September and finishes in mid-June in all grades. The primary and lower secondary school systems are compulsory, comprehensive and free of charge. At the end of lower secondary school, in Grade 8, students take a national exam and choose among several different types of upper secondary schools that last 5 years (Grades 9–13).³ Alternatively, at the end of lower secondary school, students can choose three-year regional vocational education and training. Since compulsory education lasts a total of ten years, up to age 16, it ideally includes (for students who have not repeated school years) the first two years of upper secondary school or vocational training.

Upper secondary schools can be broadly grouped into general (lyceums), technical and vocational tracks. More specifically, general programmes include traditional lyceums – the most academic-oriented options, divided into the humanistic lyceum (classical) and the scientific lyceum – and other lyceums, which include schools with an emphasis on foreign languages, social sciences and arts. The aim of lyceums is to give students a strong background to pursue higher education and to prepare them in terms of competences, methodological and substantive knowledge, and critical thinking skills (Eurydice 2023). Technical schools combine general and technical education, aimed at providing students with a strong background in technological and/or economic subjects and preparing them for skilled technical or administrative professions. Vocational schools provide students with a vocational background to access a variety of low-skilled occupations and deliver both three- and five-year programmes. Upon completion of any five-year high school programme and passing of a national exam, students are awarded a high school diploma that grants them access to college without proficiency requirements. Despite the formal openness of the system, the likelihood

³ Students also receive non-binding recommendations by their teachers during the final year of lower secondary school.

of enrolling in higher education (and even more so, the likelihood of earning a college degree) varies widely across school types (Contini and Salza 2020).

To monitor children's skills throughout their schooling careers, the National Institute for the Evaluation of the School System (INVALSI) administers Italian reading and maths standardised tests at different grades, from primary school to the end of high school. In high school, students enrolled in all tracks take these tests in Grade 10 and Grade 13, as described in Section 3.

2.2 The Covid-19 pandemic and school closure

Italy was the first Western country to impose strict social restrictions due to the wide-spread outbreak of Covid-19. During the first wave of the pandemic, in the spring of 2020, schools were closed nationwide for about 15 weeks, from the end of February until the end of the school year in mid-June. Wherever possible, face-to-face teaching was replaced by distance learning, leaving teachers, students and schools largely unprepared and struggling to cope. In the same school year, the Italian government suspended the possibility of applying grade retention – the practice of holding back low-achieving students to repeat a school year – which is common in Italy, especially in high schools (Salza 2022).⁴

Due to the new spread of Covid-19, school closures were again ordered at the beginning of the new school year. In practice, in the school year 2020/2021, schools were closed intermittently, with alternating periods of full closure, full opening, and limited closure in regions with high prevalence of infection (Camera dei Deputati 2022). Class-level closures were also based on the occurrence of cases in each class/school. Priority was given to opening primary and lower secondary schools, while high schools were closed for longer periods.⁵ When schools were closed, the replacement of face-to-face teaching with distance learning was mandatory, although the actual implementation of distance learning was very uneven across schools. When high schools were open, to ensure social distancing, only 50–75% of students could attend face-to-face lessons, which they attended in turn. In the school year 2021/2022, schools of all grades were again open and face-to-face teaching was resumed, with few derogations related to exceptional circumstances.

The pandemic hit the different regions with different severity. Moreover, although the general rules on school closure were set out in national guidelines, in school year 2020/2021 regional authorities were allowed to impose stricter measures. This led to considerable variation in school closures across the country, linked to the severity of the pandemic but also to political decisions and the sensitivities of local governors, idiosyncratic motivations and preferences. Figure 1 summarises the total weeks of school closure over school years 2019/2020 and 2020/2021 across the Italian region, which range from 23.4 (Trentino) up to 37.4 (Puglia). Puglia is the region with the longest period of school closure, almost two months longer than the Italian average.⁶

⁴ The empirical implications for our results of this policy change are discussed in Section 5.4 (Robustness checks).

⁵ The decision was based on the assumption that older students would be less harmed by distance learning and that they did not require parents to be present at home.

⁶ The total number of weeks of school closure are calculated as the sum of the weeks of school closure in 2019/2020 and 2020/2021. In 2019/2020, schools were closed at national level for about 15 weeks (with minor differences between regions according to the regional school calendars). In 2020/2021, school closures were decided both at the national and at the regional level according to the spread of the contagion and to the political choices of the regional authorities.

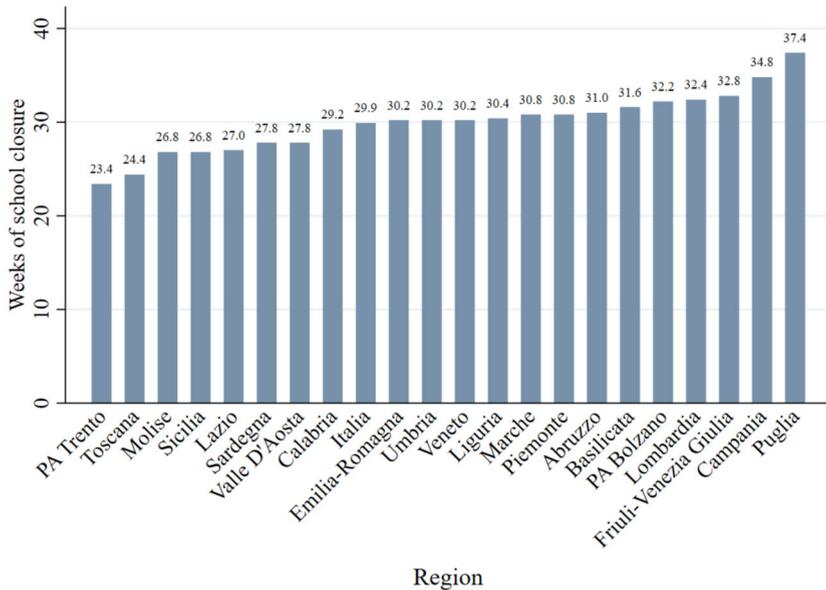


Fig. 1 Length of school closure in 2019/20 and 2020/21 school years across Italian regions. Note: Total weeks of school closure across Italian regions for the 2019/20 and 2020/21 school years. For the former school year, the weeks of closure have been measured using the regional planned school calendars provided by the Ministry of Education and Research

When compared to other Western countries, Italy presents some challenges: (1) even before the pandemic, adult literacy and numeracy levels were well below the average of OECD countries participating in the Survey of Adult Skills (PIAAC); (2) the proportion of young people with tertiary education is among the lowest in Europe, while the proportion of NEETs⁷ is among the highest (Education GPS 2023; Eurostat 2023); (3) compared to other countries, Italy lacked digital skills and adequate infrastructure for distance learning as a substitute for face-to-face teaching; (4) prior to the pandemic, Italy had one of the lowest scores in the Digital Economy and Society Index (DESI) in the European Union, one of the lowest proportions of households with a fixed broadband subscription and one of the lowest proportions of individuals with at least basic software skills (European Commission 2020); (5) Italian teachers tend to have low levels of ICT skills and little experience with blended and technology-enhanced teaching (OECD 2018; European Schoolnet 2012; Bertoletti et al. 2023); (6) Italy has one of the highest shares of children lacking individual and school learning resources among European Union countries (Blaskó et al. 2022).

In terms of counter-measures, as early as March 2020, schools received funding to improve digital tools for distance learning and technical support (Camera dei Deputati 2022). While this measure had positive effects in terms of the speed of adaptation, it also shows how unprepared schools were at the time. A budget was allocated to provide free digital equipment (PCs, tables, internet connection) to students from low socio-economic backgrounds. In the summer of 2020, a specific budget was allocated for the renovation of school buildings – to ensure physical distance in classrooms and school during the school year – and for school

⁷ Young adults not in employment, formal education or training.

staff to reduce the disruption caused by teacher contagion. In terms of remedial measures to improve student learning, no measures were taken in the summer of 2020. Instead, in the 2020/2021 school year, the state funded face-to-face teaching projects aimed at reducing learning deficits, with priority given to primary and secondary schools in disadvantaged areas. Projects were submitted by schools and then approved, with wide variations between schools in what was actually implemented. Overall, there was no uniform policy across schools, provinces and regions, and only the schools that were better equipped in terms of human resources were able to access the available funding.

3 Data and descriptive statistics

This paper exploits the data from the national standardised tests administered by INVALSI. Tests are administered to the entire population of Italian students (about 500,000 students per grade) in Grades 2, 5, 8, 10 and 13 and evaluate students' reading and maths skills.⁸ As mentioned above in Section 2, upper secondary schools in Italy can be classified into three broad tracks: general (lyceums), technical and vocational. The reading test in Grades 10 and 13 is the same across the different tracks, whereas the mathematics test has a common part and a specific part that varies between tracks.

The standardised tests in primary and lower secondary schools have been conducted in late spring every year since 2008/2009. The assessment in Grade 10 was first administered in 2011; students in Grade 13 were tested starting in 2019. Due to the pandemic, in 2020 the survey was suspended for all school stages and then administered again in 2021 and in 2022. However, the Grade 10 assessment resumed only in 2022. The tests are conducted between March and May, depending on the grade. Students in Grade 13 sit the test in March, students in lower secondary school in April, primary school children in the beginning of May, and students in Grade 10 in mid May.

An important distinction, which is often overlooked in the economic literature, is between equated and non-equated tests. In equated tests, some items appear in both assessments, allowing their "anchoring" (Bond and Lang 2013). This allows the scores of different assessments to be expressed in a common metric. The term *horizontally equated tests* refers to different assessments administered to students' from different cohorts in the same grade: when tests are horizontally equated, it is possible to make direct comparisons of the achievement levels of students who took the test in different years. The term *vertically equated tests* refers to assessments administered at different grade levels: when tests are vertically equated, it is possible to compare the results of children enrolled in different grades and to estimate achievement growth over time.

