E-learning Engagement Gap During School Closures: Differences by Academic Performance

Josep Amer-Mestre* Alaitz Ayarza-Astigarraga* Marta C. Lopes†‡

November, 2022

Click here for the latest version

Abstract

We study the impact of COVID-19 school closures on differences in online learning usage by regional academic performance. Using data from Google Trends in Italy, we find that during the first lockdown, regions with a previously lower academic performance increased their searches for e-learning tools more than higher-performing regions. Analysing school administrative and survey data before the pandemic, we find that both teachers and students in lower performing regions were using no less e-learning tools than higher performing ones. These two findings suggest that the COVID-19 shock widened the e-learning usage gap between academically lower and higher-performing regions. Exploiting the regional variation in school closure mandates during the 2020-2021 academic year, we report that the patterns detected after the first lockdown were no longer present. Regions with different previous academic performance had the same response in terms of online learning usage when faced with stricter school closures.

JEL classification: C31, C81, I24, H75

Keywords: E-learning, COVID-19, Education, Inequality, School Closures

Acknowledgments: We thank Thomas Crossley, Andrea Ichino, Anna Lavizzari, Eleonora Erittu and Pedro Freitas, for very helpful comments and suggestions. We also thank the very useful discussions at the 2nd Joint IZA and Jacobs Centre Workshop, 37th AIEL National Conference, and at the internal seminars of the European University Institute and Universidad Carlos III de Madrid. Finally, we gratefully acknowledge the financial support from the European University Institute COVID-19 Knowledge Hub and the support provided by the institution in granting access to safely secured data. The usual disclaimer applies. Declarations of interest: none.

*European University Institute

†Universidad Carlos III de Madrid; Institute of Labor Economics (IZA)

‡Corresponding author: Marta C. Lopes. e-mail: marta.cardoso@uc3m.es address: Calle Madrid 126, 28903 Getafe, Madrid, Spain.

1 Introduction

Closing schools has been one of the primary measures of governments worldwide to prevent the spread of the COVID-19 virus. As a result, teachers and students were forced to an unprecedented sudden transition from face-to-face to online schooling. Empirical evidence indicates that this pandemic brought short-term average learning losses for students (see e.g. Maldonado and De Witte, 2021; Contini et al., 2022), and disproportionate learning losses for students from lower socioeconomic backgrounds (Engzell et al., 2021). In this context, it is important to understand how different areas of a country responded to the shock, and whether existing regional disparities have been widened. In particular, the extent to which e-learning resources – which have required substantial investments by governments, schools and households – have been used across the different regions of the country during the mandatory school closures is of particular relevance.

In this paper, we study the differential response in online learning usage by regions with different pre-pandemic academic performance during school closures. Most of the literature studying the link between e-learning and academic performance has focused on the possible adverse effects on student outcomes.[1] However, little is known about how students in regions with different academic performances engage with the available e-learning tools when their in-presence class time is reduced (Figlio et al., 2013; Joyce et al., 2015).

We analyse the heterogeneity in e-learning engagement during two periods of school closures. We measure the engagement with online learning resources using real-time data via Google Trends for Italian regions and analyse two distinct periods. One from September 2016 to June 2020, which includes the period in which a nationwide school closure was implemented. A second one from November 2020 to June 2021, in which school lessons were carried out either in-person or online intermittently depending on the local spread of the virus. To measure academic performance, we use pre-pandemic average standardised test scores in reading and mathematics administered by the National Institute for the Evaluation of the Education and Training System in Italy (INVALSI).[2]

[1]The following papers have found none to negative average effects of e-learning tools on academic performance: Brown and Liedholm (2002), Fairlie and Robinson (2013), Figlio et al. (2013), Joyce et al. (2015), Beuermann et al. (2015), Bando et al. (2017), Cristia et al. (2017), and Lu and Song (2020).

[2]Established in 1997, among other tasks, INVALSI is entrusted with administering periodic tests to evaluate students' academic achievement at different levels of education.

We document four main findings. First, Google search data on selected popular e-learning platforms show a vast surge in their searches right after the nationwide school closure was introduced. We estimate that during the more than three months that schools remained closed there was an average increase of 142 percent in searches of such type of platforms, compared to the pre-pandemic period. This big and yet sustained increase in the Google searches for e-learning platforms is explained by the unprecedented transition that the schooling system had to endure in the first months of the COVID-19 pandemic.

Second, estimating a difference-in-differences specification we find that after schools were closed, regions with lower academic performance experienced a 19 percent *higher* increase in their search for online learning resources compared to higher-achieving regions. This is the opposite result found in the U.S., where it was more affluent areas, with better internet access, and fewer rural schools that saw substantially larger increases in internet search intensity of online learning resource (Bacher-Hicks et al., 2021).

Third, we report that regions with higher average academic performance did not have a higher engagement in online learning in pre-COVID-19 times. Using PISA (OECD Programme for International Student Assessment) and INVALSI surveys, we show a statistically significant negative association between academic performance and the use of online learning resources by students outside school and by teachers in-class at a regional level. Thus, we argue that Italian regions with a higher academic performance did not face the COVID-19 outbreak with a greater familiarity in the use of online learning resources.

Finally, the analysis of the 2020-2021 academic year allows us firstly to ascertain the accuracy of the Google Trends data as a valid proxy for measuring changes in e-learning platform usage. Secondly, to show that previous academic performance was no longer a relevant factor determining differences in e-learning platform usage in the new academic year.

All these results, taken together, suggest that the first months of the pandemic contributed to widening the gap in the usage of online learning resources between academically high and low performing regions in Italy. Not only did lower-performing regions have higher levels of online learning usage before the pandemic, but during the pandemic, they also increased the search for online learning resources more than their counterparts. However, we find that during the subsequent academic year, regions with different pre-pandemic academic performances did not react differently to localised school closure mandates in terms of online learning usage.

Bacher-Hicks et al. (2021) conclude that the usage of e-learning platforms will be one of the main channels through which the COVID-19 pandemic will likely widen socioeconomic gaps in the U.S. Our findings however, do not point in that direction. Instead, we show that the increase in the short-term educational gaps across academically high and low achieving regions in Italy caused by the pandemic, would not be attributed to a lower engagement with online learning resources by lower-achieving regions.

We reconcile the opposite findings of our study and those reported in Bacher-Hicks et al. (2021) by arguing that both in Italy and the U.S., it was the areas with higher preexisting levels in e-learning usage that saw a higher immediate growth in their search rates after the school closures. Thus, we claim that the heterogeneity in the transition cost faced by students and teachers is the main mechanism behind this result. In other words, we argue that in both countries, areas with higher rates of learning usage during pre-pandemic years were simply better equipped to navigate the transition from in-person to online schooling.

This study focuses on Italy, which is an interesting country to study for at least three reasons. First, due to the rapid spread of the virus and its virulence – which resulted in around 130,000 deaths as of September 2021 – it was the first country to close schools outside Asia, and the country with one of the longest school closure period. Schools remained closed nationwide from March 4 2020 until mid-September 2020. Starting in Fall 2020 and throughout the entire 2021 academic year, the Italian Government used a regionalised system to control the virus spread. Depending on regional outbreaks, each Italian region was assigned a different colour in a four-colour category system. Each of them corresponded to a different set of measures, including school closing mandates. We answer our research question using two different analyses, one for each institutional setting.

Second, Italy is a country that presents substantial regional differences in school quality and academic performance (Agasisti and Vittadini, 2012; Argentin and Triventi, 2015), with a pronounced North-South divide. In a country where achievement gaps across regions are a concern, studying how regions with different academic performances engaged in online learning during the pandemic is relevant. Third and finally, the Italian setting allows the comparison between the engagement of regions on online learning using platforms that the Government promoted at a national level. Right after the announcement of school closures, the centralised school management in the country put forward a website (*didattica a distanza*) to support schools in implementing online learning methods.

Our work adds to the literature on the effect of COVID-19-induced school closures on online learning engagement by focusing on the differential impact across regions by their academic performance. The learning time gap between low and high academic achievers has been studied using time-use survey data in Germany. Grewenig et al. (2021) show that the reduction of daily time spent learning was significantly larger for low achievers than for high achievers, while they do not find differences by students' socioeconomic status.[3]

We also contribute to the literature that exploits Google Trends data. We show that these data can provide useful, reliable, and real-time information for education-related choices not only in the U.S., but also in smaller countries, with lower initial searches on the Google search tool, such as Italy. In this regard, due to the sampling feature of Google Trends, we call attention to the need to download several samples in settings such as ours. Using nationally representative high-frequency data, our paper documents that lower-achieving regions in Italy were no less engaged in online learning during the lockdown – as measured by their searches for e-learning platforms – than high achieving regions.

The results of this paper can help inform future policy responses in education. If the performance gaps widen as a result of the pandemic, the evidence in this paper calls for a greater involvement of the Government than just providing families with access to these platforms in periods when schools are forced to close.[4] If this were the case, our paper is consistent with a subtler channel: for example, lower-achieving regions doing a less efficient use of online resources, where

[3] Andrew et al. (2020) show that the gap in the time used for learning between primary school students from high and low socioeconomic status increased in England.

[4] For example, Carlana and La Ferrara (2021) find that an intervention giving free, individual, online tutoring to disadvantaged students in Italy substantially increased students' academic performance. Angrist et al. (2020) show that SMS and phone calls to parents minimise learning loss when school close.

more searches for online learning resources do not translate into better grades.

The remainder of the paper is organised as follows. Section 2 explains the necessary institutional background for our analysis and Section 3 describes the main data used for it. Section 4 presents the empirical strategy while Section 5 shows the findings for the impact of the first nationwide school closure on the change in online learning engagement. Section 6 provides descriptive evidence on the use of e-learning by regions before the COVID-19 outbreak. Section 7 shows additional results for the 2020-2021 academic year, when school closures were imposed depending on the local spread of the virus. Finally, Section 8 concludes.

2 Institutional Background

Italy was the first European country hit by the COVID-19 in 2020. The first case of the virus in the country was confirmed by January 31st, but both the intensity and speed of new cases were unequal across the country, leading to a highly regionalised impact, as reported by Giuliani et al. (2020). By February 23rd, the first schools started closing in the two most affected regions, Lombardy and Veneto (*zona rossa*) as well as in two neighbouring regions, Piedmont and Emilia-Romagna. On March 4th, *all* schools and universities across the country closed.[5]

Schools remained closed until the end of the academic 2019/2020 year. The starting date of the 2020-2021 school year differed across some of the Italian regions, with the majority of them starting on September 14th, 2021, and each following their own discretion on school closure mandates. The next meaningful legislative change that affected the development of schooling activity was enacted by the November 3rd, 2021 decree. The new decree established a new method to classify each region into three different categories according to its epidemiological risk – yellow, orange and red.[6]

These new measures imposed online learning only to grade 9 students and above in the two

[5] Five days later, on March 9th, the president declared a national lockdown. On March 11th, all commercial activity except for supermarkets and pharmacies were closed, and on March 21st, the Italian Government closed all non-essential businesses and industries and restricted the movement of people.

