# Academic Performance and E-learning in Italy *

Josep Amer-Mestre
European University Institute

Alaitz Ayarza Astigarraga
European University Institute

Marta C. Lopes
Universidad Carlos III de Madrid, IZA, GLO

Wednesday 31st March, 2021

**Abstract**

Using data from Google Trends in Italy, we study how different regions reacted to the implementation of online learning through national-wide platforms, such as Google Classroom and WeSchool, due to the COVID-19 pandemic. Regions with a previously lower academic performance, actively searched more in Google about e-learning tools, surpassing regions with earlier school closures. Analysing school administrative data before the pandemic we find that teachers in lower performing regions were already using other e-learning tools more than higher performing regions. Complementing this with survey data for students we find that the ones in lower performing regions were also relying more on internet to follow up school lessons and attend additional tutorials. Unlike studies in other countries, our findings suggest that the COVID-19 shock may not change the lower academic performance regions behaviour in their usage of e-learning.

**JEL classification**: C81, I24
**Keywords**: E-learning, COVID-19, Google Trends, Inequality

*We thank Andrea Ichino, Thomas Crossley, Anna Lavizzari, Eleonora Erittu, Pedro Luís Silva, Pedro Freitas and Jan Stuhler for very helpful comments and suggestions. We also thank the very useful discussions at the internal seminars of European University Institute and Universidad Carlos III de Madrid. We gratefully acknowledge financial support from the European University Institute COVID-19 Knowledge Hub. Corresponding author: marta.cardoso@uc3m.es. Declarations of interest for all three authors: none

# 1 Introduction

School closures were the primary response by Governments in most affected countries by COVID-19 to control the virus' spread in Spring 2020 (ACAPS, 2020). Educational institutions were suddenly forced to undergo a quick transition from face-to-face to online teaching. Around the world, concerns were raised about potential differences in online learning usage by students with different socio-economic conditions. See Bacher-Hicks et al. (2021) and Rodríguez-Planas (2020) for analyses in the United States, Andrew et al. (2020) for England, and Engzell et al. (2020) for an evaluation in primary schools in the Netherlands.

In this paper, we look for differences in e-learning usage by academic performance. Students with better grades should get higher income returns in the future (Murnane et al., 1995). We want to test whether the sudden change in learning methods will be a contributor to potentially increase social mobility or, instead, contribute to the widening of the socio-economic inequalities.

As a case study, we focus on Italy. This country, albeit unexplored, is an interesting case for three main reasons. First, according to the COVID-19 Government Measures Dataset (ACAPS, 2020), Italy was the first country to close schools outside Asia (the epicentre of the virus outbreak). Second, Italy is a country that presents substantial regional differences in school quality (Brunello and Checchi, 2005) and academic performance (Agasisti and Vittadini, 2012). Third and finally, the centralised school management in the country put forward a website (*didattica a distanza*) established together with the school closures, to support schools in implementing online learning methods, which enable us to compare regions on the usage of nationally implemented e-learning platforms.

To measure the usage of e-learning platforms, we use real-time data on e-learning tools interest via Google Trends for each Italian region from January to June 2020 (the Spring term during which national closure of schools was implemented). To measure academic performance, we use pre-pandemic regional level of test scores in Italian and Math, performed by the National Institute for Education Evaluation and Training (INVALSI).

We find that regions that performed better before the pandemic searched relatively less for e-learning tools via Google during the first lockdown. While this finding could potentially suggest a catching-up effect from the regions with lower academic performance, we actually find that these regions were already using more e-learning before the pandemic, when we analyse older platform's usage. Exploring e-learning related questions in INVALSI and PISA (OECD Programme for International Student Assessment) at the regional level, we find that these regions were already using more Information and Technology (IT) both in class and outside school, thus suggesting that any regional differences in academic performance and e-learning usage should perpetuate in Italy. In fact when, we test for regional differences in online e-learning platforms that were popular before and after the pandemic (*Scuola.net* and *Studenti.it*) we find no statistical differences.

Our results, therefore, contrast with the findings for the United States, reported by Bacher-

Hicks et al. (2021). The authors find that areas of the country with higher income, better internet access and fewer rural schools saw substantially larger increases in search intensity. We find that the Italian regions with a comparative *disadvantage* in terms of academic performance *increased* the search intensity relatively more. Additionally to the literature that studies regional academic inequalities during COVID-19 lockdowns, our paper also contributes to two other branches of the economic literature: Google Trends, and regional differences in Italy.

Most of this economic literature using Google Trends has been focusing on Macroeconomic and Financial indicators: see, for example, Vosen and Schmidt (2011) for consumption, Baker and Fradkin (2017) for unemployment insurance, Castelnuovo and Tran (2017) for uncertainty, Dergiades et al. (2015) for sovereign spreads and Hamid and Heiden (2015) for volatility in the stock market. This paper instead uses the data at the regional level, which has been recently more popular on the study of potential regional inequalities due to the COVID-19 pandemic (see Bacher-Hicks et al. (2021), for example). In this paper, we call attention for two important features of the data that are often disregarded: 1) the comparison of levels of the index between regions, and 2) the sampling feature of Google Trends. The later is particularly important in the case of smaller countries.

The economic literature that studies regional differences in Italy found evidence for differences in economic development (Peracchi, 2008), school quality (Brunello and Checchi, 2005), and academic performance (Agasisti and Vittadini, 2012), but no study, to the best of our knowledge, has documented the regional differences in the usage of technology at schools. This finding is particularly important as the economic literature is divided between none (Fairlie and Robinson, 2013; Beuermann et al., 2015; Cristia et al., 2017; Bando et al., 2017) and negative association (Brown and Liedholm, 2002; Joyce et al., 2015) between e-learning and academic performance.