INVALSI tests have never been vertically equated. Instead, since 2019, the tests have been horizontally equated for all school grades, making it possible to express grade-specific scores in a common metric and to assess changes in results over time.⁹ These aspects have implications for the empirical analysis, discussed below in Section 4. Taking advantage of these data, this paper compares test scores of students enrolled in Grade 13 in 2020/2021 - a cohort that experienced one full year of intermittent school closure due to the pandemic - with the test scores of students enrolled in Grade 13 in 2018/2019 - a cohort that did not experience the school closure, while controlling for prior skills. The pre-Covid cohort took the INVALSI

⁸ Recently, standardised tests in English proficiency have also been introduced.

⁹ The procedure adopted by INVALSI requires that part of the items administered in the 2019 assessment are re-administered to a sub-sample of students who carried out the test in 2021.

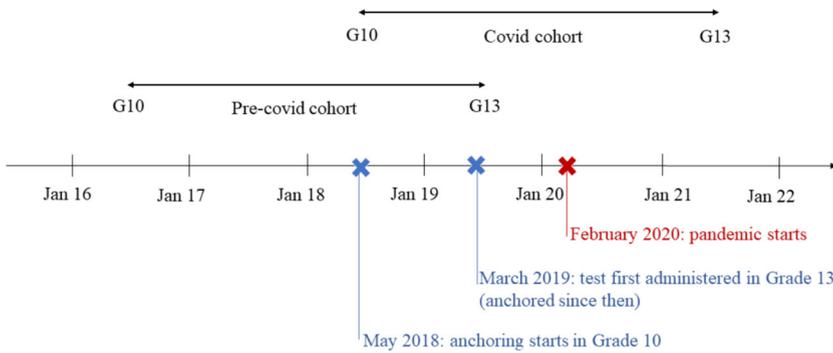


Fig. 2 Dataset structure and timeline of the anchoring of INVALSI tests

tests in spring 2019 and the Covid cohort in spring 2021. Thanks to the longitudinal nature of the survey, it is possible to link test scores in Grade 10 at the individual level. For the pre-Covid cohort, we link the dataset for Grade 13 in 2019 with the dataset for Grade 10 in 2016 and for the Covid cohort, we link the dataset for Grade 13 in 2021 with the dataset for Grade 10 in 2018 (Fig. 2).¹⁰

Given the characteristics of the linkage, the longitudinal sample consists of all students who took the tests in both Grade 10 and Grade 13, who did not repeat a school year in between (otherwise, it would not be possible to identify the same student in the Grade 10 archive three years earlier) nor dropped out of the school system. Robustness checks to account for the potential differential selection across cohorts are presented in Section 5.4.

The initial dataset recording all students in the Covid and pre-Covid cohorts who took the Grade 13 test consists of 879,786 students. A few students were excluded because they were absent from one of the two assessments (maths or Italian) in Grade 13, our outcome of interest. Others were excluded because it was not possible to match them with their prior test scores, due to absences in Grade 10, or because they experienced a grade retention in between. Longitudinal linkage has been possible for the majority of the students. Our final sample is composed of 618,226 individual observations, 289,938 in the pre-Covid cohort (47%) and 329,029 in the Covid cohort (53%) (see Table A1 in the Online Appendix for the details of the sample selection).

Table 1 reports the descriptive statistics both for the entire sample and separately for the two cohorts.¹¹ To facilitate comparability with other studies and the interpretation of the results, we rescaled test scores to have mean 0 and standard deviation 1 in the original full population. When horizontally anchored (as occurs in Grade 13), test scores are directly comparable (and standardised across cohorts). As *prima facie* evidence of a negative effect of the pandemic, we see that Grade 13 test scores are higher for the pre-Covid cohort than for the Covid cohort in both Italian and in maths.

Test scores for Grade 10 for 2016 and for 2018 are not horizontally anchored, and they are standardised by INVALSI within each cohort and not directly comparable over time. Note

¹⁰ There is some evidence of cheating in INVALSI tests (Lucifora and Tonello 2015; Bertoni et al. 2013; Angrist et al. 2017). However, this should not be an issue of major concern here. First, because the existing evidence points to lower cheating behaviour in higher grades (Lucifora and Tonello 2015). Second, we use test scores corrected by INVALSI for the risk of cheating. Third, the tests in Grade 13 - our outcome variable - are computer-based (CBT) since 2018 and correction is centralised, reducing the risk of cheating, in particular when stemming from teacher shirking (Angrist et al. 2017).

¹¹ Table 2 presents additional descriptive statistics on parental occupation and macro-area of residence.

Table 1 Descriptive statistics, overall and by cohort

Variables	Overall		Pre-Covid cohort		Covid cohort	
	mean	sd	mean	sd	mean	sd
Covid cohort	0.536					
Italian INVALSI test score G10	0.095	1.030	0.079	1.125	0.109	0.940
Maths INVALSI test score G10	0.092	1.048	0.056	1.098	0.123	1.001
Italian INVALSI test score G13	-0.082	1.032	0.130	1.013	-0.266	1.012
Maths INVALSI test score G13	0.009	1.007	0.211	0.999	-0.165	0.981
Italian teachers' mark first term G10	6.257	1.631	6.282	1.545	6.236	1.702
Maths teachers' mark first term G10	5.953	1.898	5.990	1.836	5.921	1.950
Age	18.446	0.621	18.449	0.625	18.443	0.617
Female	0.519		0.524		0.514	
Native	0.863		0.896		0.834	
Migrant first generation	0.037		0.033		0.040	
Migrant second generation	0.047		0.044		0.050	
At least one parent with university degree	0.277		0.266		0.286	
School track						
Lyceum Scientific	0.268		0.271		0.264	
Lyceum Other	0.300		0.293		0.305	
Technical	0.292		0.296		0.289	
Vocational	0.141		0.140		0.142	
Observations	618,226		289,197		329,029	

Note: G10 stands for Grade 10; G13 stands for Grade 13. Source: own elaboration on INVALSI data. Additional descriptive statistics for control variables are shown in Table A2 of the Online Appendix

that Grade 10 test scores have a mean slightly above 0 in both cohorts; this is an indication of the existence of some positive sample selection, as mentioned above.

In addition to scores in the standardised test, INVALSI collects information on teacher's marks in Italian and mathematics at the end of the first term,¹² students' socio-demographic characteristics and family background. The set of variables includes age, gender, migratory background, parents' level of education and occupation, and geographic area. All the variables used in the analysis are described in Table A3 in the Online Appendix.

4 Identification strategy

4.1 Average effects

Our starting point is a model for achievement at a given stage of schooling based on a standard education production function, with a value-added specification (Todd and Wolpin 2003):

$$Y_{1ij} = \alpha + \lambda X_{ij} + \gamma Y_{0ij} + \delta_j + \epsilon_{ij} \quad (1)$$

where Y_{1ij} is a standardised test in maths or reading set by student i in school j in a given grade (here, Grade 13); X_{ij} is a vector of time invariant controls, including the student's

¹² Marks range between 0 and 10 (6 is the pass mark), although in practice marks below 4 are extremely rare.

gender and family background (migratory background, parental education and occupation); Y_{0ij} is a vector of prior skills (standardised test scores and teacher's marks) measured at the time of the previous assessment, in Grade 10. δ_j are schools fixed effects and ϵ_{ij} are normally distributed stochastic errors.

If test scores are all measured on the same scale, the difference $Y_1 - Y_0$ represents achievement growth and may be a function of X and school fixed effects if $\gamma = 1$, or also depend on prior achievement Y_0 in the general case. If test scores at different grades are not measured on the same scale (i.e., are not *vertically* equated), the above difference does not convey any useful information. Nonetheless, model (1) is still meaningful, because parameter λ captures the extent to which sociodemographic inequalities develop between the two grades net of prior achievement (and school effects).

In this scenario, the corresponding DID model allowing to assess the average impact of the pandemic on children's learning would be:

$$Y_{1ikj} = \alpha_0 + \alpha_1 C_k + \lambda X_{ikj} + \gamma Y_{0ikj} + \delta_j + \epsilon_{ikj} \quad (2)$$

where C_k is a dummy variable equal to 1 if the child is in the Covid cohort k and 0 otherwise, and X_{ikj} and Y_{0ikj} are the explanatory variables previously defined corresponding to cohort k . α_1 is the coefficient of interest, capturing the average effect of being in the Covid cohort rather than in the pre-Covid cohort on the test scores, given previous performance and conditional on school fixed effects. The untestable identifying assumption is that, conditional on prior abilities, the performance of children in the Covid cohort would have been the same as the pre-Covid cohort had the pandemic not occurred. Such an assumption seems plausible as the two cohorts are close in time and we are conditioning on school fixed effects.

As in canonical DID, there are a treated and a control cohort and pre- and post-periods. The classical difference-in-differences model is a special case of extended DID in (2) when $\gamma = 1$ (see Online Appendix B1), which may be used only when Y_1 and Y_0 are measured on the same scale.

It is worthwhile to notice that if test scores administered at the same grade level in different years are *not* measured on the same scale (i.e. if test scores are not *horizontally* equated), the scores of different individuals in different years cannot be compared, and thus the average effects of the pandemic is not identified.

4.2 Heterogeneous effects

In addition to the average effect, we are interested in assessing how inequalities between socio-demographic groups evolved due to the pandemic. In the absence of scaling issues, one would allow coefficients and school-specific fixed effects in (2) to vary across cohorts. Naming coefficients of the pre-Covid cohort with subscript 0 and coefficients of the Covid cohort with subscript 1, we obtain the following specification:

$$Y_{1ijk} = \alpha_0 + (\alpha_1 - \alpha_0)C_k + \lambda_0 X_{ijk} + (\lambda_1 - \lambda_0)C_k X_{ijk} + \gamma_0 Y_{0ijk} + (\gamma_1 - \gamma_0)C_k Y_{0ijk} + (\delta_{jk} + \epsilon_{ijk}) \quad (3)$$

where the coefficients of interest are those of the interaction terms, capturing the extent to which the effects of individual variables and prior abilities vary before and during the pandemic. If only the constant term is allowed to vary across the two cohorts, this model boils down to (2).