[6] Under each category, the Government implemented different measures to contain the spread of COVID-19. These measures mostly regulated social gatherings and events, and the ability to move across cities and regions. Thresholds in the value of specific epidemiological indicators measured at the regional level, such as relative COVID-19 active cases and the share of occupied beds in intensive care units, determined the changes across colour zones.

lowest risk zones and extended it to grade 7 students and above for the red zone. After the Christmas holidays, grade 9 students and above were allowed to go back to in-person schooling in yellow and orange zones. However, the number of students allowed in class was capped from 50 to 75 percent of the classroom's usual capacity. This implied that nine graders and older students were organised in a bi-weekly rotation scheme between in-person and e-learning during yellow and orange zones. Table A4 in the appendix summarises all the online learning mandates and their changes in the 2020-2021 school year.[7]

Together with the measures restricting mobility, at the beginning of the COVID-19 outbreak, the Italian Ministry of Education put specific measures in place to help teachers, students and families transit from face-to-face to e-learning. At the end of February 2020, the Minister of Education announced on the radio the program *Didattica a Distanza* (distance learning, in English). On March 4th, when *all* schools closed in the country and e-learning became mandatory, the Ministry of Education's website made available a new tab with dedicated training webinars and information on different platforms that were constantly updated. The website promoted three platforms: G Suite, provided by Google, (which includes Google Classroom and Google Meetings), Microsoft Office 365 provided by Microsoft, and WeSchool, provided by the Italian main communication company. While all these platforms already existed before the pandemic, their usage was scarce relative to the high popularity that they gained as a result of the COVID-19 outbreak.[8]

[7]In January 2021, a new lower colour category was introduced, "white", where most of the measures present in the yellow category would not be in place. For schooling activity, however, this new white zone imposed the same measures as those present in its subsequent higher category, yellow.

[8]Based on the data collected by Carlana and La Ferrara (2021) on 427 teachers in 76 middle schools all over Italy, by the month of June 2021 more than 96 percent of the teachers were providing synchronous online classes, and around 85 percent of the teachers provided some asynchronous videos additionally−usually no more than one hour per week. Right after the launch of the website, on March 26th, the Italian Ministry of Education passed the Ministerial Decree n.187, which allocated resources as follows: 1) 70 million euro to buy IT devices, such as tablets or computers, to lend temporarily to students in need, as well as to help these students improve their internet connection; 2) 10 million to allow schools to equip themselves with platforms and digital tools useful for distance learning and; c) 5 million euro to train teachers on methodologies and techniques for distance teaching. Due to bureaucracy delays, however, the help did not arrive to all in need.

3 Data

This Section introduces the data sources that are used in our analysis. Namely, Google Trends, INVALSI, and PISA data, as well as additional data used to control for the COVID-19 spread and other relevant regional level data.

Google Trends

We rely on Google Trends data to measure the engagement with online learning platforms in each of the Italian regions during COVID-19. Google Trends calculates the fraction of Google searches that are devoted to a given term relative to the overall Google searches within a given geographic area and time period. This ensures that places with the most search volume are not necessarily always ranked highest. Importantly, Google does not provide the actual fraction of searches but a scaled index ranging from 0 to 100. It assigns a value of 100 to the point in time and geographic area with the highest fraction value. We detail this and other particularities of Google Trends data in Appendix B.

A relevant part of our study is to choose platforms that are exclusively designed for e-learning, to avoid confounding between work-from-home and e-learning. For example, while Google Drive can be used by teachers to upload study material, it is also a commonly used cloud storage application by firms. Thus, in our data its increase in popularity during the pandemic would be attributed to a compound effect of the increase in work-from-home and e-learning. Taking this into consideration, we restrict our keyword list to 5 different platforms exclusively dedicated to e-learning: *Studenti.it*, *Scuola.net*, *Edmodo*, *Google Classroom* and *WeSchool*. It is important to note that the first two are fundamentally different from the last three:

Studenti.it is an Italian website for studying support, managed by the Italian schooling books publisher Mondalori Media S.p.A.. It is one of the most visited websites in Italy by high school students, university students and young people looking for training and employment. The website is constantly updated, and it provides students with the subjects of study of the current school year, study material to prepare for the exams, as well as plenty of practical information, including news from the Ministry of Education.

Scuola.net is a project of La Fabbrica. La Fabbrica is a training institution for teaching staff of the Italian school accredited by the Ministry of Education. It is a website dedicated to teachers of various school grades. A platform where they can participate in free educational initiatives and access solutions for digital teaching.

While these first two are websites where students and teachers can get informed about school and teaching related issues, *Edmodo, Google Classroom and WeSchool* are e-learning platforms themselves.[9] The three of them provide similar services, including allowing teachers to set assignments, have work submitted by students, mark, return graded essays, and distribute quizzes and surveys. In a time when all schools were suddenly forced to switch to online teaching, one would expect the use of the 3 e-learning platforms to experience the most dramatic increase – compared to the other two websites.[10]

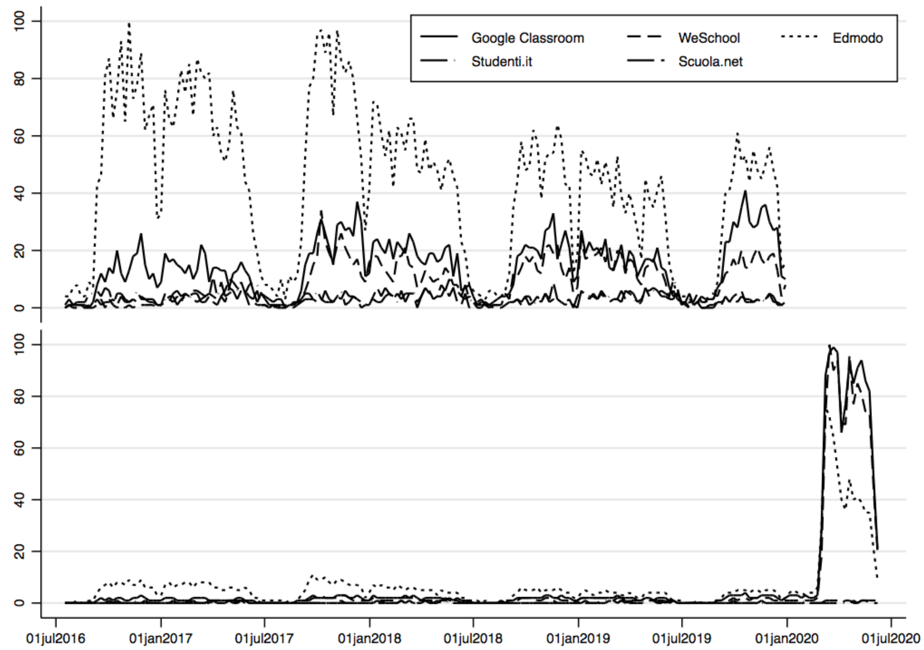
It is also worth mentioning that because of low search intensity Google Trends could not provide us with the search information on terms related to *Studenti.it* and *Scuola.net* in the two least populated Italian regions. Therefore, the time series referring to this two e-learning platforms contain slightly less observations than the those referring to the other three platforms.

We now show the raw Google Trends data for two purposes: First, to show the features of Google Trends' data and justify our choice regarding how to download our data set. Second, to validate the quality of the data. Each of the two graphs in Figure 1 shows 5 series over time, one for each of the 5 keywords described above. The two graphs in the figure correspond to two *different* downloads that differ only on the selected time window. The bottom panel contains the series downloaded for both the before-and-after the pandemic period (from June 2016 to June 2020), while the top panel plots the same data from June 2016 to December 2019. Given that we have exactly 5 keywords per graph, we downloaded the data set corresponding to each graph *all at once*. Thus, the values of this series are comparable *across* – not only within – the series.

[9]The website *didattica a distanza*, created by the Italian Ministry of Education as a way to help teachers and students to have a smoother transition into e-learning promoted three different platforms: G Suite, provided by Google (which includes Google Classroom and Google Meetings), Microsoft Office 365 (which includes Word, PowerPoint, Excel, Outlook and Teams), provided by Microsoft, and WeSchool, provided by the Italian main communication company.

[10]In fact, *Google Classroom* and *WeSchool* feature as the third and fourth most searched of all words in the list of ten trending words of Italy during the year 2020, only after *Coronavirus* and *Elezione USA* (USA elections) keywords, which take the first and second places respectively.

Figure 1: Google Searches in Italy for 5 selected keywords



Notes: This figure shows the data downloaded from Google Trends for the keywords Google Classroom, WeSchool, Edmodo, Studenti.it and Scuola.net setting the country of Italy as the geographic area. We download two bundles of 5 series each. The first bundle - top graph- contains series spanning from September 2016 to December 2019. The second bundle - bottom graph- spans from September 2016 to June 2020. Given that the series are downloaded in bundles, the series in each graph are comparable within and across across themselves.

The bottom graph in Figure 1 shows clear evidence of the dramatic increase of the online search for the three e-learning platforms in Italy right after the COVID-19 outbreak. This increase was lead by Google Classroom, which reached the highest value across all the keyword series on the week of 22-28th of March 2020, thus getting the value 100 in the graph. That same week, WeSchool was searched 91% and Edmodo was searched 60% as much as Google Classroom. Studenti.it and Scuola.net show an almost constant search behaviour over the entire time window *relative* to the other three platforms. In that same week, they were each searched 1% as much as Google Classroom.

Figure 1 helps to visualize the nature of Google Trends' data, especially when using its comparison feature by downloading the series in bundles. The top panel shows that when downloading the exact same bundle of series excluding the post-pandemic period, the series' variability increases. Thus, we consider that this figure justifies our choice on how to download our data set.

We also use the top graph in Figure 1 as supporting evidence that validates the use of Google

Trends data to understand the engagement of Italian students with online learning over time. The figure clearly shows that the index of search intensity follows the teaching activity periods along the academic year. The series experience a significant fall during the summer break and fall, to a lesser extent, during Christmas break and Easter holidays. While the level is highest for *Edmodo*, showing that it was the most searched e-learning platform in Italy before the pandemic, Google Classroom and WeSchool followed the same pattern.