The paper is organised as follows. Section 2 provides a contextual background to how the Italian Ministry of Education implemented e-learning in the country during the first COVID-19 lockdown. Section 3 describes the main data used in this study. Section 4 explains the methodology that originates the results in Section 5. Section 6 provides additional empirical evidence on the differences in e-learning for two platforms that were are widely used before the pandemic. Section 7 analysis microdata for differences in e-learning before the pandemic. Finally, section 8 concludes.

## 2 Contextual Background

Italy was the first European country to be hit by the COVID-19 in 2020. The first case of the virus in Italy was confirmed by the January 31, but both intensity and speed of new cases were unequal across the country, thus leading to a highly regionalised impact, as reported by Giuliani et al. (2020). The first schools to close were the ones in the two most affected regions, Lombardy and Veneto (*zona rossa*), on February 23.

On March 4, Italy ordered the closure of *all* schools and universities. Five days later, on March 9, the president declared a national lockdown. On March 11, all commercial activity except for supermarkets and pharmacies were closed, and on March 21, the Italian government closed all non-essential businesses and industries and restricted movement of people.

The program *Didattica a Distanza* (distance learning, in English) was firstly announced on the radio by the Minister of Education, Lucia Azzolina, at the end of February. The details were not provided immediately but on March 4, e-learning became mandatory for all schools in Italy. Since that date and throughout the Spring term, the Ministry of Education website made available a new tab with dedicated training webinars and available information on different platforms that were constantly updated.

Eventually the website stabilised in three 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.[1] While all these platforms already existed before the pandemic, their usage was relatively scarce, but quickly became popular with the program of *Didattica a Distanza*, as we show in the next section.

# 3 Data

## 3.1 Google Trends

To measure the usage of online learning platforms in each of the Italian regions during the first COVID-19 lockdown in Italy, we rely on the frequency of searches in Google about terms associated with online learning. Google search Trends has already proved to be a *sufficiently representative* data source for job search Baker and Fradkin (2017), and a good predictor of youth unemployment rate in Italy Naccarato et al. (2018). Although we do not have available data to contrast the usage of the e-learning platforms with the Google Trends index, we can also see in Figure A1 that Google Trends is a good predictor of the jump in the number of active Gmail users in Italy in the spring of 2020. Therefore, we are confident that this source should be a good proxy to measure the increase in e-learning usage in Italy during the first COVID-19 lockdown.

When using Google Trends data, the researcher should be aware of three main characteristics: 1) for each different query, Google samples a number of searches; 2) the indices are relative to the maximum interest of a term for a certain region during the requested period, therefore the results from two queries with different regions should not be compared in levels; 3) Google Trends only allows to introduce five different queries at the same time that will be referenced to the same value and therefore comparable (in levels).

[1] WeSchool, on top of providing an alternative for online classes, also offered the tool of online quizzes.

To circumvent the last point, previous studies suggested either to anchor each bundle of regions on a first region, i.e., keep a fixed region in each bundle of five regions Goldsmith-Pinkham and Sojourner (2020). However, given that the re-sampling feature of Google reveals substantially different values in our queries, we cannot be sure that the region used as an anchor is representing the same raw data in different bundles. Alternatively, one could use an anchor that has always the value 100 so that all the bundles are on the same scale. However, this approach also brings problems. First, we could not identify a region for the terms we search that always presents the highest value among all regions in all samples. Second, finding a new term-region combination that is always higher than all other searches for all samples leads to a decrease in the variation observed and in the level differences between regions.

In this paper, we are mostly interested on the change over time across regions, therefore we use the indices from separated queries instead of the comparison feature to capture as much time variation as possible. To take into account the sample feature we queried each sample 20 times by slightly moving the the limits of the period window. This is a faster trick to obtain 20 samples, available when the peak of the series is far away from the limits of the period window.

The (common) period selected for the analysis in all regions is between January 6 and June 7 to guarantee the maximum common interval of complete weeks in the academic calendar across the different regions.[2] As the selected period comprises less than 9 months, the series obtained has daily frequency.

For each day in the series of term j, in a given sample, we have the following:

$$I_{j,r,d} = 100 \frac{S_{j,r,d}/\sum_i S_{i,r,d}}{\max_{d \in [Jan6, June7]}(S_{j,r,d}/\sum_i S_{i,r,d})}$$

, that is the ratio between the popularity of term $j$ relative to the maximum popularity of that term over the time period in the query, measured on a 0 to 100 scale. The first one is measured as the ratio between the number of searches of term $j$ in region $r$ in day $d$ ($S_{j,r,d}$) and the sum of searches for all terms in that region and day ($\sum_i S_{i,r,d}$). The second is the maximum of these ratios over the time period in the query for that region. Note that after downloading all the series for 20 different samples, we averaged out the results for each day, region, and term, and therefore the upper bound of each of our series is not 100.

### 3.1.1 Selection of Keywords

To avoid confounding teleworking and e-learning, a key point in our study is to choose platforms that are exclusively designed for e-learning. For example, while Google Drive can be used by

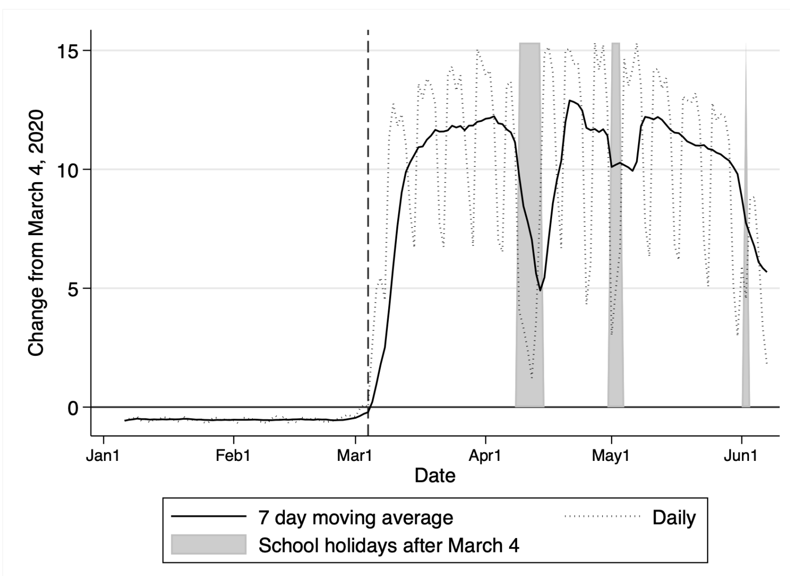
[2]Note however that there is not much variation in the start and ending dates as the regions that started the classes earlier, after the Christmas break, were Valle d'Aosta and Molise, which begun on January 4, while the region that finished later was Bolzano that ended on June 16.