Model (3) allows to analyse how inequalities across social groups evolve due to the pandemic, provided that Y_1 and Y_0 test scores are comparable across cohorts. If instead the

assessments are not horizontally equated, it may deliver biased results. In our case study, for example, while assessments in Grade 13 are equated, assessments in Grade 10 are not. Thus, Y_0 is not measured on the same scale in the two cohorts of students (Covid and pre-Covid cohort). This means that a given result in one cohort cannot be considered better or worse in absolute terms than that of another cohort. Consider the case where children's performance in Grade 10 had worsened on average between the two cohorts: the same relative position in the two cohorts would imply a lower absolute performance in the post-pandemic cohort, with the consequence of the negative impact of the pandemic being overestimated. Instead, the negative impact of the pandemic would be underestimated in the opposite scenario.

In the case test scores are not measured on the same scale, a comparison in relative terms can still be made: two children with the same score in two different cohorts may not have the same absolute performance, but they share the same relative position within their cohort distribution. The following strategy allows us to analyse how inequalities across social groups evolved during the pandemic, even in cases where some (or all) assessments are not measured on the same scale. Instead of focusing on absolute performance measures, we analyse the changes in the *relative* positions of each social group in Grade 13 before and after Covid-19 school closures, given their prior relative position.

Let us define Z_1 and Z_0 as the within-cohort standardised test scores in the two grades of interest, so that $E(Z_1) = E(Z_0) = 0$.¹³ It can be shown that if we standardise scores, single cohort models have the same structure as (1):

$$Z_{1ij} = \alpha' + \lambda' X_{ij} + \gamma' Z_{0ij} + \delta'_j + \epsilon'_{ij} \tag{4}$$

and consequently, the DiD model becomes:

$$Z_{1ijk} = \alpha'_0 + (\alpha'_1 - \alpha'_0)C_k + \lambda'_0 X_{ijk} + (\lambda'_1 - \lambda'_0)C_k X_{ijk} + \gamma'_0 Z_{0ijk} + (\gamma'_1 - \gamma'_0)C_k Z_{0ijk} + (\delta'_{jk} + \epsilon'_{ijk}) \tag{5}$$

The parameters of interest are the coefficients of the interaction terms $(\lambda'_1 - \lambda'_0)$ and $(\gamma'_1 - \gamma'_0)$ and capture the differential effects of the pandemic on the relative position in the distribution of test scores in the two cohorts: the first, by prior skills, the second by sociodemographic factors.¹⁴ The coefficient of the cohort variable has no meaningful interpretation here, as it is simply a rescaling term that ensures a 0 mean for all Z variables.

4.3 Empirical strategy in our setting

As described in Section 3, we estimate the effect of the pandemic on learning in high school using Grade 10 and Grade 13 test scores, in the Covid and the pre-Covid cohorts. The average effect is estimated using model (2). This is feasible because Grade 13 scores are measured on the same scale. However, as highlighted above, Grade 10 scores are not, meaning that Y_0 for the two cohorts are not comparable in absolute terms, but only in relative terms. In other words, what Eq. 2 does in this case is regressing same scale results relative to Grade 13 (conceivable as absolute measures of performance) on within-cohort standardised test scores in Grade 10 (conceivable as relative measures of performance). As a consequence, the

¹³ To avoid the within-cohort standardised scores to be affected by selection in the matched sample, we implemented the within-cohort standardisation in the original sample. We then performed the empirical analysis on the final analytical sample.

¹⁴ Note that Contini and Cugnata (2020) make additional arguments on the identification of the different channels responsible of changes in inequalities that do not apply to our case, because we have longitudinal data at the individual level.

model provides a realistic estimate of the average effect only if the scales employed in the two cohorts are similar, otherwise the estimation of (2) could lead to biased estimates of the impact of the pandemic.

To address this point, as a sensitivity check, we run simulations to see what happens to the estimates of the average impact of the pandemic if test scores vary across cohorts by different amounts. Given the expertise of the test designers, we consider a relatively large amount to be 0.1 standard deviations, meaning that the same score on the two Grade 10 tests could correspond to a difference in actual achievement of that size (to give a sense of what this means, according to Bloom et al. (2008), the average annual gain in reading and mathematics at Grade 10 is about 0.2 standard deviations). Thus, we add and subtract 0.1 s.d. from all individual test scores of the Covid cohort in Grade 10, estimate model (2) for each of these situations, and provide what we consider to be 'reasonable intervals' on the estimate of the average effect. We also calculate the amount of change that would correspond to a 0 effect of the pandemic on student learning, for both reading and mathematics.

With these caveats in mind, we also investigated the impact of the length of school closure and how it has influenced regional differences. First, we also estimated a version of model (2) where we include the number of weeks of closure W_{rk} (varying at the regional level and equal to 0 in the pre-Covid cohort) instead of the Covid-cohort dummy. The corresponding coefficient captures the average effect of a week of closure across the country and is approximately equal to the total effect of the pandemic divided by the average number of weeks of closure. Second, regional differences in the impact of the pandemic were investigated with a version of model (2), where being in the Covid cohort is interacted with regional dummies.¹⁵ Since the duration of school closures was defined regionally and varied significantly across regions, it was also interesting to assess whether the observed regional differences could be fully explained by the duration of school closures. From this perspective, we estimated a model that additionally includes the number of weeks of school closures W_{kr} .¹⁶ If closing weeks were entirely responsible for spatial differences, the region-specific coefficients of the Covid-cohort variable would become non-statistically significant.

To analyse heterogeneous effects, we estimated model (5) to properly account for the fact that Y_0 scores are not horizontally equated. The coefficients of the interactions between each X and the cohort variable capture how differences between groups in the relative position of the test score distributions changed before and after the pandemic, conditional on prior achievement and school characteristics. However, we are also interested in describing how overall inequalities between social groups have changed over the period of interest. To do this, we estimate a reduced form of (5) that does not include prior ability relative position or school fixed effects:¹⁷

$$Z_{ijk} = \alpha''_0 + (\alpha''_1 - \alpha''_0)C_k + \lambda''_0 X_{ijk} + (\lambda''_1 - \lambda''_0)C_k X_{ijk} + u_{ijk} \quad (6)$$

The coefficients of the interaction terms in (6) capture the gross gain (or loss) of different socio-demographic groups relative to each other that occurred during the pandemic years, which could be attributed to one of the following mechanisms: (i) 'new' X-inequalities that developed between Grades 10 and 13, given prior abilities and school features; (ii) differences

¹⁵ Building on the previous Eq. 2: $Y_{ijk} = \alpha_0 + \theta_r D_r * C_k + \lambda X_{ijk} + \gamma Y_{0ijk} + \delta_{jr} + e_{ijk}$, where θ_r are the coefficients of the interaction terms between regional dummies D_r and the Covid cohort.

¹⁶ $Y_{ijk} = \alpha_0 + \theta_r D_r * C_k + \beta W_{kr} + \lambda X_{ijk} + \gamma Y_{0ijk} + \delta_{jr} + e_{ijk}$. This model has one extra coefficient, so identification is obtained by setting one of the regions' D_r (in this case, Lombardy) to 0. The effect of the pandemic in Lombardy is represented by β_1 times the number of weeks of school closures in Lombardy. The remaining $\theta_{r,s}$ represent the additional effect in region r that is not captured by $\beta_1 W_r$.

¹⁷ In this case, we include dummies to control for school track.

related to carry-over effects of prior ability; and (iii) differences in the value-added of the schools attended. Instead, the coefficients of the interaction terms involving X variables in (5) capture the net changes attributable to channel (i) only (see Online Appendix B.2 for a formal discussion). For the sake of completeness, we also estimate a model without Z_0 but with school fixed effects, where the coefficients of the interaction terms involving X variables capture the net changes imputable to channels (i) and (ii), given school characteristics. While the overall changes cannot be definitively attributed to the pandemic, given that the prior ability distribution and school fixed effects may have changed for reasons unrelated to the pandemic itself, we can nevertheless identify some potential pandemic-related mechanisms that may have contributed to these changes. The relative changes in the value added of schools may be attributed to differences in the resources available to meet the challenges associated with school closures, including management and teacher quality, as well as ICT skills. The carry-over effects of prior skills may differ due to varying levels of attachment to school, resilience to unexpected challenges, and ICT proficiency among higher-performing students. Differences in relative learning between socio-demographic groups, net of prior achievement and school effects, could be attributed to average differences in individual resilience or to differences in parental support.

5 Results

5.1 Average learning loss

Table 2 reports the average learning loss related to the Covid-19 pandemic on students' performance in maths and in Italian for all students in Grade 13 and by school track. These figures derive from the estimation of Eq. 2, including all the available sets of controls at the individual level and school fixed effects.¹⁸

Overall, high school students suffered an average loss of 0.39 standard deviations in mathematics and 0.41 standard deviations in Italian due to the pandemic. We observe some differences across tracks; in particular, students at Scientific high schools and Technical institutes suffer the most severe losses in maths and reading, while students at Vocational institutes suffer the least.

To account for the fact that tests in Grade 10 (initial abilities) are not equated between the Covid and pre-Covid cohorts, we perform a sensitivity check by adding and subtracting 0.1 to the test scores of the Covid cohort in Grade 10, in order to simulate a difference in the metric between cohorts. We then estimate model (2) for each of these cases, and provide 'reasonable intervals' for the estimate of the average effect. These intervals are reported below each point estimates in Table 2. In addition, we compute the amount of change that would correspond to a 0 effect of the pandemic on student learning, obtaining a value of almost 0.8 standard deviations for both reading and mathematics. This figure is really large, nearly 4 times the average annual gain at Grade 10, according to Bloom et al. (2008).

To assess the medium-term impact of the pandemic, we also repeat the analysis for students in Grade 13 in 2021/2022. In this case, we compare the test scores of students enrolled in Grade 13 in 2021/2022 with the test scores of students enrolled in Grade 13 in 2018/2019, controlling for prior knowledge in Grade 10.¹⁹ Students who completed high school in 2021/2022 attended face-to-face classes for the entire school year, but experienced intermit-

¹⁸ Full results are available from the authors upon request.

¹⁹ Descriptive statistics for this sample of students are available from the authors upon request.