Finally, as a further check on the validity of Google Trends' data, Figure A1 shows that Google Trends is a good predictor of the sudden increase in the number of active Gmail users in Italy in the spring of 2020. We believe that together with Figure 1, this is convincing evidence of the validity of Google Trends' data as a measure of engagement in online learning in Italy.

INVALSI

To measure academic performance at the regional level, we use data collected by INVALSI, the National Institute for the Evaluation of the Education and Training System. It organizes yearly standardised tests to assess students' performance at primary school (2^{nd} and 5^{th} grades), at lower secondary school (8^{th} grade), and at higher secondary school (10^{th} and 13^{th} grades).

For the purpose of this paper, we focus on students evaluated in the 10^{th} grade, i.e. higher secondary education. First of all, as students go up on the education system, many of them have extra motivation to study to get access to university, for which there are national entry exams. Second, we give preference to the 10^{th} rather than the 13^{th} grade, as these are the students about to complete mandatory education.

In the 10^{th} grade, students are administered two tests, one on the subjects of reading and one on mathematics, by an external examiner. In Table 1 we present the regional rankings of the 2018-2019 academic year.[11] We observe the classic North-South divide for both reading and mathematics. Table 1 shows evidence that all regions below the median of both tests are located below Emilia-Romagna. While the ranking position of each individual region is not the same in reading and mathematics, the bundle of regions that lie above the median is the same for both

[11] INVALSI grades are reported according to the WLE (Weighted likelihood estimates) of individual parameters of the Rasch model (Rasch, 1993) where 200 matches the national average.

subjects. In all our analysis we use the regional average INVALSI scores for the reading exam.

Table 1: Regional Average Scores in INVALSI for 10^{th} Graders

Region	Average reading	Ranking reading	Average math	Ranking math
Valle d'Aosta	218	1	224	1
Lombardia	217	2	220	2
Trento	217	3	218	3
Veneto	216	4	217	4
Friuli-Venezia Giulia	213	5	214	5
Emilia-Romagna	211	6	212	6
Piemonte	210	7	212	7
Marche	210	8	211	8
Liguria	206	9	207	9
Bolzano	206	10	207	10
Umbria	206	11	206	11
Lazio	205	12	205	12
Abruzzo	204	13	203	13
Toscana	203	14	200	14
Molise	199	15	198	15
Basilicata	196	16	194	16
Puglia	196	17	193	17
Campania	192	18	188	18
Sicilia	192	19	184	19
Calabria	189	20	184	20
Sardegna	187	21	182	21

Notes: This table reports INVALSI regional average scores in the 2018-2019 school year. Data obtained from the 2019 INVALSI report.

COVID-19 and Other Data

We now describe the three control variables that we employ in our empirical strategy and their data sources. First, we control for the total number of COVID-19 cases reported daily for each region, provided by the Ministry of Health's website. Given that COVID-19 first and more severely hit the North of the country, we condition on the number of confirmed COVID-19 to account for different trends in the virus spread that would induce different searches in e-learning platforms. Note that all the regions closed all their schools at a similar time, less than 15 days apart, as explained in Section 2.

Second, we control for the regional share of households with internet access in 2019, obtained from the National Statistics Institute (ISTAT), and collected by the Annual Questionnaire of Multiscopes for households in Italy. Although virtually all Italian households live in areas covered by broadband internet – in 2017 the European Commission estimated that 99% of all Italian households lived in areas covered by fixed broadband (European Commission, 2017) – not all households use this service. Additionally, as can be seen in Figure A2, all territories have access to similar levels of average download internet speed levels.

Finally, we include a northern dummy which follows the ISTAT terminology for statistical purposes. This dummy takes the value one for Emilia-Romagna, Friuli-Venezia Giulia, Lombardy, Piedmont, Trentino-Alto Adige, Valle d'Aosta, and Veneto. Italy's North-South divide in terms of cultural, socioeconomic and labor market characteristics is well documented. Thus, this dummy accounts for the North-South differences in all these characteristics, which may in turn drive differences on academic performance and e-learning usage.

4 Empirical Strategy

This section lies down the empirical strategy used to analyse the impact that the 2020 school closures had on the change of e-learning platform usage by level of academic performance across regions. Given hat the first case of the COVID-19 virus in Italy was confirmed on January 31st 2020, regional school closures were implemented as follows: the regions of Piemonte, Emilia-Romagna, Lombardy and Veneto closed on February 23rd 2020, Marche and the province of Trento on February 24th, Liguria on March 1st, and on March 4th all the remaining regions closed their schools. Soon after each closure, teaching was moved to online platforms and schools across the country remained closed until the end of the 2019/2020 academic year.

We estimate the following model to study whether there were regional differences on the change in search intensity of e-learning platforms after the school closures by their academic performance:

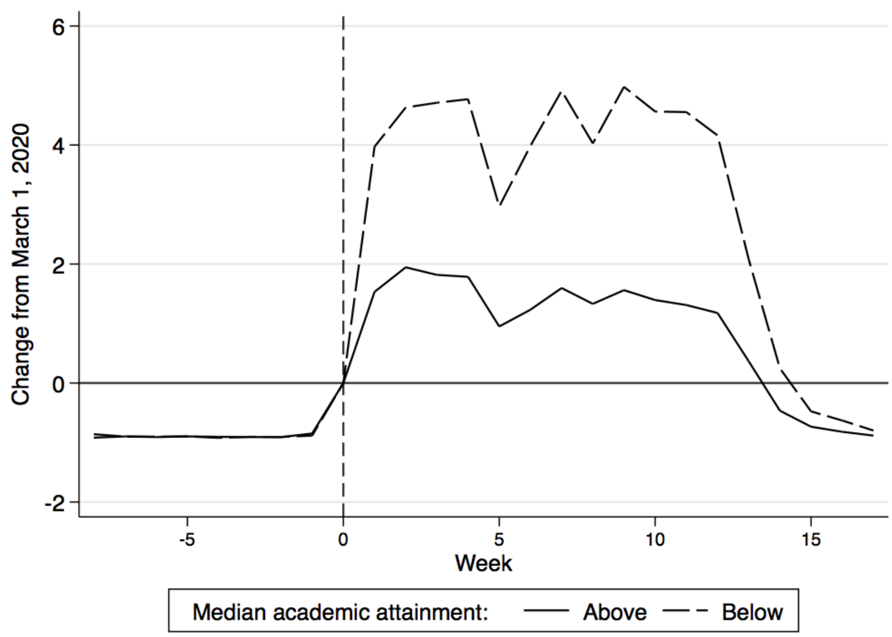
$$\ln(G.T.Index_{j,r,w}) = \alpha_0 + \alpha_1 \mathbb{1}AfterSchoolClosure_{r,w} + \beta_2 INVALSIScore_r +$$
$$+ \beta_3 \mathbb{1}AfterSchoolClosure_{r,w} \times INVALSIScore_r +$$
$$+ \gamma \ln(TotalCases_{r,w}) + X'\delta + \lambda_j + \phi_w + \epsilon_{j,r,w} \tag{1}$$

$\ln(G.T.Index_{j,r,w})$ is the log of Google Trends index for term j in region r in week w. Note that because the index includes zeros we shift it by one unit so that the dependent variable is defined for all weeks in our time window. $\mathbb{1}AfterSchool\ Closure_w$ is an indicator variable that takes value 1 after the week schools closed in region r and 0 before. $INVALSIScore_r$ represents the average score obtained in the 2019 INVALSI test for reading in region r. This variable has been standardised – i.e. demeaned and divided by its standard deviations. $\ln(TotalCases_{rw})$ is the total number of notified COVID-19 cases in region r in week w, to capture the potential increase in the need to use more e-learning rather than alternative in-person resources. X is a matrix of (time-invariant) regional characteristics, which includes: the share of households with internet access at home, to approximate internet usage and the total amount of terms searched in that region; and a dummy for whether the region is in the North of the country, to capture invariant regional characteristics of that part of the country, as well as the fact that they were firstly hit by the virus. To account for seasonality factors, fixed effects for the academic year and week of the year are introduced in ϕ_w. λ_j are platform fixed effects. Finally, $\epsilon_{j,r,w}$ is the error term. Our coefficient of interest is β_3, and it measures the effect of one standard deviation increase in INVALSI scores on the change of e-learning usage after schools closed relative to the period before school closures. Standard errors are clustered at the region level and bootstraped 1000 times to account for the low number of regions in our case study. All coefficients are weighted by the 2019 population values in each region to obtain nationally representative results, and the time window we use is between June 27th 2016 and June 7th 2020.

5 Results

We begin by providing a descriptive visualisation of the difference in online learning usage across regions with different levels of academic performance. Figure 2 depicts the average search indices for regions above and below the 2019 median INVALSI score within a time period of 25 weeks. It clearly illustrates that while academically high and low performing regions have a similar pattern both before and after school's closure, the increase in search intensity is substantially different, with regions below the median 2019 INVALSI score searching more than those above.

Figure 2: Google Trends Search Index for Google Classroom
by Academic Performance



Notes: This figure plots weekly changes of the Google Trends search index for the term *Google Classroom* in two groups of regions relative to March 1, 2020. Search index represented under below (above) the 2019 median INVALSI score contain the population weighted mean of the search index for the regions with a score in reading below (above) the national median. Regional mean scores in reading are extracted from the 2019 INVALSI report corresponding to Grade 10 students. Regional population shares used for the weights correspond to 2019 and are extracted from ISTAT.

Table A1 in the Appendix reports the results from an event study quantifying the total change in the usage of e-learning platforms due to school closures. The first column shows how, on average, regions increased the search of the e-learning platforms terms by 143%, relative to the period before school closures. Compared to southern regions, northern regions had on average a 7% higher increase in their searches of e-learning platforms during the entire period.

To perform a more exhaustive analysis, we estimate the regression equation (1) and present the results in Table 2. The first column pools all search data across the main five e-learning platforms at the national level, while results for each of them are shown in columns 2 to 6. In the third row of the first column we observe that after the closure of schools regions differed in their changes of e-learning platform searches depending on their previous academic performance in the 2019 INVALSI test. Specifically, we estimate that regions scoring one standard deviation above the average INVALSI score in reading had 19% lower change in their internet searches about e-learning platforms. As expected, regions that reported more COVID-19 cases are associated with higher levels of internet searches of e-learning platforms.