teachers to upload study material, it is also a commonly used application by firms. Thus, its increase in popularity during the pandemic would be attributed to a compound effect of the increase in teleworking and e-learning that our data does not allow to disentangle.

Taking this into consideration, we restricted our keyword list to Google Classroom and WeSchool. During the period in analysis, Google Classroom was the most popular platform, according to the Google Trends Index for Italy as a whole. Even though WeSchool was only searched 64% of the times before the schools' closure, its popularity raised to being 85% as popular as Google Classroom, after March 4. Even though both platforms were searched at different rates, they had a similar pattern over time. To focus on the evolution of Google Searches over time we plot the Google Trends index for Google Classroom as this platform was more popular.

Figure 1 plots the evolution of the search intensity for Google Classroom during the period of our analysis. In dots we have the original daily series and in full line we plot the 7 days moving average to highlight the main changes across weeks. The graph shows that the Google Searches over this term substantially rose after March 4, declined temporarily over the break of Easter and the bank holiday of May 1, and finally decayed smoothly as the schooling period ended. The overall trend of the Google Trends Index indicates that the searches are more likely to proxy for the continuous usage of the platform, rather than the installation and information gathering of the platform.

Figure 1: Google Trends Search Index for Google Classroom in Italy




Note: This figure plots daily changes of the Google Trends search index in Italy for the term *Google Classroom* relative to March 4, 2020. The solid line corresponds to the 7 day moving average computed by taking the average of each daily value together with the values from the preceding 6 days. Shaded areas mark the official school holidays including the adjacent weekends that happened after the national closing of schools on March 4, date represented by the dashed line.

### 3.2 INVALSI

To measure academic performance at the regional level, we use data from INVALSI, the National Institute for the Valuation of the Education and Training System.[3] This institute organizes yearly standardised tests to assess students' performance at primary school (2$^{\text{nd}}$ and 5$^{\text{th}}$ grades), at lower secondary school (8$^{\text{th}}$ grade), and at higher secondary school (10$^{\text{th}}$ and 13$^{\text{th}}$ grades).

For the purpose of this paper, we focus on the students evaluated at the 10$^{\text{th}}$ grade. First of all, in Italy, there are more students enrolled in secondary schools than in primary school. Second, as students go up on the education system, many of them have extra motivation to study to get access to university. Third and finally, we give preference to the 10$^{\text{th}}$ rather than the 13$^{\text{th}}$ grade, as these are the students that are about to complete mandatory education.[4]

At the 10$^{\text{th}}$ grade, two tests are administered to all students in the presence of an external examiner: Italian and Mathematics. In Table 1 we present the regional rankings in the academic year of 2018/2019.

Table 1: Regional Average Grades in INVALSI 10$^{\text{th}}$ Grade Tests

| Region | Average Italian | Ranking Italian | Average Math | Ranking Math |
|---|---|---|---|---|
| Lombardy | 213 | 1 | 215 | 2 |
| Veneto | 213 | 2 | 216 | 1 |
| Friuli-Venezia Giulia | 209 | 3 | 214 | 3 |
| Valle d'Aosta | 208 | 5 | 205 | 10 |
| Trentino-Alto Adige | 208 | 4 | 211 | 4 |
| Emilia-Romagna | 207 | 6 | 210 | 5 |
| Piedmont | 206 | 7 | 207 | 7 |
| Umbria | 205 | 9 | 207 | 8 |
| Liguria | 205 | 8 | 206 | 9 |
| Marche | 204 | 10 | 208 | 6 |
| Tuscany | 200 | 11 | 203 | 11 |
| Abruzzo | 199 | 12 | 200 | 12 |
| Lazio | 198 | 13 | 196 | 14 |
| Basilicata | 196 | 14 | 196 | 13 |
| Molise | 194 | 15 | 195 | 15 |
| Apulia | 193 | 16 | 191 | 16 |
| Campania | 189 | 17 | 186 | 17 |
| Sicily | 187 | 18 | 184 | 18 |
| Sardinia | 183 | 19 | 178 | 19 |
| Calabria | 181 | 20 | 176 | 20 |

This table reports the regional average grades for the academic year 2018/2019. The dashed line divides the regions that are above and below the median across regional average grades.

[3]INVALSI Data is available at http://invalsi-serviziostatistico.cineca.it/ upon request. Methodological information is available at: http://invalsi-areaprove.cineca.it.

[4]For the purpose of categorising regions according to the academic performance it makes no statistical difference between using the 8$^{\text{th}}$, 10$^{\text{th}}$, or 13$^{\text{th}}$ grades as their correlation at the regional level is approximately 85%.

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. We observe that regions in the South of the country tend to perform worse in both subjects but the ranking in Italian is not the same as in Mathematics. However, even if regions change position on each subject's ranking, none of them changes substantially relative to the median.

### 3.3 COVID-19 Cases and Other Regional Data

Before providing additional details on the empirical strategy we describe here the source of our three crucial control variables. First, we control for the total number of COVID-19 cases reported daily for each region, provided by the Italian Ministry of Health on their website. Given that the North of the country was firstly and more severely hit by COVID-19, we want to clean our analysis from different trends in the virus spread that would induce different searches in e-learning platforms.