Table 2 Impact of Covid-19 on maths and Italian test scores, in Grade 13 overall and by school track

	Grade 13 (1)	Lyceum Scientific (2)	Lyceum Other (3)	Technical (4)	Vocational (5)
Maths					
Covid	-0.389*** (0.004)	-0.415*** (0.007)	-0.359*** (0.006)	-0.400*** (0.006)	-0.299*** (0.007)
Intervals	[-0.338***; -0.439***]	[-0.376***; -0.455***]	[-0.312***; -0.405***]	[-0.355***; -0.445***]	[-0.255***; -0.344***]
Italian					
Covid	-0.410*** (0.004)	-0.415*** (0.007)	-0.399*** (0.007)	-0.446*** (0.006)	-0.327*** (0.007)
Intervals	[-0.358***; -0.462***]	[-0.369***; -0.460***]	[-0.350***; -0.449***]	[-0.394***; -0.499***]	[-0.276***; -0.378***]
Obs.	618,226	166,859	185,426	180,543	85,398
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). School FE: School fixed effects. Standard errors in parentheses are clustered at the class level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

tent school closures due to the pandemic in previous years from March 2020 to June 2021, when they were enrolled in Grades 11 and 12. The results, reported in Table A4 in the Online Appendix, show that the 2022 cohort suffered a loss of a similar magnitude as the 2021 Covid cohort with respect to the pre-Covid cohort (0.37 s.d. in mathematics and 0.42 s.d. in Italian). This underlines the inadequacy of a simple return to face-to-face teaching to make up for previous losses in the absence of specific policies.

As a term of comparison, in their meta-analysis (Betthäuser et al. 2023) point to a learning loss of 0.14 s.d. on average across grades and subjects. This loss persists over time during the two years following the start of the pandemic. The authors report no substantial differences between primary and secondary schools, with some studies finding greater losses for younger children and other studies finding the opposite. However, of the 42 studies included in their review, only a minority concerned upper secondary school, while most research focused on primary school and, to a lesser extent, lower secondary school.

Our estimates are also much larger than the available evidence for Italy in the lower stages of schooling, where the average learning difference is estimated between -0.13 s.d. and -0.29 in maths and +0.06 s.d. and -0.08 in reading, depending on the period covered, the grade and the estimation strategy (Contini et al. 2022; Borgonovi and Ferrara 2023). The fact that the learning loss is much larger in Grade 13 is probably due to the longer duration of school closure that high school students have been exposed to. The magnitude of the loss is so large, that students may have experienced not only a delay in their expected learning growth, but possibly even a loss of competences already acquired, with long-lasting effects on their future educational and working life.

To the best of our knowledge, only two existing studies have focused on students close to the end of upper secondary school, and they are both from middle-income countries.

Lichand et al. (2022) estimates that in 2020 in Brazil the dropout risk more than triplicated and average learning loss in maths and reading amounted to 0.32 s.d. for students in Grades 6 through 12, with some variation by grade but no distinctive difference between lower and upper secondary school. In Colombia, Vegas (2022) estimates a learning deficit of 0.2 s.d. for students in Grade 11. Worryingly, the estimated effect for Italy is even larger and thus requires urgent action for these young adults.

5.2 Length of school closure

As discussed in Section 4.3, we take advantage of the regional variation in school closure weeks to estimate the effect of one week of school closure on mathematics and reading learning. We present two specifications in Table 3. First, instead of including a dummy variable for being in the Covid or pre-Covid cohort in Eq. 2, we include a continuous variable corresponding to the number of weeks of school closure in each region, equal to 0 for all students in the pre-Covid cohort (column 1). The results show an average learning loss of 0.013 s.d. per week of school closure in both mathematics and Italian. In a second specification, we include the same continuous variable but focus only on students in the Covid cohort (column 2). This specification overcomes the issue of non-horizontal-anchoring for Grade 10. The results indicate a slightly smaller loss: 0.010 in mathematics and 0.012 in Italian.

Table 3 Impact of COVID-19 on maths and Italian test scores by weeks of school closure, Grade 13

	Final Sample (1)	Covid cohort Sample (2)
		Maths
Weeks school closure	-0.013*** (0.000)	-0.010*** (0.001)
		Italian
Weeks school closure	-0.013*** (0.000)	-0.012*** (0.001)
Obs.	618,226	329,029
Initial Abilities	Yes	Yes
Socio-Demogr.	Yes	Yes
School FE	Yes	Yes
Pre-Covid cohort	Yes	No

Note: The final sample (1) consists of all students in pre-Covid and Covid cohorts who performed both Italian and maths INVALSI assessment tests in G13 and were successfully matched with the observations with INVALSI sample in Grade 10. The Covid cohort sample (2) includes only students from the Covid cohort who performed both Italian and maths INVALSI assessment tests in G13 and were successfully matched with the observations with INVALSI sample in Grade 10. 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). School FE: School fixed effects. Standard errors in parentheses are clustered at the class level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

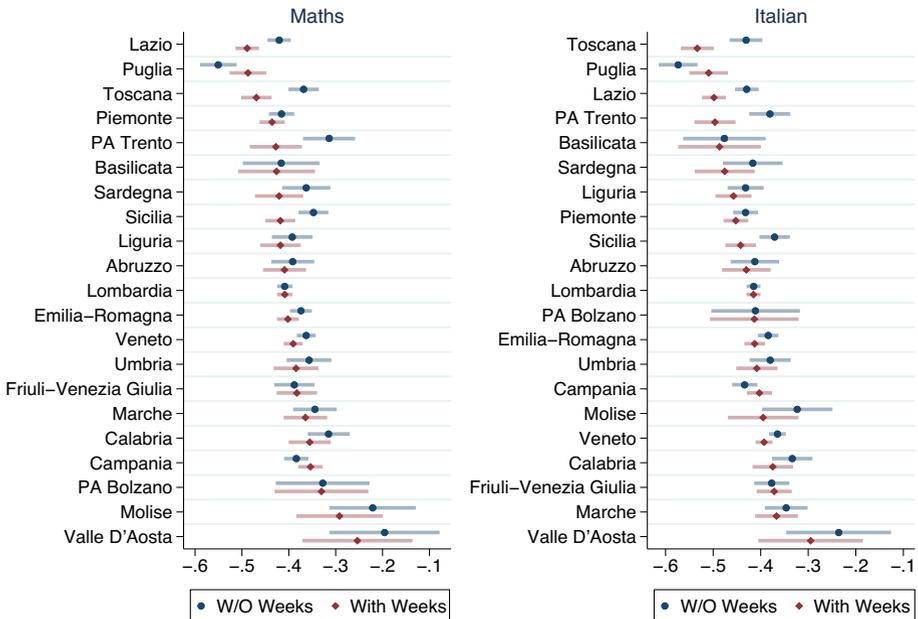


Fig. 3 Estimated effect of COVID-19 on maths and Italian test scores by region, Grade 13, with and without weeks of school closure. Note: Estimates are generated with an OLS difference-in-differences regression using the 2018/19 and 2020/21 waves of the INVALSI dataset. The estimation models control for initial abilities (maths and Italian INVALSI test scores, teacher-assigned marks in the subject related to the assessment test in Grade 10), and socio-demographic characteristics (gender, first and second generation migrant status, age, parental occupations, and high-educated parents - at least one parent has a tertiary degree). 95 percent confidence intervals are reported with horizontal lines

Figure 3 shows the heterogeneous impact of the pandemic by region, controlling and not controlling for the number of weeks of school closure. Learning losses vary significantly across regions when we do not control for school closures (blue dots and lines). Learning losses in maths vary between 0.55 s.d. (Puglia) and 0.20 s.d. (Valle d’Aosta and Molise). Reading learning losses vary between 0.58 s.d. (Puglia) and 0.22 s.d. (Valle d’Aosta). We replicated the analysis, also controlling for the number of weeks of school closures (red lines). Regional differences are reduced as expected, but only slightly; thus, we may conclude that weeks of school closure do not fully explain regional differences.

5.3 Effects on inequalities

In order to analyse the impact of the pandemic on learning inequalities and to address the potential problem related to unanchored pre-test scores, we estimate models (5) and (6) in terms of z-scores, including interaction terms with all the explanatory variables for which we want to assess changes in inequalities between the pre-Covid and Covid cohorts (prior achievement, gender, parental education, migrant background, geographical area).

To begin, we focus on results relative to prior skills. In Fig. 4, we report the average marginal effects of the corresponding interaction term in Eq. 5. Overall, for a one standard deviation increase in test scores in Grade 10, the corresponding test scores in Grade 13 increase by 0.11/0.16 s.d. (Italian/maths) more in the Covid cohort than in the previous

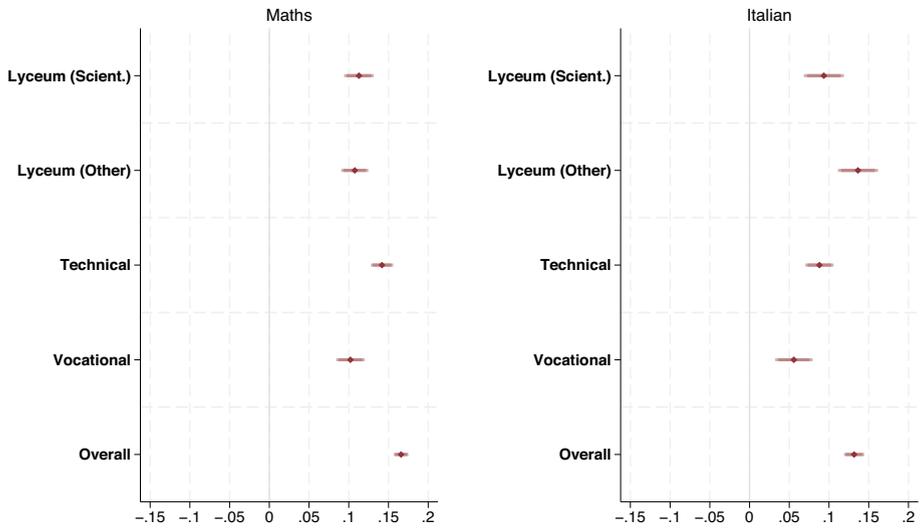


Fig. 4 Estimated effect of COVID-19 on learning inequalities by prior skills, Grade 13. Note: Estimates are generated with the OLS model in equation (5) using the 2018/19 and 2020/21 waves of the INVALSI dataset. The estimation models control for socio-demographic characteristics (gender, first/second generation migrant status, age, and high-educated parents - at least one parent has a tertiary degree), and school fixed effects. 90 percent and 95 percent confidence intervals are reported with horizontal lines

cohort. This means that previously low-performing children lost more than high-performing ones during the pandemic, and inequalities by ability have widened significantly. The results are consistent with most of the existing literature (notable exceptions are Birkelund and Karlson 2022, Contini et al. 2022).²⁰ If we look at the results by school type, we can see that this trend is more pronounced in lyceums for Italian and in technical schools for mathematics. The pandemic has thus exacerbated even more the educational inequalities among students with different initial skills - amplifying the risk of increasing inequalities in the long run.