Table 2: Difference-in-Difference Results

	(1) All	(2) GC	(3) WS	(4) Ed	(5) Sc	(6) St
INVALSI Score	0.066	0.096	0.059	0.047	0.097	0.030
	(0.202)	(0.095)	(0.047)	(0.101)	(0.537)	(0.484)
After Regional Schools Closure	0.988***	2.436***	2.466***	2.212***	-1.085*	-1.337***
	(0.192)	(0.105)	(0.136)	(0.160)	(0.560)	(0.450)
After Regional Schools Closure * INVALSI Score	-0.188***	-0.285***	-0.193***	-0.088	-0.132	-0.268***
	(0.039)	(0.037)	(0.033)	(0.057)	(0.094)	(0.100)
North	0.002	0.106	0.124	0.159	-0.267	-0.113
	(0.249)	(0.130)	(0.078)	(0.191)	(0.650)	(0.621)
ln(COVID-19 Cases)	0.118***	0.092***	0.109***	0.006	0.182***	0.226***
	(0.021)	(0.014)	(0.015)	(0.019)	(0.063)	(0.054)
Share of Internet Access	0.009	0.011	-0.007	0.005	0.012	0.023
	(0.027)	(0.014)	(0.007)	(0.019)	(0.074)	(0.068)
Constant	0.684	-0.100	1.054*	1.145	0.931	0.391
	(2.092)	(1.078)	(0.569)	(1.421)	(5.734)	(5.230)
Observations	19,776	4,120	4,120	4,120	3,708	3,708
Platform FEs	Yes	-	-	-	-	-
Academic year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Week of the year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Bootstrap replications	1,000	1,000	1,000	1,000	1,000	1,000

Notes: This table reports the results from estimating equation 1 by ordinary least squares during the period June 27th 2016 to June 7th 2020. The dependent variable is the logarithm of the Google Search Index for selected E-learning platforms. *After Schools Closure* takes value 1 when schools closed in each region and 0 before. *INVALSI Score* represents the average score obtained in 2018 in the INVALSI test for Italian language. This variable has been standardised (demeaned and divided by its standard deviations) hence its units are standard deviations. *North* takes value 1 for Emilia-Romagna and all regions above it, and 0 otherwise. *Share of Internet Access* contains the share of households in each region that had internet access in 2019. *ln(COVID-19 Cases)* contains the total number of COVID-19 cases reported in each region and day. GC stands for Google Classroom, WS for WeSchool, Ed for Edmodo, Sc for Scuola.net and St for Studenti.it. All regression coefficients are weighted by each region's population and include fixed effects for each searched platform, academic year and week of year. Bootstrapped standard errors are clustered by region and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Importantly for the identification strategy, the results show that in the period before the school closures the regional academic performance was not an economic nor statistically significant factor associated with an increase of e-learning platform searches (between 3% and 9%), as indicated by the first coefficient of each row.

Analysing the difference-in-differences results by platform we conclude that the differential effect of academic performance on the change of e-learning platform searches after the school closures was driven by three of the five main platforms used in Italy, namely Google Classroom, WeSchool and Studenti.it. In contrast, previous academic performance was not detected to play a statistically significant role in the changes of searches related to Edmodo and Scoula.net platforms.

We find very similar results when using two alternative definitions of school closures, reported in Table A3. The first alternative definition uses the March 4th 2020 as the school closing date for all regions. The second one drops all observations between February 15th and March 15th 2020, and uses the latter date as the school closure date.[12]

With a similar data set but for the case of the U.S., Bacher-Hicks et al. (2021) show that economically more developed areas of the country saw substantially larger increases in search intensity for online learning platforms. Bearing in mind that areas of the United States with higher income are also areas with higher average SAT scores (Chetty et al., 2020), our analysis shows that the opposite effect took place in Italy. From their result the authors conclude that the pandemic will widen achievement gaps due to these areas' different engagement with online learning resources during the lockdown. Our findings, however, do not support the idea of differences in e-learning platform usage being a relevant factor when trying to explain the differences in immediate school outcomes.

The reasons why our analysis yields opposite findings to those reported for the U.S. are potentially related to the difference in the pre-pandemic usage of e-learning platforms across these two groups of regions. We analyse this for the case of Italy in the next section and we find that, before the outbreak of the pandemic, Italian regions with lower academic achievement were already using e-learning platforms more than those with higher academic achievement. Based on the pre-pandemic levels reported in their analysis and on the results from previous studies (see for instance Vigdor et al., 2014) we know that this was not the case in the U.S. There, students in high academically performing areas were already using e-learning resources more than those in

[12] Table A2 in the Appendix reports the same type of analysis for the event study.

low performing areas before the outbreak of COVID-19.

By acknowledging that in both countries the areas with higher preexisting usage rates of e-learning resources are the ones experiencing higher immediate increases, we are able to reconcile the results of both studies. Finally, we argue that the main mechanism behind this result is related to the cost of transitioning from in-person to online schooling. Students and teachers in areas with higher preexisting usage rates of e-learning resources were more likely to face a significantly lower transition cost compared to those with lower pre-pandemic usage rates. These lower costs could be related to the fact that the required software and hardware was already available in those regions, as well as to the higher skills and familiarity in using them of their teachers and students.

6 Online Learning Engagement before the Pandemic

Google Trends' Index values, allows us to study which set of regions *changed* the search intensity more as a result of the pandemic induced school closures. We just showed that contrary to findings in other studies, during the pandemic in Italy, academically lower performing regions *increased* the engagement with online learning platforms more than academically higher performing regions.

To interpret this result it is important to explore which of two opposite mechanisms, both consistent with our finding, is likely to have prevailed: 1) A catching-up-effect where academically lower-performing regions faced the COVID-19 outbreak with a lower *level* of engagement, and thus, had a bigger room for improvement; or 2) a gap-widening effect where academically lower-performing regions already had a higher engagement, and during the pandemic widened this gap even more.

For this, we need to compare the *levels* of engagement with e-learning before the pandemic across regions with different academic performances. Unfortunately, our Google Trends' Index data set does not allow to do so, and thus, we have to rely on other data sources.[13]

[13] As explained in Section 3, the value of the index for a given term in each of the series – corresponding to each of the regions – is a value relative to each series' own peak i.e if Lombardy takes the value of 70 and Campania takes the value of 50 on the index on a given date for a given term, it means that in that particular date, that term was searched 70% as much as in its most searched day in Lombardy and 50% as much as in its most searched day in Campania. We still do not know whether in that day and for that term, Lombardy had a higher search intensity than Campania or the opposite was true.

To analyse the relationship between academic performance and the *level* of online learning usage before the pandemic, we would ideally like to have the number of users and accesses, by region, to each of the three e-learning platforms and two websites that we use in our main analysis. Unfortunately, the data is not available. Thus, we have to rely on other data sources, and we use PISA and INVALSI as they are the two most complete surveys related to education in Italy. Taken together, they present a piece of consistent descriptive evidence that academically higher-performing regions were *not* using online learning more *before* the pandemic.

6.1 Use of e-learning tools before the pandemic by students

PISA (Programme for International Student Assessment) is an international standardised survey to 15-year-old students that comprises of a cognitive test on reading, mathematics and science, and complementary questionnaires to assess students' attitudes and motivations. Two surveys, the ICT Familiarity Questionnaire and the Educational Career Questionnaire are relevant to us. While the questionnaires have a very rich set of questions, the caveat of PISA is that not all the regions participate in every wave. We use PISA 2015 because it includes the better-suited regions for this study, Lombardy and Campania.[14] The two regions are among the most populated regions and have already been used as representative cases of the north-south divide in Italy in other studies (Acconcia and Graziano, 2017). We provide results comparing the two of them, where we use Lombardy as an example of the academically higher-performing regions of the North and Campania as an example of the lower-performing regions of the South.

From the various questions available, we focus on three that assess the ICT usage and availability outside school, as the availability and usage at school will be discussed in the data reported from teachers to INVALSI in the next subsection. Panel A in Table 3 reports differences in the usage of ICT resources for schoolwork and Panel B to attend additional instructions which are not part of students' mandatory school schedule.

[14]PISA 2015 provides data for Bolzano, Campania, Lombardy and Trento, while PISA 2018 provides data for Bolzano, Toscana, Sardegna and Trento. Note that both Bolzano and Trento (which form Trentino-Alto Adige) have a considerably lower share of publicly managed schools and therefore might be using e-learning differently than schools managed by the State. Excluding these two regions, PISA 2018 does not include any other region from the "above median performance" group we consider in our main analysis. Therefore, PISA 2015 is best suited for our analysis.

Table 3: ICT Usage in Selected Regions

Variable: Proportion of students	(1) Campania	(2) Lombardy	(3) Difference	(4) Italy
Panel A				
Outside school, at least once a week				
- for schoolwork	0.626	0.567	0.060***	0.591
	(0.013)	(0.013)	[0.001]	(0.009)
- to follow up school lessons	0.602	0.415	0.187***	0.504
	(0.014)	(0.013)	[0.000]	(0.009)
- for doing homework on computer	0.423	0.343	0.080***	0.362
	(0.014)	(0.012)	[0.000]	(0.009)
- for doing homework on mobile	0.416	0.266	0.150***	0.322
	(0.014)	(0.012)	[0.000]	(0.009)
Panel B				
Additional Math Instructions				
- Internet tutoring by a person or app	0.235	0.162	0.073***	0.185
	(0.017)	(0.016)	[0.002]	(0.011)
- Video recorded	0.168	0.069	0.099***	0.111
	(0.015)	(0.011)	[0.000]	(0.009)
Additional Italian Instructions				
- Internet tutoring by a person or app	0.275	0.226	0.049	0.263
	(0.020)	(0.023)	[0.112]	(0.016)
- Video recorded	0.155	0.103	0.052**	0.130
	(0.017)	(0.016)	[0.027]	(0.012)

Notes: The data reported in Panels A and B come from PISA 2015 ICT Familiarity Questionnaire and Educational Career Questionnaire respectively. Columns 1,2, and 4 report the proportion of students that answered positively to each of the metrics. Standard errors are reported in parenthesis. Column 3 reports the difference between Campania and Lombardy. The stars ,***,**,*, in this column indicate whether the difference is statistically significant at 1%,5%, and 10%, respectively. The p-values associated with the differences tests are reported in square brackets. All averages are weighted by the PISA final trimmed non-response adjusted student weights.

Panel A shows clear evidence that in the year 2015 students in Campania were already using e-learning technologies for schoolwork outside school more than students in Lombardy. Students in Campania were 10.4% more likely to use the internet for schoolwork, 45.1% more likely to use the internet to follow-up school lessons, 23.3% more likely to do their homework using a computer and 56.4% more likely to do them using a mobile phone. As reported in the third column of Table 3, all these differences are statistically significant at a 1% level. Panel B shows that in 2015 students in Campania were also more likely to use ICT in their additional instructions (not part of the student's mandatory school schedule) in both mathematics and reading.[15]

6.2 Use of e-learning tools before the pandemic by teachers

Together with the tests described in Section 3, INVALSI carries out surveys to students, teachers and school principals. The advantage of INVALSI over PISA is that the former includes a representative random sample of schools for every region in Italy. We begin by showing computer availability in schools by regions above and below the median 2019 INVALSI score.