Second, we control for the share of home internet usage by region in 2019, obtained from ISTAT (National Statistics Institute), and collected by the Annual Questionnaire of Multiscopes for households in Italy. Although virtually all Italian households live in areas that are covered by broadband internet, not all households consume this service.[5] Within those that use internet we do not need to further control for internet speed as we can see in Figure A2 that all territories have access to similar levels of average download 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 This control is particularly useful as we exclude from our analysis other cultural characteristics that might drive differences in academic performance and e-learning usage.

## 4 Empirical Strategy

The time window for all the specifications in our main analysis is between the January 6 and June 7, 2020. To estimate the average effect of COVID-19 on the access to e-learning platforms across regions we perform a simple before and after analysis relative to the date of schools closure:

$$\begin{aligned}\ln(1 + G.T.Index_{j,r,d}) = \alpha_0 + \alpha_1 \mathbb{1}AfterSchoolClosure_d + \\ + \gamma \ln(1 + TotalCases_{rd}) + X'\delta + \lambda_j + \epsilon_{j,r,d} \qquad (1)\end{aligned}$$

$\ln(1 + G.T.Index_{j,r,d})$ is the log of Google Trends index for term $j$ in region $r$ in day $d$. Note

[5] In 2017 the European Commission estimated that 99% of all Italian households lived in areas covered by fixed broadband, Commission (2017).

that because the index includes zeros we shift it by one unit so that the dependent variable is defined. $\mathbb{1}AfterSchool\ Closure_d$ is an indicator variable that takes the value 1 for the days after March 4 and 0 otherwise. $\ln(TotalCases_{rd})$ is the total number of COVID-19 cases for region $r$ in day $d$, to capture the potential increase in the need to use more e-learning rather than alternative live sources. X is a matrix of regional characteristics which includes the percentage of students in the population, to capture the potential relative interest for searching these words in Google; the share of the population using internet, to capture amount of terms searched in that region; and a dummy for whether the region is located in the South of the country to capture other culture characteristics that are common across regions, as well as the fact that the North of the country was firstly hit by the virus. $\delta$ is the vector of coefficients associated with the regional variables. $\lambda_j$ are searched terms fixed effects. Finally, $\epsilon_{j,r,d}$ is the error term. The main coefficient of interest in this regression is given by $\alpha_1$, which indicates the average increase in the search intensity across all the regions in the period after the school closures.

To study whether there were regional differences on the search intensity change after the school closures we split the regression above by academic performance, with the following difference-in-differences specification:

$$\begin{aligned}\ln\left(1+G.T.Index_{j,r,d}\right) = \alpha_0 + \alpha_1 \mathbb{1}AfterSchoolClosure_d + \beta_2 \ln(INVALSIScore_r) + \\ + \beta_3 \mathbb{1}AfterSchoolClosure_d \times \ln(INVALSIScore_r) + \\ + \gamma \ln(1+TotalCases_{rd}) + X'\delta + \lambda_j + \epsilon_{j,r,d}\end{aligned} \tag{2}$$

$\ln(INVALSIScore_r)$ is the log of the average regional score in the Italian language test performed by INVALSI. Our coefficient of interest is $\beta_3$, which measures the differential effect of having higher grades in INVALSI test on the search intensity after the school closures, relative to the period before the schools closure.

Finally, to show how the search intensity evolved over time before and after the date of schools closure, we re-estimate the same regression on a week-by-week basis, where we use only Mondays of each week. Formally, we mean:

$$\begin{aligned}\ln(1+G.T.Index_{j,r,w}) &= \textstyle\sum_{i=1}^{8+13}(\alpha_i \mathbb{1}Week_w) + \beta_1 \ln(INVALSIScore_r) + \\ &+ \textstyle\sum_{i=}^{8+13}(\beta_{i+1} \mathbb{1}Week_w \times \ln(INVALSIScore_r)) + \\ &+ \gamma \ln(1+TotalCases_{rd}) + X'\delta + \lambda_j + \epsilon_{j,r,d}\end{aligned} \tag{3}$$

$\mathbb{1}Week_w$ is an indicator variable which takes the value 1 on the respective calendar week w. Our coefficients of interest in this regression are all the ones associated to the weekly interaction terms.

# 5 Results

The results for the estimated coefficients of regression 1 are presented in Panel A of Table 2. The first column presents the main coefficient of interest in a regression with platform fixed effects but no controls. The second column includes the regional controls, as well as the total number of daily cases by region. In columns 3 and 4 we repeat the full specification for each platform separately.

In the first column we observe that, on average, each region increased the search of the e-learning terms by four times, relative to the period before school closures. Adding controls to the regression decreases the magnitude of the main coefficient but does not affect it's statistically significance. While the share of internet usage seems to not be relevant when we pool the two platforms, the regions in the North searched less relative to their peak, than their counterpart regions. On the other hand, the number of total COVID-19 cases has the expected sign as the higher the number of cases the larger the search intensity.

When we split the first regression by platform we observe that the increase in search intensity was stronger for Google Classroom than for WeSchool, and that the latter was particularly less used in North. While the share of internet usage positively contributes to more searches for the Google Classroom, this variable is not statistically significant for the WeSchool platform.

The results for the estimated coefficients of regression 2 are presented in Panel B of Table 2. As before, the first column presents the main coefficients of interest in a regression with platform fixed effects but no controls. The second column includes the regional controls, as well as the total number of daily cases by region. In columns 3 and 4 we repeat the full specification for each platform separately.