Next, we describe the results on inequalities by socio-demographic dimensions (gender, parental education, migrant background). The average marginal effects of being in the Covid cohort by socio-demographic characteristics and conditional on prior abilities are reported in red (marked with a diamond) (Eq. 5), while the unconditional effects (not controlling for prior abilities and school fixed effects) are presented in green (triangle) (Eq. 6). The former can be thought of as the pandemic effect when comparing students with the same relative positions of previous performance and same school characteristics, the latter captures the variation in the overall learning gaps between socio-demographic groups. We also estimated a model with no prior performance but with school fixed effects (in blue; marked with a circle), in order to capture total effects not driven by carry-over effects of prior inequalities, but the results end up being usually in between conditional and unconditional ones.

The results for gender differences are shown in Fig. 5. Overall, the relative position of girls compared to boys can be seen to improve after the pandemic, particularly in Italian, but also in mathematics in Technical and Vocational schools (no gender differences are observable in Scientific lyceums). One possible explanation for this finding is that girls are more disciplined and self-controlled than boys (Duckworth and Seligman 2006). During school

²⁰ The discrepancies with the Contini et al. (2022) findings can be attributed to the varying levels of schooling, the duration and extent of school closures, and the geographical scope of the data.

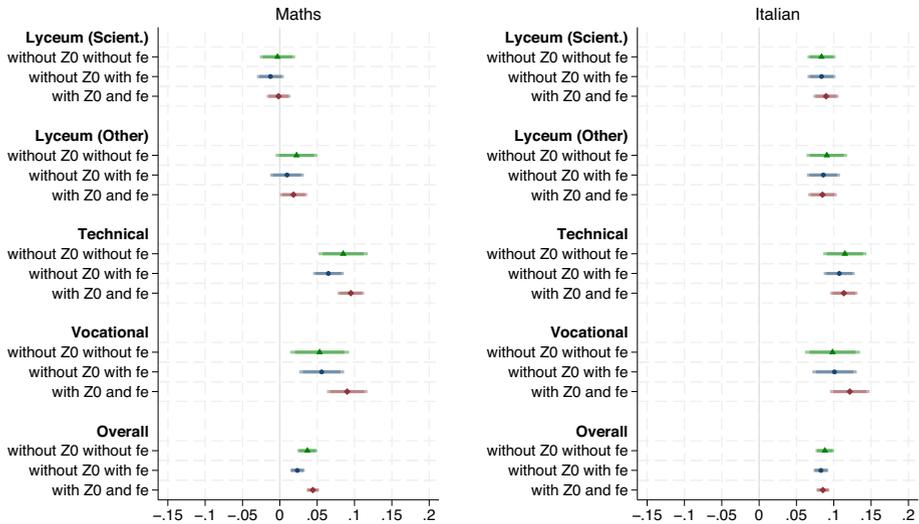


Fig. 5 Estimated effect of COVID-19 on learning inequalities by gender (girls vs boys), Grade 13. Note: Estimates are generated with the OLS models in equations (5) and (6) using the 2018/19 and 2020/21 waves of the INVALSI dataset. Z_0 is the student's prior ability in Grade 10 in maths (left-hand side) and in Italian (right-hand side), measured with INVALSI test scores in Grade 10 for maths and Italian standardised at the cohort level. Fe are school fixed effects. In all models we control for socio-demographic characteristics (gender, first/second generation migrant status, age, and high-educated parents - at least one parent has a tertiary degree). When we consider Grade 13 overall, we include a school track variable. 90 percent and 95 percent confidence intervals are reported with horizontal lines

closures self-discipline is particularly important, because in an online learning environment there is less feedback and less interaction between students and teachers (De Paola et al. 2023).²¹ Given the finding that better performers lose less and given that, on average, girls perform worse than boys in mathematics, it is not surprising that the relative improvement for girls found without controlling for prior achievement (and school fixed effects) is smaller than that observed when we do include prior achievement in the model.

Figure 6 presents differences by parental education. Overall, these inequalities remained virtually unchanged: most of the observed effects are small and statistically insignificant. This result is in line with existing studies on lower grades in Italy (Borgonovi and Ferrara 2023), which highlights an Italian specificity rather than a grade specificity and calls for further reflection. Why is it that, in Italy, contrary to theoretical predictions and international findings, there is no evidence that students from disadvantaged backgrounds have suffered the greatest learning losses? Unfortunately, we do not have a fully convincing explanation for this result, and more research is certainly needed. However, we can imagine a few hypotheses. It is possible that highly educated parents in highly skilled occupations were more likely to continue working during the pandemic, either physically or remotely, with even more intense work schedules than before, making it difficult for them to support their children effectively. Using data collected in Italy in spring 2020, Del Boca et al. (2020) show that both women and men spend less time with their offsprings if they continue to work away from home. In the

²¹ De Paola et al. (2023) find that online teaching during Covid-19 reduced the performance of university students. However, the effects differed greatly according to the students' tendency to procrastinate costly activities such as studying. If the same is true for younger students, this could explain the overall improvement of girls compared to boys.

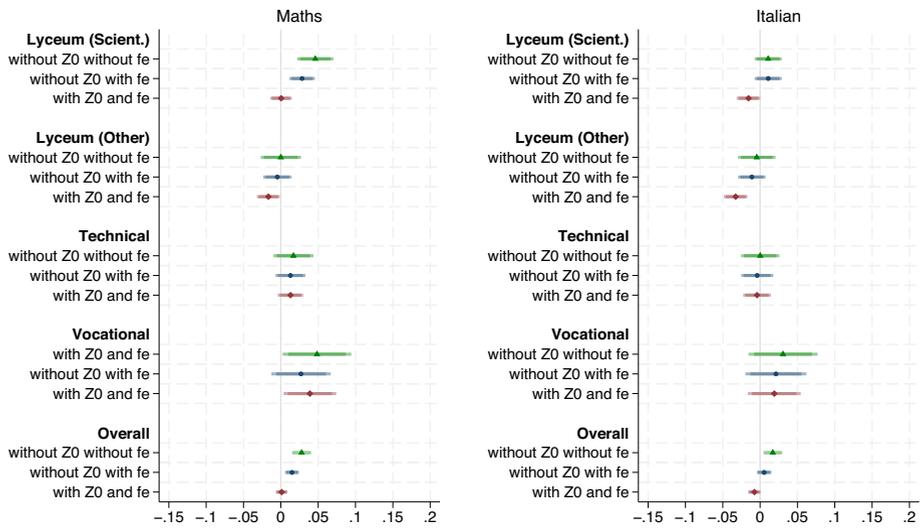


Fig. 6 Estimated effect of COVID-19 on learning inequalities by parental education ((high vs low), Grade 13. Note: Estimates are generated with the OLS models in equations (5) and (6) using the 2018/19 and 2020/21 waves of the INVALSI dataset. Z_0 is the student's prior ability in Grade 10 in maths (left-hand side) and in Italian (right-hand side), measured with INVALSI test scores in Grade 10 for maths and Italian standardised at the cohort level. Fe are school fixed effects. In all models we control for socio-demographic characteristics (gender, first/second generation migrant status, age, and high-educated parents - at least one parent has a tertiary degree). When we consider Grade 13 overall, we include a school track variable. 90 percent and 95 percent confidence intervals are reported with horizontal lines

U.S. Bansak and Starr (2021) revealed that in households experiencing a loss of employment income due to COVID-19, parents devoted more time to help children with schoolwork than households that did not face income loss. Such an effect could have been more pronounced in Italy, where many low-skilled workers were not allowed to go to work during the first lockdown (in spring 2020) and thus remained at home. On the contrary, white-collar and high-skilled office workers were overwhelmed with the need to learn how to use ICT tools in order to continue their activities remotely. This may be an Italian peculiarity, given the low level of digital literacy that most people had before the pandemic (European Commission 2020).

Figure 7 shows the results by immigrant background. Children from a migrant background end up improving slightly relative to native children with the same prior performances. This is indeed an unexpected finding. One possible explanation is that due to the significant disadvantages that migrant pupils face at school, they have to work harder to achieve the same results as natives. Therefore, when we run the analyses controlling for prior achievement, migrants may perform better after the pandemic because they are likely to be endowed with higher unobservable non-cognitive skills and/or resilience. However, as migrant pupils perform worse on average and the lowest achievers lose more, they have lost further ground overall relative to natives. Overall, we find that the migrant-native gap increased by 0.09 standard deviations in mathematics and 0.07 standard deviations in Italian during the pandemic. When we estimate the model without prior achievement but with school fixed effects, the estimated gap lies somewhere between the net effect and the total gap, implying that both channels, carry-over effects of prior inequalities and differential school effects, contribute to the overall loss of migrants relative to natives.