Figure 3 plots the proportion of Italian language and mathematics teachers reporting having access to a computer and their usage in class during their lessons in every academic year between 2013 and 2019. The top figure suggests that until the academic year of 2017/2018, in higher academically performing regions a higher proportion of teachers had access to a computer to conduct their lessons. However, by 2018 and onward the two rates converged. The bottom figure shows the same pattern for computer usage and confirms that, in the last two academic years, there has been no difference in computer usage between low and high academic performing regions.

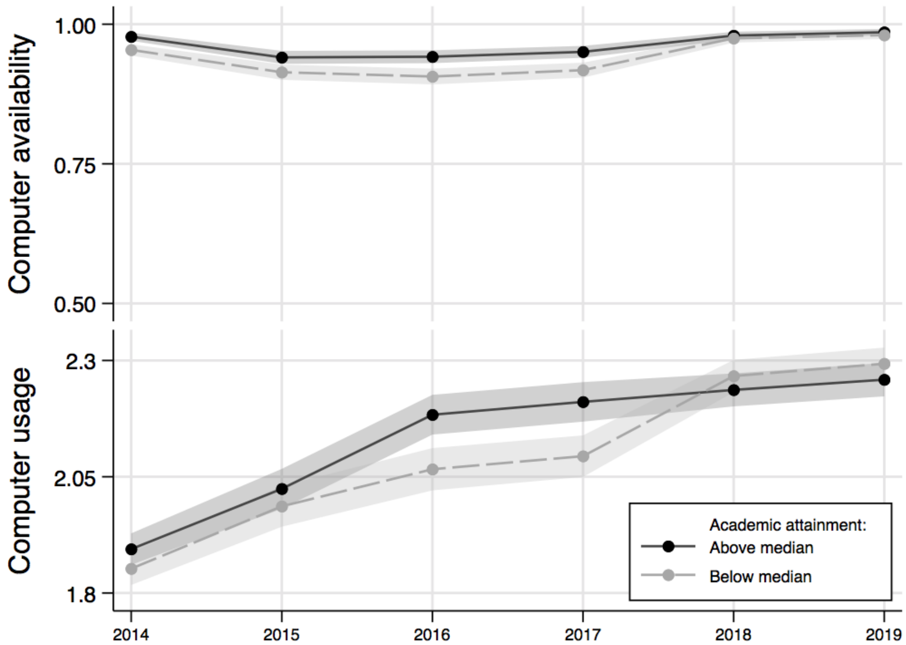
Unfortunately, the INVALSI surveys to students do not ask about their use of online tools outside school. However, the teachers' questionnaire includes a question of our interest: "Thinking about the didactic activity you carried out this year, please indicate how often you carried out the following activities: use of e-learning platforms.", the response options being 0 =Never or almost never; 1 = Sometimes; 2 = Often; 3 = Always or almost always.

Figure 4 plots Grade 10 teachers' reported usage of e-learning platforms when conducting their didactic activity in academic year 2018/2019 in reading and mathematics classes by their INVALSI score on those subjects in that year. Results show that lower academic performance regions are the ones associated with a greater level of e-learning usage in school by teachers.

Summing up, with these two data, we show that lower-performing regions were not using online learning tools at a lower level than higher-performing regions before the pandemic, allowing us to reconcile our results with those found by Bacher-Hicks et al. (2021) in the U.S.

[15]Despite Campania being a much poorer region than Lombardy, one could wonder if results are driven by higher access to ICT, by students in Campania. We use the ICT Familiarity Questionnaire, which asks about device availability at home, and find that this is not the case.

Figure 3: Computer Availability and Usage by Teachers in Class by Academic Performance

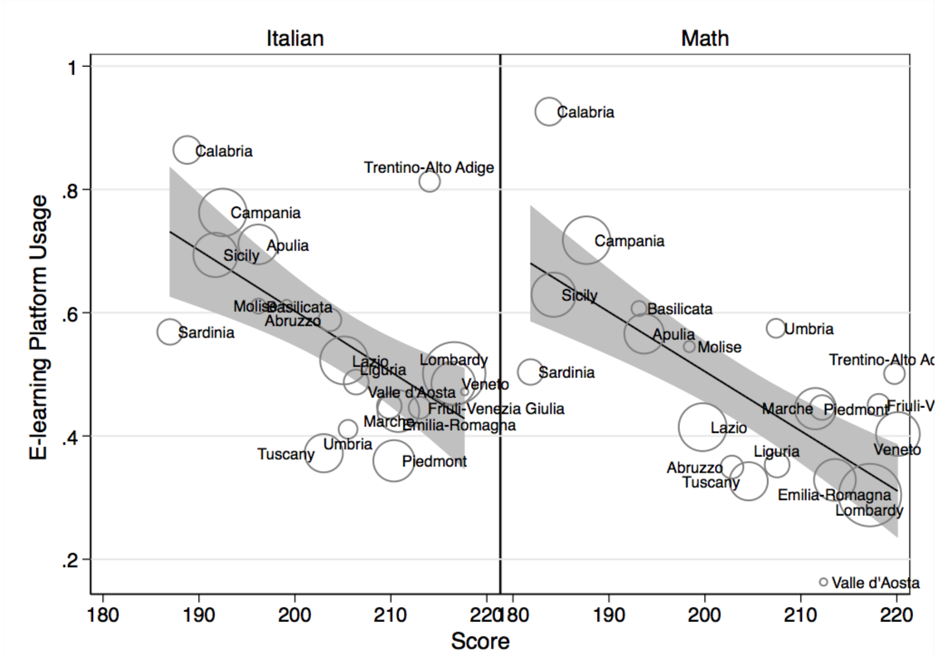


Notes: The figure plots the proportion of Italian language and mathematics teachers reporting having access to a computer, panel (a), and their usage, panel (b), in class during their lessons. Values are taken from a specific responses to question D6a administered by INVALSI to Grade 10 teachers of both subjects from 2013 to 2019, which states: *How much did you use the computer in lessons with the students of your class in the last school year?* Panel (a) plots one minus the share of teachers who responded *No computer present in school.* Panel (b) plots the group average of the following response options: *0 = No computer present in school; 1 = I don't use it; 2 = Occasional use; 3 = Regular use.* Below (above) the median contains the mean of the responses in the regions with a mean score in each subject below (above) the national median, respectively. Regional mean scores in both subjects are extracted from the 2019 INVALSI report corresponding to Grade 10 students.

7 Additional Results

As explained in section 2, from November 6th, 2020 onwards the Italian Government categorised regions in three categories using colours: yellow, orange and red, according to the different levels of the spread of the virus. Under each category, different measures were implemented to contain the spread of COVID-19 during the second wave. The new measures imposed online learning to grade 7 and above in the Red zone, while grade 9 students and above had to also follow their lessons online in the two lowest risk zones, yellow and orange. After the Christmas holidays, grade 9 students and above were allowed to go back to in-person schooling during yellow and orange. However, the number of students allowed in class was capped from 50 to 75 percent of the classroom's usual capacity. This implied that nine graders and older students were organised in a bi-weekly rotation scheme between in-person and e-learning during yellow and orange zones.

Figure 4: Teachers' E-learning Platform Usage by Students' Academic Performance



Notes: This figures shows the correlation between the reported usage of e-learning platforms by teachers when conducting their didactic activity in each region with the average results for the 2018/2019 INVALSI tests in reading and mathematics at Grade 10. The usage values for e-learning platforms are taken from the responses to question: *Thinking about the didactic activity you carried out this year, please indicate how often you carried out the following activities: e) use of e-learning platforms.* With the following response options: *0 = Never or almost never; 1 = Sometimes; 2 = Often; 3 = Always or almost always.* Sizes of circles correspond to the population share of each region, in 2019. The solid line corresponds to a linear fit weighted by the population share of each region. The shaded area corresponds to a 95% confidence interval of the linear fit.

Table A4 in the appendix summarises all the online learning mandates and their changes in the 2020-2021 school year. In this section we want to exploit the variation in the change of online learning platforms usage imposed by the new zone classification to analyse whether the pattern discovered after the first nationwide schools closure still persisted.

For this, we use Google Trends' data from the 2020-2021 academic year alone together with daily information on the assigned colour zone for each region. Importantly for our analysis, regions were declared at different moments and with different frequencies into the most restrictive category, the Red zone. Table A5, in the Appendix, summarises the descriptive statistics of the assignment of the regions to each of the colour zones. The data collection ranges from the beginning of the new zoning system, November 6th, to the end of the 2020-2021 academic year, June 18th.

In this analysis we seek to exploit the variation introduced by categorising regions into different colour zones to analyse two subjects. First, we are interested in testing the accuracy of our Google

Trends' measure as a proxy for e-learning usage by exploiting the regional variation in online learning mandates. Second, we want to understand whether the result found for the previous academic year, that is, that students in lower academic performing regions increased their usage of e-learning platforms more, is still present in the new course. To perform each of these two analyses, we pool daily Google Trends data on the three main e-learning platforms used across the different levels of education – Google Classroom, WeSchool and Edmodo – and estimate the two following specifications:

$$\ln(G.T.Index_{j,r,d}) = \alpha_0 + \alpha_1 \mathbb{1}RedZone_{r,d} + \alpha_2 \mathbb{1}OrangeZone_{r,d} + \beta_1 INVALSIScore_r \\ + \gamma_1 \ln(TotalCases_{r,d}) + X'\delta + \lambda_j + \phi_w + \epsilon_{j,r,d} \tag{2}$$

$$\ln(G.T.Index_{j,r,d}) = \alpha_0 + \sum_c \alpha_c \mathbb{1}ZoneC_{r,d} + \beta_2 INVALSIScore_r + \\ + \sum_c \delta_c \mathbb{1}ZoneC_{r,d} \times INVALSIScore_r + \\ +_1 \ln(TotalCases_{r,d}) + X'\gamma + \lambda_j + \phi_w + \epsilon_{j,r,d} \tag{3}$$

In equation (2) we are interested in measuring the correlation of changes of colour zones with changes in e-learning usage, measured by α_1 and α_2. $\mathbb{1}RedZone_{r,d}$ and $\mathbb{1}OrangeZone_{r,d}$ take value 1 if region r in day d was declared to be in the Red or Orange zones and zero otherwise, respectively. Since weekends and national holidays are removed from our studied sample, the base group of the colour indicator variables aggregate both the Yellow and the White zones. Thus, the base group is expected to contain the periods with low e-learning usage. λ_j are platform fixed effects and ϕ_w week of the year fixed effects. The rest of the variables are defined as explained in equation (1). Standard errors are clustered at the region level. We bootstrap the standard errors 1000 times to account for the low number of regions in our case study. All coefficients are weighted by the 2019 population values in each region to obtain nationally representative results.