In the first column we observe that the different behaviour after the date of schools closure differed by regions with different INVALSI scores. Namely, regions with better academic performance have searched relatively less than regions with lower academic performance after that date. To provide a better visualisation of this difference we plot in Figure A3 the average search indices for regions above the median INVALSI score and for regions below the median INVALSI score. The figure clearly illustrates that while academically high and low performing regions have a similar pattern both before and after school's closure, the jump in the search intensity is substantially different. Unlike Bacher-Hicks et al. (2021), who find that areas of the United States with higher income (revealed to be areas with average lower SAT scores, by Chetty et al. (2020)) demonstrated substantially larger increases in search intensity, our graph shows the opposite for the case of Italy. It was the academically low performing regions who experienced a larger increase in the search intensity.

Adding controls to the regression now increases the magnitude of the main coefficient both because, as shown in Table 1, the INVALSI scores are higher in the North of the country, and also because the number of daily cases was higher in the regions with better INVALSI scores. The

Table 2: Before-After and Difference-in-Differences Results

| Variables | (1) Both Platforms | (2) Both Platforms | (3) Google Classroom | (4) WeSchool |
|---|---|---|---|---|
| **Panel A** | | | | |
| $\mathbb{1}$After Schools Closure | 2.910*** | 1.707*** | 1.862*** | 1.552*** |
| | (0.019) | (0.065) | (0.091) | (0.093) |
| $\mathbb{1}$North | | -0.293*** | -0.184*** | -0.403*** |
| | | (0.025) | (0.034) | (0.036) |
| Share of Internet Usage | | 0.003 | 0.010*** | -0.003 |
| | | (0.002) | (0.003) | (0.003) |
| ln(COVID-19 Cases) | | 0.164*** | 0.140*** | 0.188*** |
| | | (0.008) | (0.011) | (0.011) |
| Platform FEs | Yes | Yes | - | - |
| Adjusted R-squared | 0.770 | 0.795 | 0.800 | 0.793 |
| Observations | 6,160 | 6,160 | 3,080 | 3,080 |
| **Panel B** | | | | |
| $\mathbb{1}$After Schools Closure | 23.602*** | 33.695*** | 37.744*** | 29.647*** |
| | (2.000) | (2.089) | (2.930) | (2.944) |
| ln(INVALSI Score) | 4.405*** | 4.582*** | 5.233*** | 3.931*** |
| | (0.310) | (0.362) | (0.510) | (0.506) |
| – × $\mathbb{1}$After Schools Closure | -3.884*** | -5.975*** | -6.739*** | -5.211*** |
| | (0.379) | (0.403) | (0.565) | (0.568) |
| $\mathbb{1}$North | | -0.073*** | -0.038 | -0.108*** |
| | | (0.025) | (0.034) | (0.036) |
| Share of Internet Usage | | -0.004 | 0.002 | -0.009** |
| | | (0.003) | (0.004) | (0.004) |
| ln(COVID-19 Cases) | | 0.124*** | 0.114*** | 0.134*** |
| | | (0.009) | (0.012) | (0.012) |
| Platform FEs | Yes | Yes | - | - |
| Adjusted R-squared | 0.874 | 0.886 | 0.888 | 0.886 |
| Observations | 6,160 | 6,160 | 3,080 | 3,080 |

Note: This table reports the results from estimating equations 1, in Panel A, and 2, in Panel B. The dependent variable is the logarithm of the Google Search Index for *Google Classroom* and *WeSchool*. *AfterSchoolsClosure* takes value 1 after March 4, 2020, and 0 before. *INVALSI Score* contains the regional average score of the 2018/2019 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. Columns 1 and 2 show the results using the full sample of observations (without and with controls), while columns 3 and 4 report the each regression for each searched term. The regression coefficients are weighted by each region's population. Heteroskedasticity robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

signs of these two coefficients are the same as in the first regression.

Analysing the difference-in-differences results by platform we conclude that not only the jump in Google searches was larger for Google Classroom but that this one was relatively smaller for regions with better INVALSI scores. The differential effect is similar in the case of the WeSchool platform. While the North dummy and the Share of Internet Access are statistically significant for the search of WeSchool the same is not observed in the case of Google Classroom.

Finally, to observe how the difference in search intensity evolved along time, within each period (before and after the schools closure) we plot in Figure A4 the estimated effects on the weekly version of the difference-in-differences specification in regression 3. While there are no statistically significant differences in the period before relatively to the week of school closures, in the period after there is an average decrease of 8% towards regions with 1% better INVALSI scores. In the panel below in the same figure we also replicate the same exercise across Mathematics grades and we find that even though the magnitude of the effect is smaller the results are not qualitatively different.

# 6 Alternative platforms that were popular before and after the pandemic - Studenti.it and Scuola.net

Our results in the previous section suggest that academically lower performing regions in Italy used e-learning platforms *more* during the national lockdown. Therefore, if the usage of these platforms is a good proxy for studying during the pandemic, one should not expect that this pandemic will widen school performance gaps in the country.

However, note that the search for Google Classroom (similarly to WeSchool) before the pandemic was extremely low in Italy. In this section, we repeat the same exercise with two e-learning websites that were relatively more popular before the pandemic (6 times larger, according to Google Trends), and continued to be used during the pandemic.

*Studenti.it* is an Italian website for studying support, managed by the Italian schooling books publisher Mondalori Media S.p.A.. It was created in 1996 as a blog of students and in the early 2000s gained its popularity as the students started to exchange their notes on the website. *Scuola.net* is instead more dedicated to professors, as a platform of digital solutions for teaching. This way, we should also be able to capture whether the search intensity differences by region were driven by the professors or the students.

First, we replicate the same exercise as with the e-learning tools implemented during the pandemic, but extend the pre-pandemic period until 2016.[6]. We follow the Google Trends search intensity of *Studenti.it* and *Scuola.net* between the summer of 2016 and the summer of 2020. Note that, unlike before, the data was requested for a period that is longer than 9 months, and therefore

[6]Note that before this year the data on Google Trends was less accurate

presents weekly frequency. In these regressions we also include month and year fixed effects to control for seasonality.[7]

The results are shown in Table 3. Like before, the first column presents the main coefficients of interest without controls. Column 2 adds the regional controls and the number of daily COVID-19 cases and, finally, columns 3 and 4 show the results for each platform separately.