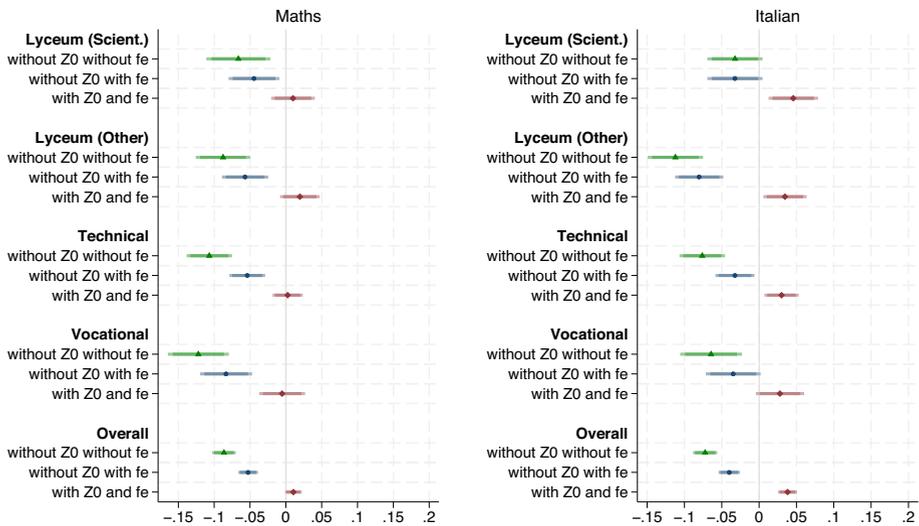


Fig. 7 Estimated effect of COVID-19 on learning inequalities by migrant status (migrant vs native), Grade 13. Note: Estimates are generated with the OLS models in equations (5) and (6) using the 2018/19 and 2020/21 waves of the INVALSI dataset. With the term migrant we refer to students born either in Italy or outside Italy from non-Italian parents (first- and second-generation migrants). Z_0 is the student's prior ability in Grade 10 for maths (left-hand side) and in Italian (right-hand side), measured with INVALSI test scores in Grade 10 for maths and Italian standardised at the cohort level. Fe are school fixed effects. In all models we control for socio-demographic characteristics (gender, first/second generation migrant status, age, and high-educated parents - at least one parent has a tertiary degree). When we consider Grade 13 overall, we include a school track variable. 90 percent and 95 percent confidence intervals are reported with horizontal lines

The effect of the pandemic on geographical achievement gaps²² is shown in Fig. 8. When comparing equally proficient students in Grade 10, students living in the South can be seen to have improved significantly over those living in the Northern regions.²³ This improvement is impressive, particularly in mathematics. It should be noted, however, that achievement gaps along the North-South divide have always been large, with southern students vastly underperforming (INVALSI 2022).²⁴ Thus, as the gap between high and low achievers widened, not conditional on prior achievement the gap appears essentially unchanged.²⁵

5.4 Robustness checks

To confirm the validity of our results, we now perform robustness checks based on model (2).

²² Note that since these differences are not identified with school fixed effects, these results derive from the estimation of a version of also model (5) that does not include them.

²³ When we control for Z_0 , students in the Centre are in between those of the North and South for maths. For Italian, they are close to the South. Unconditional on Z_0 and school fixed effects, students in the Centre of Italy are not significantly different from students in the North. Results available from the authors upon request.

²⁴ The reasons for these differences have been attributed to the role of contexts and school quality (Bratti et al. 2007).

²⁵ The pandemic's effects on inequalities are similar when examining the cohort of students who graduated high-school in 2022 - available from the authors upon request.

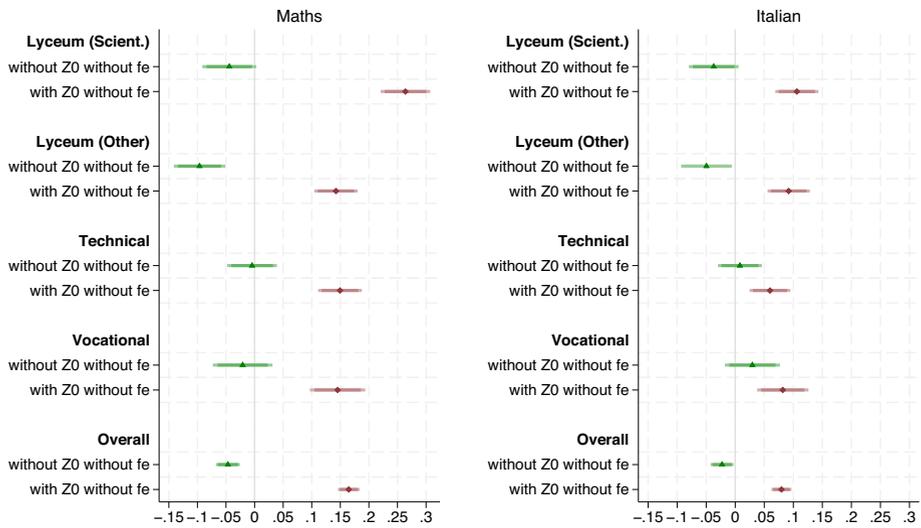


Fig. 8 Estimated effect of COVID-19 on learning inequalities by geographical area (South vs North), Grade 13 Note: Estimates are generated with the OLS models in equations (5) and (6) using the 2018/19 and 2020/21 waves of the INVALSI dataset. Z_0 is the student's prior ability in Grade 10 in maths (left-hand side) and in Italian (right-hand side), measured with INVALSI test scores in Grade 10 for maths and Italian standardised at the cohort level. In both models (with and without Z_0) we control for socio-demographic characteristics (gender, first/second generation migrant status, age, and high-educated parents - at least one parent has a tertiary degree). When we consider Grade 13 overall, we include a school track variable. 90 percent and 95 percent confidence intervals are reported with horizontal lines

The first issue to address is that our analytical sample consists of students who participated in assessment in Grades 10 and 13 and who did not repeat a school year between the two grades. As mentioned in Section 3, this feature implies that the analytical samples used for the difference-in-differences analysis are to some extent positively selected. By analysing the data of students in Grade 13 without controlling for prior performance, we can compare the estimates deriving from the total population of students who took the tests in Grade 13 with those of the selected population that we were able to link with the test data in Grade 10. The results are shown in Table 4. In column 1, we report the results for our final sample, while in column 2, we report those of the full sample.²⁶ Estimates of the learning loss are very similar in the two samples for both Italian and maths. The small differences are consistent with the expectations: since in our main analyses we found a greater loss for previously poorly achieving students, a smaller learning deficit should be observed in the analytical (selected) sample rather than in the full sample. This suggests that even in the difference-in-differences analysis the level of bias should be small, and that, if anything, the learning loss is slightly underestimated. Also note that when not controlling for prior abilities, the estimated learning loss is somewhat smaller than the results when we include Grade 10 test scores in the model (-0.33 s.d. in maths and -0.36 s.d. in Italian, vs -0.39 s.d. and -0.41 s.d. in our preferred specification).

²⁶ Since data on parental education and occupation were not available for the initial sample (because the information was retrieved from the Grade 10 assessment), we controlled for the student ESCS (Economic, Social and Cultural Status) instead. When comparing the final sample estimates derived from using ESCS with those derived from using parental education and occupation we find very similar results.

Table 4 The impact of COVID-19 on maths and Italian test scores in Grade 13, not controlling for prior achievements in Grade 10 - final and initial sample

	Final sample (1)	Initial sample (2)
	Maths	
Covid cohort	-0.329*** (0.004)	-0.334*** (0.003)
	Italian	
Covid cohort	-0.358*** (0.004)	-0.369*** (0.004)
Obs.	618,226	852,862
Initial Abilities	No	No
Socio-Demogr.	Yes	Yes
School FE	Yes	Yes

Note: The final sample (1) consists of all students in pre-Covid and Covid cohorts who performed both Italian and maths INVALSI assessment tests in G13 and were successfully matched with the observations with INVALSI sample in Grade 10. The initial sample (2) includes all the students in pre-Covid and Covid cohort who performed both Italian and maths INVALSI assessment tests in G13. Since the variables for parental occupation and high-educated parents (at least one parent has a tertiary degree) are not available for the initial sample, in columns (1) and (2) we control for student ESCS. Socio-demographic controls include gender, first and second generations migrant status, age, and student ESCS. Standard errors in parentheses are clustered at the class level. School FE: School fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A second robustness check takes into consideration the fact that due to the school closure that occurred in spring 2020, the Ministry of Education suspended grade retention for the current school year (acknowledging that schools were unprepared to cope with the new situation, remote learning was not mandatory and only oral exams were allowed). This results in a lack of full comparability between the two cohorts, which were subjected to different rules: in the pre-Covid cohort, Grade 12 students with low results were exposed to the risk of being retained, whereas this was not the case for those in the Covid cohort. For this reason, the group of Grade 13 students in the Covid cohort could be to the same extent poorer performing than the corresponding group in the pre-Covid cohort. Consequently, the risk is to overestimate the negative effects of the pandemic on student learning. To account for this imbalance, we derived the proportion of students who were held back between Grade 12 and Grade 13 in school year 2018/2019 from the statistics of the Italian Ministry of Education: 3.33% in Scientific lyceums, 2.95% in Other lyceums, 7.15% in the Technical track and 10% in the Vocational track (Ministero dell'Istruzione 2020). To simulate what would have happened if grade retention had been applied, we removed the corresponding proportions of retention for each track from the lowest performing students in the Covid cohort sample. The results, presented in Table 5, column 2, are very similar to the main estimates already shown in Table 2 and reported again in column 1. Again, the minimal observed differences go in the expected direction, with the new estimates being slightly smaller than the main ones.