In (3) we test if regions in the same colour zone, and therefore with the same online learning mandate, present a different change in their e-learning usage according to their average academic grade at the 2019 INVALSI test on reading. To do so, we interact the indicator variables associated with each region's colour zone and day ($\mathbb{1}ZoneC_{r,d}$ equals one if region r is in colour c, c being either red or orange at day d) with the standardised INVALSI score.

The first column in Table 4 reports the estimates of equation (2). As expected, compared with periods in which regions are declared Yellow or White, the change in the usage of e-learning increased more as online learning mandates were declared for a higher number of students; that is, when regions turned into orange or red zones, respectively. Also, it is no surprise that the change in search of e-learning resources is larger when changing to Red zone, 34.5 percent, than when doing it to Orange, 9.7 percent. These estimates are statistically significant at the 1 percent level. These first results confirm that our measure of changes in the Google searches of e-learning platforms is a proper proxy for the actual change in the platforms' usage.

Table 4: Difference-in-Differences Results - Academic Year 2020-2021

	(1)	(2)
INVALSI Score	0.004	-0.005
	(0.134)	(0.136)
Orange Zone	0.097***	0.093***
	(0.033)	(0.034)
Red Zone	0.345***	0.342***
	(0.105)	(0.123)
INVALSI Score x Orange Zone		-0.021
		(0.040)
INVALSI Score x Red Zone		0.065
		(0.100)
North	-0.407*	-0.406*
	(0.212)	(0.211)
ln(COVID-19 Cases)	0.679***	0.679***
	(0.055)	(0.056)
Share of Internet Access	-0.038*	-0.037*
	(0.021)	(0.021)
Constant	-2.672*	-2.735*
	(1.480)	(1.486)
Observations	8,160	8,160
Platform FEs	Yes	Yes
Week of the year FEs	Yes	Yes

Note: This table reports the results from estimating equation (2) and (3). The sample used contains daily observations from November 6th 2020 to June 7 2021, except for weekends and national holidays. The dependent variable is the logarithm of the Google Search Index for *Google Classroom*. *Red Zone* and *Orange Zone* take value 1 when a region is, respectively, red or orange zone in a certain day and 0 otherwise. *INVALSI Score* contains the regional average score of the 2019 INVALSI test in Italian. *North* takes value 1 for Emilia-Romagna and all regions above, and 0 otherwise. *Share of Internet Usage* contains the share of households in each region that used internet in 2019. *COVID-19 Cases* contains the total number of COVID-19 cases reported in each region and day. Bootstrapped standard errors are clustered by region and reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5 reproduces the same type of analysis using data only for one of the three main platforms each time. The largest change in the searches when entering in each new colour zone was coming from WeSchool platform, followed closely by the change in searches for Google Classroom. No statistically significant changes are detected for the changes in searches of Edmondo.

Table 5: Difference-in-Differences Results by Platform - Academic Year 2020-2021

	(1) GC	(2) GC	(3) WS	(4) WS	(5) Ed	(6) Ed
INVALSI Score	0.029	0.036	0.153	0.138	-0.171	-0.188
	(0.110)	(0.104)	(0.143)	(0.155)	(0.349)	(0.349)
Orange	0.093**	0.088**	0.168**	0.169**	0.029	0.021
	(0.037)	(0.036)	(0.071)	(0.076)	(0.086)	(0.094)
Red	0.228***	0.226**	0.522***	0.522***	0.285	0.280
	(0.078)	(0.090)	(0.158)	(0.176)	(0.189)	(0.207)
INVALSI Score x Orange		-0.043		0.021		-0.042
		(0.027)		(0.084)		(0.080)
INVALSI Score x Red		0.017		0.045		0.132
		(0.057)		(0.143)		(0.131)
North	-0.322***	-0.321***	-1.053***	-1.054***	0.155	0.157
	(0.113)	(0.111)	(0.251)	(0.253)	(0.499)	(0.499)
ln(COVID-19 Cases)	0.326***	0.327***	0.810***	0.809***	0.902***	0.901***
	(0.061)	(0.061)	(0.078)	(0.079)	(0.138)	(0.138)
Share of Internet Access	-0.012	-0.011	-0.082***	-0.082***	-0.020	-0.018
	(0.017)	(0.016)	(0.023)	(0.023)	(0.055)	(0.056)
Constant	0.424	0.361	-0.695	-0.695	-7.745*	-7.871*
	(1.375)	(1.351)	(1.775)	(1.800)	(4.098)	(4.123)
Observations	2,720	2,720	2,720	2,720	2,720	2,720
Week of the year FEs	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results from estimating equation (2) and (3) for each platform. The sample used contains daily observations from November 6th 2020 to June 7 2021, except for weekends and national holidays. The dependent variable is the logarithm of the Google Search Index for *Google Classroom*. *Red Zone* and *Orange Zone* take value 1 when a region is, respectively, red or orange zone in a certain day and 0 otherwise. *INVALSI Score* contains the regional average score of the 2019 INVALSI test in Italian. *North* takes value 1 for Emilia-Romagna and all regions above, and 0 otherwise. *Share of Internet Usage* contains the share of households in each region that used internet in 2019. *COVID-19 Cases* contains the total number of COVID-19 cases reported in each region and day. GC stands for Google Classroom, WS for WeSchool, Ed for Edmodo. Bootstrapped standard errors are clustered by region and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The second column reports the estimates of equation (3). We first observe how the estimated change in the search level of the average region in terms of the 2019 INVALSI score when the region is declared to be in the Orange or the Red zones are very similar to those in column 1. The coefficients of the interaction of each colour zone indicator variable with the 2019 INVALSI score are very small and not statistically different from 0. Implying that the regions with different levels of academic achievement did not change their searching behaviour differently during the establishment of each colour category.

We interpret this absence of different changes in e-learning usage between high and low academic performance regions during stricter school closure mandates as suggesting that, after half a year from the COVID-19 outbreak, all regions adapt their online learning behaviour in the same way. That is, the gap between higher and lower e-learning usage regions did not widen.

8 Conclusion

In this paper, we study whether academically high and low performing regions had a different response, in terms of changes in their e-learning usage, to schools' closure mandates imposed by the spread of the COVID-19 epidemic in Italy. We use real-time Google searches to measure the change in the use of several popular e-learning platforms. We divide our analysis into two periods: one spanning from September 2016 to June 2020, which includes the pre-covid time window and the time in which a nationwide school closure was implemented. A second one spanning from November 2020 to June 2021, in which lessons were carried out in-person or online intermittently depending on the local spread of the virus. To measure academic performance, we rely on pre-pandemic average standardised test scores in reading and mathematics administered by INVALSI.

We begin by using real-time Google search data to study the *change* in the use of online learning tools between regions with different academic performances due to the pandemic. We first document a substantial increase in the usage of e-learning platforms nationwide. Then, using a difference-in-differences specification, we find that regions with lower academic performance *increased* their search for online learning resources *more* after the nationwide school closure was implemented. We further document that previous academic performance was no longer a relevant factor determining changes in e-learning platform usage in the subsequent academic year. We interpret this result as evidence favouring all regions having the same online learning behaviour when faced with stricter school closure mandates during the subsequent academic year.

Several particularities of Google search data limit the policy recommendations that can be extracted from our analysis. Most notably, the fact that the Google Trends data is expressed using a scaled index rather than the actual search levels (or their share over total searches) is indeed its most restricting feature. In particular because it only allows us to speak about the magnitude of the change in Google searches on e-learning resources and not about their usage levels. In our case we rely on alternative data sources to document the level of online learning usage in low and high academically performing regions.

We overcome this limitation by relying on PISA and INVALSI surveys to document the *level* of online learning usage in regions with different average academic performances before the outbreak of the pandemic. We find that regions with a higher average academic achievement did *not* have a higher engagement in online learning in pre-COVID times. This allows us to argue that Italian regions with higher academic performance did *not* face the COVID-19 outbreak with greater familiarity in the use of online learning resources.

Our results, taken together, suggest that the first months of the pandemic contributed to widening the gap in the use of online learning resources between academically high and low performing regions in Italy. Combining different data, before 2020 and during the subsequent academic year, we have ruled out the channel of a lower engagement in online learning resources by students in lower-achieving regions. The empirical evidence in this paper suggests the need for a greater involvement of governments than just providing households and schools with access to online learning platforms.

References

Acconcia, G. and Graziano, P. R. (2017). *The Youth Guarantee: The Cases of Milan (Lombardy) and Naples (Campania)*, pages 339–348. Palgrave Macmillan UK.

Agasisti, T. and Vittadini, G. (2012). Regional economic disparities as determinants of student's achievement in italy. *Research in Applied Economics*, 4(2):33.

Andrew, A., Cattan, S., Costa Dias, M., Farquharson, C., Kraftman, L., Krutikova, S., Phimister, A., and Sevilla, A. (2020). Inequalities in children's experiences of home learning during the covid-19 lockdown in england. *Fiscal Studies*, 41(3):653–683.

Angrist, N., Bergman, P., and Matsheng, M. (2020). School's out: Experimental evidence on limiting learning loss using "low-tech" in a pandemic. Technical report, National Bureau of Economic Research.

Argentin, G. and Triventi, M. (2015). The north-south divide in school grading standards: New evidence from national assessments of the italian student population. *Italian Journal of Sociology of Education*, 7(2).

Bacher-Hicks, A., Goodman, J., and Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics*, 193:104345.

Bando, R., Gallego, F., Gertler, P., and Fonseca, D. R. (2017). Books or laptops? the effect of shifting from printed to digital delivery of educational content on learning. *Economics of Education Review*, 61:162–173.

Beuermann, D. W., Cristia, J., Cueto, S., Malamud, O., and Cruz-Aguayo, Y. (2015). One laptop per child at home: Short-term impacts from a randomized experiment in peru. *American Economic Journal: Applied Economics*, 7(2):53–80.

Brown, B. W. and Liedholm, C. E. (2002). Can web courses replace the classroom in principles of microeconomics? *American Economic Review*, 92(2):444–448.

Carlana, M. and La Ferrara, E. (2021). Apart but Connected: Online Tutoring and Student Outcomes during the COVID-19 Pandemic. *HKS Working Paper*, RWP21-001.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2020). Income Segregation and Intergenerational Mobility Across Colleges in the United States*. *The Quarterly Journal of Economics*, 135(3):1567–1633.