Table 3: Difference-in-Differences Results for Scuola.net and Studenti.it

| | (1)<br>Both Platforms | (2)<br>Both Platforms | (3)<br>Scuola.net | (4)<br>Studenti.it |
|---|---|---|---|---|
| $\mathbb{1}$After Schools Closure | -4.643 | 6.830 | -0.079 | 13.738 |
| | (6.239) | (7.108) | (10.510) | (8.920) |
| ln(INVALSI Score) | 0.481 | -0.783 | -0.443 | -1.122 |
| | (0.344) | (0.703) | (1.064) | (0.908) |
| $- \times$ $\mathbb{1}$After Schools Closure | 0.915 | -1.412 | -0.081 | -2.744 |
| | (1.179) | (1.366) | (2.018) | (1.716) |
| $\mathbb{1}$North | | -0.064 | -0.123 | -0.005 |
| | | (0.055) | (0.083) | (0.071) |
| Share of Internet Usage | | 0.029*** | 0.026*** | 0.032*** |
| | | (0.006) | (0.009) | (0.007) |
| ln(COVID-19 Cases) | | 0.107*** | 0.086* | 0.129*** |
| | | (0.032) | (0.049) | (0.040) |
| Observations | 7,524 | 7,524 | 3,762 | 3,762 |
| Adjusted R-squared | 0.191 | 0.196 | 0.192 | 0.201 |
| Platform FEs | Yes | Yes | - | - |
| Year FEs | Yes | Yes | Yes | Yes |
| Month FEs | Yes | Yes | Yes | Yes |

Note: This table reports the results from estimating equation 2 for the platforms of *Scuola.net* and *Studenti.it*. The dependent variable is the logarithm of the *Google Trends Index* for the term specified in the column heading. *After Schools Closure* takes value 1 after March 4 2020 and 0 before. *INVALSI Score* contains the average score of the 2018 INVALSI test in Italian language competences for Grade 10 students in each region. *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. All the specifications use the full sample of observations, namely daily values of the Search Intensity Index from July 3 2016 to June 28 2020. The regression coefficients are weighted by each region's population. Heteroskedasticity robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

[7]Note also that the number of regions included in these regressions is only 18 as Valle d'Aosta and Molise did not report enough searches for these websites during the period in analysis.

Unlike before, we observe that regions with different academic performance did not differentiate their behaviour during the pandemic, relative to the previous years. As the series used in these regressions are scaled relatively to their own peak, we do not know, however, if there was still a difference between regions with different academic performance at each point in time.

In Figure A5 we re-scale the indices relative to the starting of the series. Although regions with different academic performance do not exhibit a consistent difference on the search for *Studenti.it*, we observe that regions with a score in INVALSI test below the median increased relatively more their searches for *Scuola.net* over time than regions with INVALSI test scores above the median.

These results indicate that the differences that already existed before the pandemic might be driven by the professors rather than the students usage of e-learning. To further investigate this point we rely on two additional data sources of microdata: PISA and INVALSI. The former provides information of e-learning usage by students and the latter by teachers for some years before the pandemic.

# 7 Students and Teachers E-learning Before the Pandemic

## 7.1 Students - PISA

To analyse the usage of e-learning technologies by students in regions with different academic performance, we use data from PISA (Programme for International Student Assessment). PISA is an international standardised survey to 15-year-old students that is comprised of a cognitive test on reading, mathematics and science, and complementary questionnaires to assess students' attitudes and motivations. In this section we focus on the ICT Familiarity Questionnaire and the Educational Career Questionnaire.

While these questionnaires include a very rich set of questions, the caveat is that not all the regions participate in every wave. 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) do not have 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 we use PISA 2015 and compare Lombardy with Campania. The two regions are among the most populated regions in Italy and have already been used as representative cases of the north-south divide in Italy in other studies (Acconcia and Graziano, 2017). In this analysis 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. Panels A and B in Table 4 report differences in the usage of ICT resources for schoolwork, and additional instructions, respectively.

Table 4: ICT usage

| Variable: Proportion of students | Campania (1) | Lombardy (2) | Difference (3) | Italy (4) |
|---|---|---|---|---|
| **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) |

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 brackets. All averages are weighted by the PISA final trimmed non-response adjusted student weights.

Panel A in Table 4 shows clear evidence that already in the year 2015 students in Campania were using e-learning technologies for schoolwork outside school more than students in Lombardy. Students in Campania were 10.4% more likely to use internet for schoolwork, 45.1% more likely to use 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 4, all these differences are statistically significant at a 1% level.

Panel B shows that students in Campania in 2015 were also more likely to use ICT in their additional instructions (not part of the student's mandatory school schedule) in both Mathematics and Italian. In both disciplines and regions, the internet tutoring was more common than the video recorded instructions but the differences across regions in Mathematics were statistically significant for both types of ICT. As for Italian, only video recorded instructions revealed a statistically significantly higher proportion of students using it in Campania, compared to Lombardy.

To evaluate whether the higher usage of ICT from students in Campania compared to Lombardy is driven by the ICT availability, in Table A1 we report the ICT access at home in the year 2015, where access is defined by having the digital device available and have used it at least once. Although students in Campania report higher availability of desktops, the proportion of students with laptops and tablets in Lombardy is statistically higher. When we look at internet, the proportion of students with internet connection is statistically smaller than that in Lombardy, but not when considered mobile internet. Therefore we conclude that the results above are not driven by students in Campania having higher access to ICT.