Third, one potential additional issue with national assessments performed during the Covid-19 pandemic is attrition bias. As pointed out by Werner and Woessmann (2023) in their study on Germany, a larger fraction of students did not participate in the assessments during

Table 5 The impact of COVID-19 on maths and Italian test scores, in Grade 13, accounting for grade retention and outlier regions

	Main results (1)	Accounting for grade retention ¹ (2)	Without Puglia (3)	Without Campania (4)	Without Calabria (5)
Maths					
Covid cohort	-0.389*** (0.004)	-0.386*** (0.004)	-0.382*** (0.004)	-0.391*** (0.004)	-0.391*** (0.004)
Italian					
Covid cohort	-0.410*** (0.004)	-0.404*** (0.004)	-0.403*** (0.004)	-0.408*** (0.004)	-0.412*** (0.004)
Obs.	618,226	601,117	583,892	554,457	599,157
Initial Abilities	Yes	Yes	Yes	Yes	Yes
Socio-Demogr.	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes

Note: ¹Grade retention was suspended in the school year 2019-20. We drop a share of low performing students in the Covid cohort according to Grade 12 retention in the school year 2018/2019 (3.33% Lyceum Scientific, 2.95% Lyceum Other, 7.15% Technical and 10% Vocational). 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 generation migrant status, age, parental occupations, and high-educated parents (at least one parent has a tertiary degree). School FE: School fixed effects. Standard errors in parentheses are clustered at the class level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

the pandemic than during normal times. If these students are low achievers, as one would expect, then the learning deficit is underestimated. In our data, we can measure attrition as the proportion of students who participated in the Grade 10 assessments and not in the Grade 13 ones, separately for the Covid and the pre-Covid cohort. Consistent with expectations, attrition in the Covid cohort (28%) is larger than in the pre-Covid cohort (21%), with large regional variation (Table A5 in the Online Appendix). Attrition bias can be reduced by controlling for prior ability, as done in our main analyses. Nevertheless, to get a sense of the possible bias that this problem introduces, we estimate the probability of taking the assessment in Grade 13 separately for the pre-Covid and Covid cohort. 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 indicating if the student participated in the Grade 13 assessment. Columns 1 and 2 of Table A6 in the Online Appendix report the estimates controlling only for maths and Italian test scores in Grade 10. As expected, high achievers in Grade 10 are more likely to participate in the Grade 13 assessment and slightly more in the Covid cohort. This result is confirmed when controlling for gender, parental education and migratory background (Columns 3 and 4). Results indicate that socio-demographic variables also predict the probability of participation: girls, students with highly educated parents and natives are more likely to participate in Grade 13 assessment, conditional on their prior ability. Overall, these results suggest that non-participation could lead to a small underestimation of learning loss, in line with Werner and Woessmann (2023).

Lastly, we perform an additional robustness check, re-estimating the main model by excluding two outlier regions that experienced a much longer school closure during the pandemic (Puglia, 37.4 weeks, and Campania, 34.8 weeks, as compared to the national aver-

age 29.9; see Fig. 3) and Calabria, which has a high attrition rate in the Covid cohort. Puglia also had a much larger proportion of attrition in the Covid cohort than the rest of Italy: the proportion of students present in Grade 10 who did not participate in Grade 13 assessment was 59% (Table A5 in the Online Appendix). There is no substantial difference with the main results (Table 5, columns 3, 4, and 5).

6 Discussion and conclusions

This paper focuses on the learning loss due to Covid-19 for students at the end of upper secondary school in Italy, a country that was already lagging behind other rich countries before the pandemic in terms of GDP, human capital accumulation, learning outcomes, tertiary attainment and labour market outcomes for young people.

Our contribution to the existing literature is twofold. First, although the literature on the effects of Covid on learning at the primary and lower secondary levels is now quite extensive, there is still a lack of empirical evidence on the effects for older students. Using rich panel data from national standardised tests for the whole student population, repeated over different cohorts, this paper analyses the learning loss associated with the pandemic and how inequalities between socio-demographic groups have changed.

Second, from a methodological perspective, we clarify when standard DID models can be used and when extended DID models are needed. To summarise, the standard DID model can be used in cases where all test scores are measured on the same scales, i.e. are both horizontally and vertically equated, and achievement growth as children get older does not depend on prior achievement. The extended DID model should be used to estimate average and heterogeneous effects in the other cases, provided that test scores in the same grade are on the same scale. While average effects cannot be estimated when same-grade tests are not horizontally equated, heterogeneous effects can be analysed using extended DID model on within-cohort and grade standardised test scores, which delivers information on changes in the relative position of the test score distributions between groups.

Against this background, focusing on students who were first hit by the pandemic during Grade 12, we estimate two sets of difference-in-differences models. With the first one, we estimate the average effect of the pandemic on student learning at the end of Grade 13, comparing the performance of students in the pandemic cohort (measured in spring 2021) with that of students in a pre-pandemic cohort (measured in spring 2019), while controlling for prior skills at the end of Grade 10. As the Grade 10 assessments were not horizontally equated, these estimates are based on the untestable assumption that the prior distribution of skills did not change between the two cohorts. The second set of models does not require this assumption and aims to investigate whether and how inequalities have changed during the pandemic period by analysing the relative position of the different groups, defined by prior performance, gender, parental education, migratory background and geographical area.

Our main findings can be summarised as follows. The average impact of the pandemic is extremely large in both mathematics (-0.39 s.d.) and reading (-0.41 s.d.), with no marked differences between tracks. The negative effects vary widely across regions, even after controlling for regional differences in the duration of school closures, suggesting that other contextual factors matter. Altogether, these estimates are much larger than those obtained for lower grades in Italy (Borgonovi and Ferrara 2023; Contini et al. 2022), suggesting that the disruption was much greater in high school than in earlier grades. While the losses experienced by younger children are of great concern because of the cumulative nature of learning,

the learning losses experienced by students at the end of their school careers can also be critical. These individuals are about to enter either the labour market or tertiary education, with major shortcomings compared to the past. In Italy, the situation is especially worrisome, as even before the pandemic, the level of adult maths and reading competencies according to the Survey of Adult Skills (PIAAC) was very low, and the percentage of NEET was very high. Moreover, for students at the end of Grade 13 in 2022, the learning loss is of similar magnitude, suggesting that effects of shocks are long-lasting and that recovery can be difficult, because learning is a cumulative process and later steps rely on concepts that need to be well understood beforehand.

In terms of inequalities, consistent with most of the literature, we find that previously lower achieving students experienced the largest losses. The relative position of girls compared to boys improved after the pandemic, that is, boys lost more than girls, with opposite effects in terms of inequality: the gender gap in reading (in favour of girls) increased, whereas the gender gap in mathematics (in favour of boys) decreased. Conditional on prior abilities, the learning gap between students with a migratory background (first and second generations) and natives decreased. We speculate that this result could be due to unobservable non-cognitive skills and resilience that helped the migrant students more than the natives. Note, however, that the overall inequality (unconditional on initial ability) between migrants and natives actually increased during the pandemic due to the large initial achievement gap in favour of native children.

Our results show no significant differences related to parents' education, while most of the international evidence emphasises the exacerbation of inequalities based on parents' socio-economic background. These results are in line with other results for Italy in the lower grades (Borgonovi and Ferrara 2023). In the results' section we speculate on possible explanations, pointing to the specificity of the Italian case, where many low-skilled workers were forced to stay home (not working), while high-skilled workers who worked remotely had to learn how to use ICT tools they had no familiarity with (as already mentioned, before the pandemic Italy had very low digital skills among the adult population).

Our results strongly call for educational policies to support the formation of human capital for this generation of students, and in particular for the most fragile groups. Indeed, more research is needed to better understand the medium-term legacy of the pandemic and counteract the negative impact on the development of skills and the professional futures of boys and girls. However, it is already clear that the price paid by the younger generation for the pandemic is very high, with likely long-term consequences for this generation and for society as a whole. Urgent remedial action is needed to compensate for these losses and to support the human capital formation of students of all ages, including those in high school and university, who are entering the labour market with a very heavy burden. In the absence of education policies that effectively address these gaps, there is a high risk of an increase in university dropout, the proportion of NEETs, as well as a sharp decline in employment prospects, wages and, ultimately, national growth.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10888-025-09708-2>.

Acknowledgements We gratefully acknowledge financial support from the University of Torino, the University of Trento, Collegio Carlo Alberto, MUR FISR. We would like to thank the Statistical Office of INVALSI for both data and fruitful collaboration. We also recognise the valuable contribution of Alessio Rossi and Chiara Ardito as research assistants. We are grateful to several seminar and conference participants for their helpful comments.

Author Contributions All authors jointly developed the research idea. MLDT and DP acquired fundings. DC and DP developed the empirical models, CM and LS performed the empirical analyses, all authors interpreted the findings. DC and DP wrote the main text and all authors reviewed the manuscript.

Funding Open access funding provided by Università degli Studi di Trento within the CRUI-CARE Agreement. The research leading to these results received funding from the University of Torino, the University of Trento, Collegio Carlo Alberto, and the Italian Ministry of University and Research (MUR FISR - Fondo Integrativo Speciale per la Ricerca 2020, grant n. FISR2020IP_02236).

Data Availability Data used in the article are available for request from INVALSI at <https://invalsi-serviziostatistico.cineca.it>.

Declarations

Competing Interests The authors declare no competing interests.

Ethics, Consent to Participate, and Consent to Publish declarations Not applicable.

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

- Almeida, M., Challa, M., Ribeiro, M., Harrison, A.M., Castro, M.C.: Editorial perspective: The mental health impact of school closures during the COVID-19 pandemic. *J. Child Psychol. Psychiatry* **63**(5), 608–612 (2022)
- Angrist, J.D., Battistin, E., Vuri, D.: In a small moment: Class size and moral hazard in the Italian Mezzogiorno. *American Econ. J. Appl. Econ.* **9**(4), 216–249 (2017)
- Bacher-Hicks, A., Goodman, J., Green, J.G., Holt, M.K.: The COVID-19 pandemic disrupted both school bullying and cyberbullying. *American Econ. Rev. Insights* **4**(3), 353–370 (2022)
- Banko-Ferran, D., Gihleb, R., Giuntella, O.: The Impact Of COVID-19 On Mental Health. In: Zimmermann, K.F. (ed.) *Handbook of Labor. Human Resources and Population Economics*. Springer, Cham (2023)
- Bansak, C., Starr, M.: COVID-19 shocks to education supply: how 200,000 US households dealt with the sudden shift to distance learning. *Review of Econ. Household* **19**(1), 63–90 (2021)
- Battisti, M., Maggio, G.: Will the last be the first? School closures and educational outcomes. *European Econ. Rev.* **154**, 104405 (2023)
- Bertoletti, A., Soncin, M., Cannistrà, M., Agasisti, T.: The educational effects of emergency remote teaching practices - The case of COVID-19 school closure in Italy. *PLoS ONE* **18**(1), e0280494 (2023)
- Bertoni, M., Brunello, G., Rocco, L.: When the cat is near, the mice won't play: The effect of external examiners in Italian schools. *J. Public Econ.* **104**, 65–77 (2013)
- Bethhäuser, B.A., Bach-Mortensen, A.M., Engzell, P.: A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behav.* **7**, 375–385 (2023)
- Birkelund, J.F., Karlson, K.B.: No evidence of a major learning slide 14 months into the COVID-19 pandemic in Denmark. *European Soc.* **1–21**, (2022). <https://doi.org/10.1080/14616696.2022.2129085>
- Björkegren, E., Svaleryd, H., Vlachos, J.: Remote instruction and student mental health: Swedish evidence from the pandemic. *Am. Econ. J. Econ. Pol.* **16**(4), 491–517 (2024)
- Blaskó, Z., Costa, P.d., Schnepf, S.V.: Learning losses and educational inequalities in Europe: Mapping the potential consequences of the COVID-19 crisis. *J. European Soc. Policy* **32**(4), 361–375 (2022)
- Bloom, H.S., Hill, C.J., Black, A.R., Lipsey, M.W.: Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *J. Res. Educ. Effective.* **1**(4), 289–328 (2008)