Contini, D., Di Tommaso, M. L., Muratori, C., Piazzalunga, D., and Schiavon, L. (2022). The COVID-19 Pandemic and School Closure: Learning Loss in Mathematics in Primary Education. *SSRN Electronic Journal*.

Cristia, J., Ibarrarán, P., Cueto, S., Santiago, A., and Severín, E. (2017). Technology and child development: Evidence from the one laptop per child program. *American Economic Journal: Applied Economics*, 9(3):295–320.

Engzell, P., Frey, A., and Verhagen, M. D. (2021). Learning loss due to school closures during the covid-19 pandemic. *Proceedings of the National Academy of Sciences*, 118(17).

European Commission (2017). Europe's digital progress report 2017: Connectivity. Technical report, European Commission.

Fairlie, R. W. and Robinson, J. (2013). Experimental evidence on the effects of home computers on academic achievement among schoolchildren. *American Economic Journal: Applied Economics*, 5(3):211–40.

Figlio, D., Rush, M., and Yin, L. (2013). Is it live or is it internet? experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, 31(4):763–784.

Giuliani, D., Dickson, M. M., Espa, G., and Santi, F. (2020). Modelling and predicting the spatio-temporal spread of covid-19 in italy. *BMC infectious diseases*, 20(1):1–10.

Grewenig, E., Lergetporer, P., Werner, K., Woessmann, L., and Zierow, L. (2021). Covid-19 and educational inequality: how school closures affect low-and high-achieving students. *European economic review*, 140:103920.

Joyce, T., Crockett, S., Jaeger, D. A., Altindag, O., and O'Connell, S. D. (2015). Does classroom time matter? *Economics of Education Review*, 46:64–77.

Lu, Y. and Song, H. (2020). The effect of educational technology on college students' labor market performance. *Journal of Population Economics*, 33(3):1101–1126.

Maldonado, J. E. and De Witte, K. (2021). The effect of school closures on standardised student test outcomes. *British Educational Research Journal*.

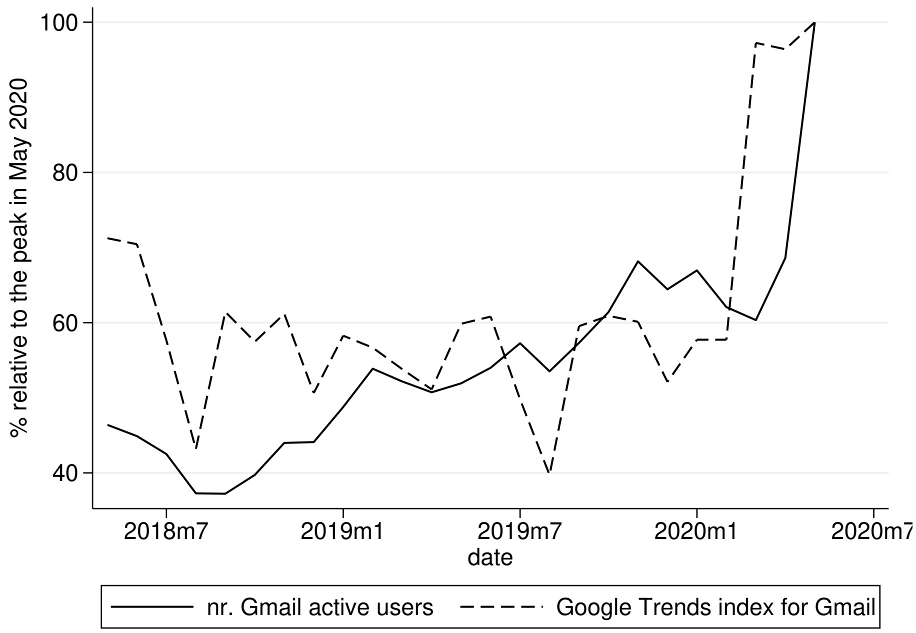
Rasch, G. (1993). *Probabilistic models for some intelligence and attainment tests.* ERIC.

Vigdor, J. L., Ladd, H. F., and Martinez, E. (2014). Scaling the Digital Divide: Home Computer Technology and Student Achievement. *Economic Inquiry*, 52(3):1103–1119. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecin.12089.

Appendices

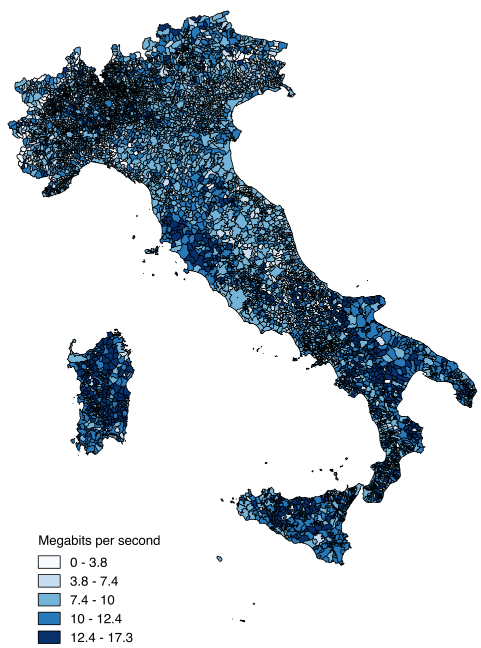
Appendix A: Appendix A: Supplementary Materials

Figure A1: Number of Active Gmail Users and its Google Trends Index



Notes: This figure plots the average monthly number of active users of Gmail, provided by AirnowData, and the average monthly Google Trends index for Gmail, between May 2018 and May 2020. Both series are rescaled relative to the peak in May 2020.

Figure A2: Geographic Distribution of ADSL Download Speed in 2018



Notes: This figure plots the average ADSL download speed in each Italian municipality in December 2018. Lighter colours indicate no data or low download speeds while darker colours represent higher average download speeds. Source: Autorità per le Garanzie nelle Comunicazioni (AGCOM).

Table A1: Results of Before-After Analysis on Google Search Index

	(1) All	(2) GC	(3) WS	(4) Ed	(5) Sc	(6) St
After Regional Schools Closure	1.428***	3.098***	2.917***	2.413***	-0.788**	-0.687**
	(0.198)	(0.218)	(0.209)	(0.198)	(0.385)	(0.297)
North	0.073***	0.209***	0.186***	0.213***	-0.155***	-0.091**
	(0.018)	(0.014)	(0.015)	(0.017)	(0.058)	(0.044)
ln(COVID-19 Cases)	0.062***	0.008	0.052**	-0.020	0.144***	0.144***
	(0.023)	(0.024)	(0.023)	(0.022)	(0.044)	(0.034)
Share of Internet Access	0.015***	0.020***	-0.002	0.010***	0.023***	0.025***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.006)	(0.005)
Constant	0.165	-0.848***	0.606***	0.750***	0.089	0.248
	(0.150)	(0.114)	(0.118)	(0.144)	(0.473)	(0.375)
Observations	19,776	4,120	4,120	4,120	3,708	3,708
Adjusted R-squared	0.482	0.888	0.881	0.809	0.218	0.240
Platform FEs	Yes	-	-	-	-	-
Academic year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Week of year FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results from an event study analysis using the period of June 27th 2016 to June 7th 2020. The dependent variable is the logarithm of the Google Search Index for selected E-learning platforms. *After Regional Schools Closure* takes value 1 when schools closed in each region and 0 before. *North* takes value 1 for Emilia-Romagna and all regions above it, and 0 otherwise. *Share of Internet Access* contains the share of households in each region that had internet access in 2019. *ln(COVID-19 Cases)* contains the total number of COVID-19 cases reported in each region and day. GC stands for Google Classroom, WS for WeSchool, Ed for Edmodo, Sc for Scuola.net and St for Studenti.it. All regression coefficients are weighted by each region's population and include fixed effects for each searched platform, academic year and week of year. Bootstrapped standard errors are clustered by region and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Results of Before-After Analysis on Google Search Index: Alternative School Closures

	(1) All	(2) All	(3) GC	(4) GC	(5) WS	(6) WS	(7) Ed	(8) Ed	(9) Sc	(10) Sc	(11) St	(12) St
After 4 March	1.242***		2.714***		2.537***		1.974***		-0.634*		-0.551*	
	(0.201)		(0.321)		(0.314)		(0.302)		(0.380)		(0.288)	
Before 15 Feb. after 15 Mar.		1.956***		4.131***		3.825***		2.963***		-0.901*		-0.580
		(0.259)		(0.132)		(0.123)		(0.162)		(0.496)		(0.379)
North	0.076***	0.088***	0.215***	0.229***	0.191***	0.204***	0.216***	0.228***	-0.155***	-0.148**	-0.091**	-0.076*
	(0.018)	(0.018)	(0.013)	(0.013)	(0.014)	(0.014)	(0.017)	(0.016)	(0.058)	(0.059)	(0.045)	(0.045)
ln(COVID-19 Cases)	0.085***	0.005	0.056	-0.100***	0.099***	-0.044***	0.032	-0.079***	0.126***	0.150***	0.128***	0.135***
	(0.023)	(0.030)	(0.037)	(0.016)	(0.036)	(0.014)	(0.035)	(0.019)	(0.043)	(0.055)	(0.033)	(0.042)
Share of Internet Access	0.015***	0.015***	0.020***	0.021***	-0.002	-0.001	0.010***	0.011***	0.023***	0.022***	0.025***	0.024***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.006)	(0.006)	(0.005)	(0.005)
Constant	0.172	0.125	-0.835***	-0.952***	0.619***	0.520***	0.769***	0.681***	0.083	0.109	0.243	0.293
	(0.150)	(0.150)	(0.115)	(0.105)	(0.117)	(0.110)	(0.145)	(0.138)	(0.473)	(0.479)	(0.375)	(0.378)
Observations	19,776	19,392	4,120	4,040	4,120	4,040	4,120	4,040	3,708	3,636	3,708	3,636
Adjusted R-squared	0.481	0.485	0.884	0.893	0.877	0.884	0.804	0.810	0.218	0.219	0.239	0.243
Platform FEs	Yes	Yes	-	-	-	-	-	-	-	-	-	-
Academic year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week of the year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results from an event study analysis using the period of June 27th 2016 to June 7th 2020. The dependent variable is the logarithm of the Google Search Index for selected E-learning platforms. *After March 4* takes value 1 after March 4 2020 and 0 before. *Before 15 Feb. After 15 March* takes value 1 after March 15 2020 and 0 before 15 February. *North* takes value 1 for Emilia-Romagna and all regions above it, and 0 otherwise. *Share of Internet Usage* contains the share of households in each region that had internet access in 2019. *ln(COVID-19 Cases)* contains the total number of COVID-19 cases reported in each region and day. GC stands for Google Classroom, WS for WeSchool, Ed for Edmodo, Sc for Scuola.net and St for Studenti.it. All regression coefficients are weighted by each region's population and include fixed effects for each searched platform, academic year and week of year. Bootstrapped standard errors are clustered by region and reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A3: Difference-in-Difference Results: Alternative School Closures