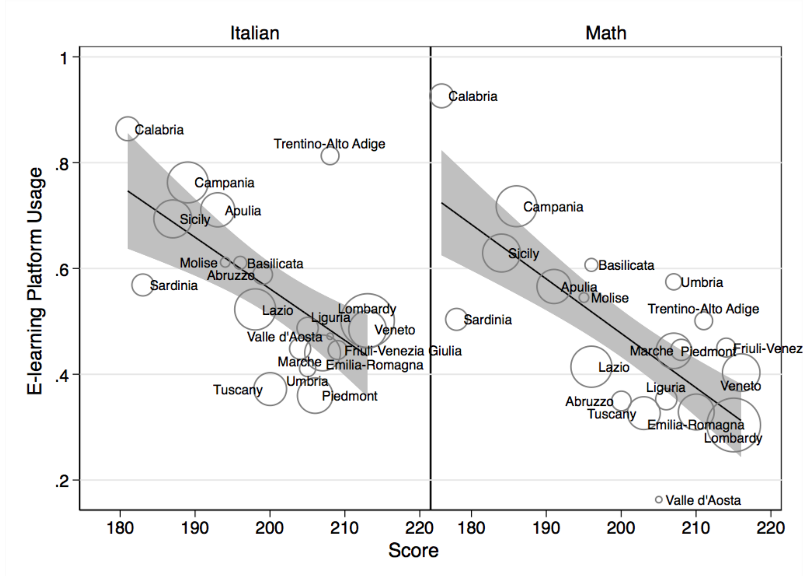
## 7.2 Teachers - INVALSI

Together with the tests described in Section 3.2, INVALSI carries out surveys to students, teachers and school principals, allowing for a comprehensive evaluation of the Italian education system. Surveys to school teachers provide us with valuable evidence of the usage of e-learning technology in their didactic activity before the pandemic. In this section we analyse the answers provided to the 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 being 0 =Never or almost never; 1 = Some times; 2 = Often; 3 = Always or almost always.

Figure 2, plots the average regional response to this question for teachers in Italian and Mathematics at Grade 10, across the average regional score in the same subject, according to the INVALSI tests in the academic year of 2019/2020. In both subjects there exists a negative relationship between the usage that teachers make of e-learning platforms when conducting their lessons and the regional mean academic performance of their students.

Comparing this with the academic year of 2017/18, where the question was also available, we observe a similar pattern. In Figure A6, we plot the average e-learning platform usage reported by Italian and Mathematics' teachers in the two consecutive academic years preceding the pandemic, splitting regions by whether they are above or below the median regional scores for Grade 10. In regions with an average score below the median, teachers were more likely to report a higher usage of e-learning platforms. Although the differences are similar for both subjects, Italian language teachers responded with slightly higher frequent usage in both academic years.

Figure 2: Teachers' E-learning Platform Usage by Students' Academic Performance




Note: 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 Italian 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 = Some times; 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.

Although there are substantial differences in the usage of e-learning, we do not observe the same when it comes to computer availability and usage in schools. In Figure A7 we plot the share of teachers that reported having access to a computer and its usage during their lessons between 2013 and 2019. Panel a) suggests that, until the academic year of 2017/2018, more teachers in higher academically performing regions had access to a computer to conduct their lessons, but in the last two academic years the two rates converged, especially in Italian. Moreover, if we analyse computer usage by conditioning on those teachers reporting having access to a computer during their lessons we observe no differences between regions with different academic performances. In Panel b), we observe that, once again, Italian language teachers report a higher usage of computers in classes but no difference between regions that are above and below the median scores. We take all this evidence of teachers in lower academically performing regions displaying a higher usage of e-learning platforms during the years preceding the pandemic, as suggestive evidence of these teachers being more prepared to swiftly shift their lessons to online learning after the national school closure.

# 8 Conclusion

With the enforcement of lockdown and school closures, Italy, like many other countries worldwide in 2020, was forced to implement e-learning tools in an unprecedented way. While some online resources such as *Studenti.it* and *Scuola.net* were already widely used by students and teachers before the pandemic to support education outside class, the usage of Google Classroom and WeSchool, implemented with the government program *Didattica a Distanza* was virtually zero before the academic year of 2019/2020.

As Italy is a country with large regional differences in both school quality, academic performance, and economic development, we study whether the usage of newly implemented e-learning tools was statistically different by regions with different academic performance. We employ a similar methodology as Bacher-Hicks et al. (2021) that studies regional differences by socio-economic status, recurring to the Google Trends. Surprisingly we find the opposite effects than those found for the United States, i.e., regions with lower academic performance (and lower average socio-economic status) had higher increases in the search of these tools.

We further investigate the origins of these differences, by using questionnaires to students (PISA) and teachers (INVALSI) on the usage of computer, internet, and e-learning tools before the pandemic. The analysis of these results show that both students and teachers in regions with lower academic performance were already using relatively more information and technology, and e-learning platforms before the national schools closure.

# References


**ACAPS**, "COVID-19 government measures dataset," 2020. data available at https://www.acaps.org/special-report/covid-19-government-measures-update-0.

**Acconcia, Giuseppe and Paolo R. Graziano**, *The Youth Guarantee: The Cases of Milan (Lombardy) and Naples (Campania)*, London: Palgrave Macmillan UK,

**Agasisti, Tommaso and Giorgio Vittadini**, "Regional economic disparities as determinants of student's achievement in Italy," *Research in Applied Economics*, 2012, *4* (2), 33.

**Andrew, Alison, Sarah Cattan, Monica Costa Dias, Christine Farquharson, Lucy Kraftman, Sonya Krutikova, Angus Phimister, and Almudena Sevilla**, "Inequalities in children's experiences of home learning during the COVID-19 lockdown in England," 2020.

**Bacher-Hicks, Andrew, Joshua Goodman, and Christine Mulhern**, "Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time," *Journal of Public Economics*, 2021, *193*, 104345.