- Bond, T.N., Lang, K.: The evolution of the Black-White test score gap in Grades K-3: The fragility of results. *Rev. Econ. Stat.* **95**(5), 1468–1479 (2013)
- Borgonovi, F., Ferrara, A.: COVID-19 and inequalities in educational achievement in Italy. *Res. Soc. Stratific. Mobility.* **83**, 100760 (2023)
- Bratti, M., Checchi, D., Filippin, A.: Geographical differences in Italian students' mathematical competencies: Evidence from Pisa 2003. *Giornale degli Econ. Annal. Econ.* **66**(3), 299–333 (2007)
- Camera dei Deputati.: Le misure adottate a seguito dell'emergenza Coronavirus (COVID-19) per il mondo dell'istruzione (scuola, istruzione e formazione professionale, università, Istituzioni AFAM). (2022) <https://temi.camera.it/leg18/temi/le-misure-adottate-a-seguito-dell-emergenza-coronavirus-covid-19-per-il-mondo-dell-istruzione-scuola-istruzione-e-formazione-professionale-universit-istituzioni-afam.html>
- Carlana, M., La Ferrara, E., Lopez, C.: Exacerbated inequalities: The learning loss from COVID-19 in Italy. *Aea papers and proceedings (Vol. 113, pp. 489–493)* (2023)
- Contini, D., Cugnata, F.: Does early tracking affect learning inequalities? Revisiting difference-in-differences modeling strategies with international assessments. *Large-scale Assess. Educ.* **8**(1), 14 (2020)
- Contini, D., Di Tommaso, M.L., Muratori, C., Piazzalunga, D., Schiavon, L.: Who lost the most? Mathematics achievement during the COVID-19 pandemic. *The BE J. Econ. Anal. Policy* **22**(2), 399–408 (2022)
- Contini, D., & Salza, G.: Too few university graduates. Inclusiveness and effectiveness of the Italian higher education system. *Socio-Econ. Plan. Sci.* **71**, 100803, (2020)
- De Paola, M., Gioia, F., Scoppa, V.: Online teaching, procrastination and student achievement. *Econ. Educ. Rev.* **94**, 102378 (2023)
- Del Boca, D., Oggero, N., Profeta, P., Rossi, M.: Women's and men's work, housework and childcare, before and during COVID-19. *Rev. Econ. Household* **18**, 1001–1017 (2020)
- Duckworth, A.L., Seligman, M.E.: Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *J. Educ. Psych.* **98**(1), 198 (2006)
- Education GPS.: Report on Italy. OECD Publishing, Paris. (2023). <http://gpseducation.oecd.org>
- Engzell, P., Frey, A., Verhagen, M.D.: Learning loss due to school closures during the COVID-19 pandemic. *Proceed. National Acad. Sci.* **118**(17), e2022376118 (2021)
- European Commission.: The Digital Economy and Society Index (DESI). Italy. (2020). <https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2020>
- European Schoolnet.: Survey of Schools: ICT in Education. Country profile: Italy. (2012). https://ec.europa.eu/information_society/newsroom/image/document/2018-3/italy_country_profile_2FC554D7-A7D2-ECAC-E56720235DEE9BDD_49443.pdf
- Eurostat.: Educational attainment statistics. (2023). https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Educational_attainment_statistics
- Eurydice.: National Education Systems: Overview. Italy. (2023) <https://eurydice.eacea.ec.europa.eu/national-education-systems/italy/overview>
- Fuchs-Schündeln, N., Krueger, D., Ludwig, A., Popova, I.: The long-term distributional and welfare effects of COVID-19 school closures. *The Econ. J.* **132**(645), 1647–1683 (2022)
- Haelermans, C., Jacobs, M., van der Velden, R., van Vugt, L., van Wetten, S.: Inequality in the Effects of Primary School Closures due to the COVID-19 Pandemic: Evidence from the Netherlands. *AEA Papers Proceed.* **112**, 303–307 (2022)
- Hampf, F., Wiederhold, S., Woessmann, L.: Skills, earnings, and employment: exploring causality in the estimation of returns to skills. *Large-scale Assess. Educ.* **5**, 1–30 (2017)
- Hanushek, E.A., Woessmann, L.: The economic impacts of learning losses. OECD Publishing, Paris (2020)
- INVALSI.: Rapporto prove Invalsi 2022. Roma: INVALSI. (2022) https://invalsi-areaprove.cineca.it/docs/2022/Rilevazioni_Nazionali/Rapporto/Rapporto_Prove_INVALSI_2022.pdf
- Jack, R., Halloran, C., Okun, J., Oster, E.: Pandemic schooling mode and student test scores: evidence from us school districts. *American Econ. Rev. Insights* **5**(2), 173–190 (2023)
- Jack, R., Oster, E.: COVID-19, school closures, and outcomes. *J. Econ. Perspect.* **37**(4), 51–70 (2023)
- Kofoed, M.S., Gebhart, L., Gilmore, D., Moschitto, R.: Zooming to class? experimental evidence on college students' online learning during COVID-19. *American Econ. Rev. Insights* **6**(3), 324–340 (2024)
- Lichand, G., Doria, C.A., Leal-Neto, O., Fernandes, J.P.C.: The impacts of remote learning in secondary education during the pandemic in Brazil. *Nature Human Behav.* **6**(8), 1079–1086 (2022)
- Lucifora, C., Tonello, M.: Cheating and social interactions. Evidence from a randomized experiment in a national evaluation program. *J. Econ. Behav. Organiz.* **115**, 45–66 (2015). <https://doi.org/10.1016/j.jebo.2014.12.006>
- Maldonado, J.E., De Witte, K.: The effect of school closures on standardised student test outcomes. *British Educ. Res. J.* **48**(1), 49–94 (2022). <https://doi.org/10.1002/berj.3754>

- Maldonado, J.E., De Witte, K., Declercq, K.: The effects of parental involvement in homework: two randomised controlled trials in financial education. *Empir. Econ.* **62**(3), 1439–1464 (2022)
- Mazrekaj, D., De Witte, K.: The impact of school closures on learning and mental health of children: Lessons from the COVID-19 pandemic. *Perspect. Psych. Sci.* **19**(4), 686–693 (2024)
- Ministero dell'Istruzione.: Focus "Esiti degli scrutini del secondo ciclo di istruzione". Anno Scolastico 2018/2019. Roma: Ufficio Gestione Patrimonio Informativo e Statistica (2020)
- Neidhöfer, G., Lustig, N., Tommasi, M.: Intergenerational transmission of lockdown consequences: prognosis of the longer-run persistence of COVID-19 in latin america. *The J. Econ. Inequal.* **19**(3), 571–598 (2021)
- OECD.: Talis 2018 results (volume I). Teachers and school leaders as lifelong learners. OECD Publishing, Paris (2018)
- OECD: PISA 2022 Results (Volume I and II) - Country Notes: Italy. OECD Publishing, Paris (2023)
- Patrinos, H., Vegas, E., Carter-Rau, R.: An Analysis of COVID-19 Student Learning Loss. *Oxford Res. Encycloped. Econ. Finance* (2023). <https://doi.org/10.1093/acrefore/9780190625979.013.893>
- Racine, N., McArthur, B.A., Cooke, J.E., Eirich, R., Zhu, J., Madigan, S.: Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. *JAMA Pedia.* **175**(11), 1142–1150 (2021)
- Salza, G.: Equally performing, unfairly evaluated: The social determinants of grade repetition in Italian high schools. *Res. Soc. Stratific. Mobility* **77**, 100676 (2022)
- Sandner, M., Patzina, A., Anger, S., Bernhard, S., Dietrich, H.: The COVID-19 pandemic, well-being, and transitions to post-secondary education. *Rev. Econ. Household* **21**(2), 461–483 (2023)
- Schult, J., Mahler, N., Fauth, B., Lindner, M.A.: Did students learn less during the COVID-19 pandemic? Reading and mathematics competencies before and after the first pandemic wave. *Sch. Effective. Sch. Improve.* **33**(4), 544–563 (2022)
- Todd, P.E., Wolpin, K.I.: On the specification and estimation of the production function for cognitive achievement. *The Econ. J.* **113**(485), F3–F33 (2003)
- UNDP.: COVID-19 and Human Development: Assessing the Crisis, Envisioning the Recovery. New York: United Nations Development Programme. (2020). <https://hdr.undp.org/content/covid-19-and-human-development-assessing-crisis-envisioning-recovery>
- UNESCO.: Monitoring of school closures. (2023) <https://www.unesco.org/en/covid-19/education-response>
- UNICEF.: COVID-19 and children. UNICEF data hub. (2022). <https://data.unicef.org/covid-19-and-children/>
- Vegas, E.: COVID-19's impact on learning losses and learning inequality in Colombia. *Center for Universal Education at Brookings*, **599**, (2022)
- Werner, K., & Woessmann, L.: The Legacy of COVID-19 in Education. *Economic Policy*, eiad016, <https://doi.org/10.1093/epolic/eiad016> (2023)

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

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.