VARIABLES	(1) All	(2) All	(3) GC	(4) GC	(5) WS	(6) WS	(7) Ed	(8) Ed	(9) Sc	(10) Sc	(11) St	(12) St
INVALSI Score	0.067	0.066	0.099	0.093	0.062	0.057	0.051	0.047	0.095	0.107	0.029	0.024
	(0.201)	(0.203)	(0.095)	(0.097)	(0.047)	(0.048)	(0.102)	(0.103)	(0.536)	(0.532)	(0.484)	(0.485)
After 4 March	0.752***		1.959***		1.984***		1.625***		-0.851		-1.176***	
	(0.197)		(0.229)		(0.259)		(0.304)		(0.539)		(0.454)	
After 4 March * INVALSI Score	-0.216***		-0.335***		-0.244***		-0.156**		-0.103		-0.266***	
	(0.044)		(0.054)		(0.063)		(0.062)		(0.076)		(0.089)	
Before 15 Feb. after 15 Mar.		1.435***		3.564***		3.705***		3.122***		-1.939**		-2.045**
		(0.457)		(0.204)		(0.239)		(0.415)		(0.969)		(0.969)
Before 15 Feb. after 15 Mar. * INVALSI Score		-0.126**		-0.141***		-0.034		0.032		-0.244*		-0.330**
		(0.056)		(0.035)		(0.034)		(0.072)		(0.133)		(0.162)
North	0.002	0.010	0.107	0.116	0.125	0.134*	0.159	0.169	-0.269	-0.276	-0.115	-0.099
	(0.248)	(0.250)	(0.131)	(0.133)	(0.078)	(0.080)	(0.193)	(0.196)	(0.651)	(0.645)	(0.622)	(0.621)
ln(COVID-19 Cases)	0.146***	0.066	0.150***	-0.034	0.167***	-0.030	0.076**	-0.098*	0.153**	0.270**	0.206***	0.305***
	(0.022)	(0.052)	(0.024)	(0.025)	(0.026)	(0.026)	(0.037)	(0.050)	(0.067)	(0.111)	(0.053)	(0.109)
Share of Internet Access	0.009	0.008	0.010	0.011	-0.008	-0.007	0.005	0.005	0.012	0.011	0.023	0.023
	(0.027)	(0.027)	(0.014)	(0.014)	(0.007)	(0.008)	(0.019)	(0.019)	(0.075)	(0.074)	(0.068)	(0.068)
Constant	0.691	0.690	-0.081	-0.141	1.074*	1.032*	1.164	1.124	0.923	1.024	0.380	0.395
	(2.081)	(2.108)	(1.075)	(1.098)	(0.567)	(0.589)	(1.428)	(1.448)	(5.743)	(5.686)	(5.238)	(5.228)
Observations	19,776	19,776	4,120	4,120	4,120	4,120	4,120	4,120	3,708	3,708	3,708	3,708
Platform FEs	Yes	Yes	-	-	-	-	-	-	-	-	-	-
Academic year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week of the year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the results from estimating equation 1 by ordinary least squares during the period June 27th 2016 to June 7th 2020. The dependent variable is the logarithm of the Google Search Index for selected E-learning platforms. *After March 4* takes value 1 after March 4 2020 and 0 before. *Before 15 Feb. After 15 March* takes value 1 after March 15 2020 and 0 before 15 February. *INVALSI Score* represents the average score obtained in 2018 in the INVALSI test for Italian language. This variable has been standardised (demeaned and divided by its standard deviations) hence its units are standard deviations. *North* takes value 1 for Emilia-Romagna and all regions above it, and 0 otherwise. *Share of Internet Usage* contains the share of households in each region that had internet access in 2019. *ln(COVID-19 Cases)* contains the total number of COVID-19 cases reported in each region and day. GC stands for Google Classroom, WS for WeSchool, Ed for Edmodo, Sc for Scuola.net and St for Studenti.it. All regression coefficients are weighted by each region's population and include fixed effects for each searched platform, academic year and week of year. Bootstrapped standard errors are clustered by region and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Online Learning Mandates 2020-2021 School Year

Dates	Zone	*Scuola Primaria* Grades 1-5	*Scuola Secundaria di Primo Grado* Grade 6	Grades 7 and 8	*Scuola Secundaria di Secondo Grado* Grades 9-13
November 6, 2020 - January 6, 2021 (DPCM November 3, 2020)	Yellow	in-person	in-person	in-person	e-learning
	Orange	in-person	in-person	in-person	e-learning
	Red	in-person	in-person	e-learning	e-learning
January 7 - March 5, 2021 (DL January 5, 2021)	Yellow	in-person	in-person	in-person	50-75% in-person
	Orange	in-person	in-person	in-person	50-75% in-person
	Red	in-person	in-person	e-learning	e-learning
March 5 - School end (DPCM March 2, 2021)	Yellow	in-person	in-person	in-person	50-75% in-person
	Orange	in-person	in-person	in-person	50-75% in-person
	Red	e-learning	e-learning	e-learning	e-learning

Note: This table reports the changes in the online learning mandates that took place during the 2020-2021 school year. "50-75% in-person" means that 50 to 75 percent of the students were allowed to attend in-person lessons. From March 5, 2021 regions and autonomous provinces were allowed to impose a color increase within their territories if specific epidemiological conditions were met. Implying that different colors could be imposed within a region.

Table A5: Descriptive Statistics on the Colour System during Schooling Days of 2020/2021

Region	First Date Red Zone	Last Date Red Zone	Nr of Times Red Zone	Share of Days in Red	Orange	Yellow	White
Abruzzo	22/11/2020	05/04/2021	5	.16	.52	.31	.01
Apulia	24/12/2020	25/04/2021	4	.24	.35	.41	0
Basilicata	24/12/2020	05/04/2021	5	.13	.43	.45	0
Calabria	06/11/2020	11/04/2021	4	.22	.39	.4	0
Campania	15/11/2020	18/04/2021	5	.34	.2	.47	0
Emilia-Romagna	24/12/2020	11/04/2021	4	.18	.39	.43	0
Friuli-Venezia Giulia	24/12/2020	11/04/2021	4	.18	.29	.49	.04
Lazio	24/12/2020	05/04/2021	5	.13	.21	.66	0
Liguria	24/12/2020	05/04/2021	4	.06	.44	.49	.01
Lombardy	06/11/2020	11/04/2021	5	.32	.29	.4	0
Marche	24/12/2020	05/04/2021	4	.15	.35	.5	0
Molise	24/12/2020	05/04/2021	5	.16	.21	.59	.04
Piedmont	06/11/2020	11/04/2021	4	.28	.29	.42	0
Sardinia	24/12/2020	02/05/2021	5	.16	.34	.36	.14
Sicily	24/12/2020	05/04/2021	5	.13	.51	.36	0
Trentino-Alto Adige	10/11/2020	06/04/2021	6	.34	.38	.28	0
Tuscany	15/11/2020	11/04/2021	5	.21	.38	.41	0
Umbria	24/12/2020	05/04/2021	4	.06	.6	.33	.01
Valle d'Aosta	06/11/2020	09/05/2021	5	.35	.33	.33	0
Veneto	24/12/2020	06/04/2021	4	.15	.25	.59	.01

Note: This table reports the descriptive statistics of the colour system in Italian regions between November 6, 2020 and June 8, 2021. Trentino-Alto Adige takes the highest colour in the scale of the two autonomous provinces of Bolzano and Trento in order to make it compatible with the Google Trends data.

Appendix B: Appendix B: Particularities of Google Trends Data

Google Trends only allows it users to download a maximum of 5 time series at once. While all series need to refer to the same time period, they can vary in the term searched for and the geographic area of reference. The scaling of the fraction values is then performed for the time series that are downloaded together. This means that only those series that are scaled together have values readily comparable one to another. Because only 5 time series can be commonly scaled, we proceeded by downloading a single series for each term, region and time period. That is, we download a single series per term j region r and time window T, where the index $I_{j,r,t}$, constructed by Google Trends, is the ratio between the popularity of term j relative to the maximum popularity of that term over the time period T in geographic area r, measured on a 0 to 100 scale. It is calculated as follows:

$$I_{j,r,t} = 100 \frac{S_{j,r,t}/\sum_{i \in I} S_{i,r,t}}{\max_{t \in T}(S_{j,r,t}/\sum_{i \in I} S_{i,r,t})}$$

The numerator is measured as the ratio between the number of searches of term j in region r at point t ($S_{j,r,t}$) and the sum of searches for all terms $i \in I$ in that region and point in time ($\sum_{i \in I} S_{i,r,t}$). The denominator is the maximum of these ratios over the time period T for term j and region r. The index $I_{j,r,t}$ is the outcome variable of our regressions. Following Bacher-Hicks et al. (2021), we use the logarithm of Google Trends' index to interpret estimates as percent changes.

When extracting Google Trends data, one should note two characteristics of Google Trends. First, Google Trends uses a representative sample of all Google searches. This is very important, in particular, when extracting data for geographic areas with low search volume. We take this point seriously and download 20 different series for each term-region-time-period combination to make sure that we are using representative samples. However, Google Trends renews the publicly available sample at unknown times. To ensure that each downloaded series belongs to a different sample we use slightly different time periods – a couple of days apart – by region and keyword in each of the 20 downloads. This process can only be used when, as in our case, one is certain that the maximum value of the series will not lie at the beginning or the end of the downloaded series. Then, we average the results by term, region and point in time across all the 20 samples. Thus, the upper bound of our index is not necessarily 100.

Finally, Google Trends provides different frequency data, depending on the time span of the series that one wants to download. For series spanning 9 months or less, it provides daily frequency data. For series spanning more than 9 months, it only provides weekly frequency data. Given that the time window we are interested in spans from June 27th, 2016 to June 7th, 2020 – corresponding to the end of the academic year – we use weekly frequency. This frequency allows us to control for seasonality effects on the engagement with e-learning technology by including month and year fixed effects. The analysis performed for the 2020-2021 academic year spans from November 6th, 2020 to June 7th, 2021, and uses daily frequency data.