**Baker, Scott R and Andrey Fradkin**, "The impact of unemployment insurance on job search: Evidence from Google search data," *Review of Economics and Statistics*, 2017, *99* (5), 756–768.

**Bando, Rosangela, Francisco Gallego, Paul Gertler, and Dario Romero Fonseca**, "Books or laptops? The effect of shifting from printed to digital delivery of educational content on learning," *Economics of Education Review*, 2017, *61*, 162–173.

**Beuermann, Diether W, Julian Cristia, Santiago Cueto, Ofer Malamud, and Yyannu Cruz-Aguayo**, "One laptop per child at home: Short-term impacts from a randomized experiment in Peru," *American Economic Journal: Applied Economics*, 2015, *7* (2), 53–80.

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

**Brunello, Giorgio and Daniele Checchi**, "School quality and family background in Italy," *Economics of Education Review*, 2005, *24* (5), 563–577.

**Castelnuovo, Efrem and Trung Duc Tran**, "Google it up! a google trends-based uncertainty index for the united states and australia," *Economics Letters*, 2017, *161*, 149–153.

**Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan**, "Income Segregation and Intergenerational Mobility Across Colleges in the United States*," *The Quarterly Journal of Economics*, 02 2020, *135* (3), 1567–1633.

**Commission, European**, "Europe's Digital Progress Report 2017: Connectivity," 2017.

**Cristia, Julian, Pablo Ibarrarán, Santiago Cueto, Ana Santiago, and Eugenio Severín**, "Technology and child development: Evidence from the one laptop per child program," *American Economic Journal: Applied Economics*, 2017, *9* (3), 295–320.

**Dergiades, Theologos, Costas Milas, and Theodore Panagiotidis**, "Tweets, Google trends, and sovereign spreads in the GIIPS," *Oxford Economic Papers*, 2015, *67* (2), 406–432.

**Engzell, Per, Arun Frey, and Mark Verhagen**, "Pre-analysis plan for: Learning inequality during the COVID-19 pandemic," 2020.

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

**Giuliani, Diego, Maria Michela Dickson, Giuseppe Espa, and Flavio Santi**, "Modelling and predicting the spatio-temporal spread of Coronavirus disease 2019 (COVID-19) in Italy," *Available at SSRN 3559569*, 2020.

**Goldsmith-Pinkham, Paul and Aaron Sojourner**, "Predicting Initial Unemployment Insurance Claims Using Google Trends," Technical Report, Working Paper (preprint), Posted April 3 2020.

**Hamid, Alain and Moritz Heiden**, "Forecasting volatility with empirical similarity and Google Trends," *Journal of Economic Behavior & Organization*, 2015, *117*, 62–81.

**Joyce, Ted, Sean Crockett, David A Jaeger, Onur Altindag, and Stephen D O'Connell**, "Does classroom time matter?," *Economics of Education Review*, 2015, *46*, 64–77.

**Murnane, Richard J, John B Willett, and Frank Levy**, "The growing importance of cognitive skills in wage determination," *Review of Economics and Statistics*, 1995, *2* (77), 251–266.

**Naccarato, Alessia, Stefano Falorsi, Silvia Loriga, and Andrea Pierini**, "Combining official and Google Trends data to forecast the Italian youth unemployment rate," *Technological Forecasting and Social Change*, 2018, *130*, 114–122.

**Peracchi, Franco**, "Height and economic development in Italy, 1730-1980," *American Economic Review*, 2008, *98* (2), 475–81.

**Rasch, Georg**, *Probabilistic models for some intelligence and attainment tests.*, ERIC, 1993.

**Rodríguez-Planas, Núria**, "Hitting Where It Hurts Most: COVID-19 and Low-Income Urban College Students," Technical Report, IZA Discussion Paper 2020.

**Vosen, Simeon and Torsten Schmidt**, "Forecasting private consumption: survey-based indicators vs. Google trends," *Journal of Forecasting*, 2011, *30* (6), 565–578.

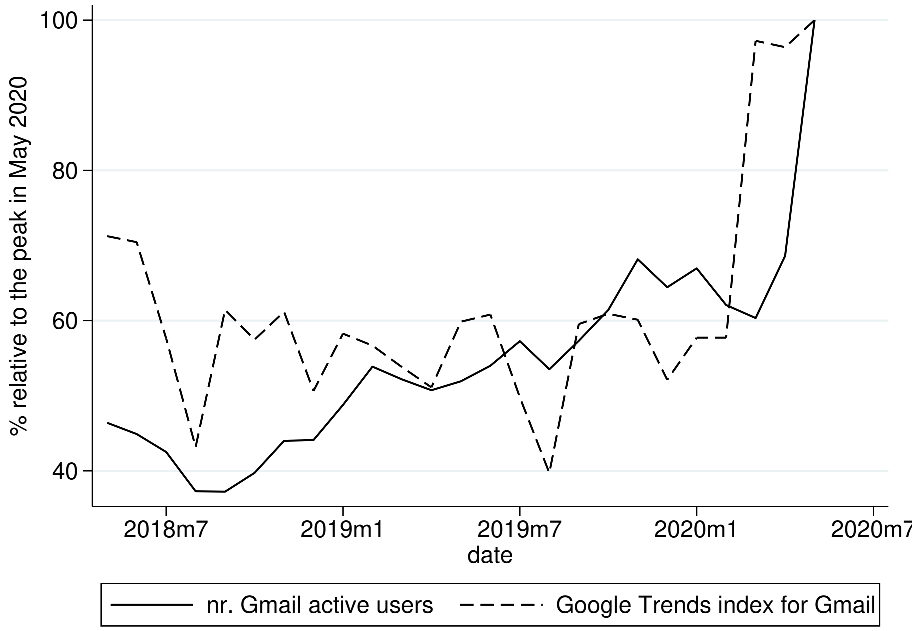
# Appendix

Table A1: ICT access at home

| Variable: Proportion of students | Campania | Lombardy | Difference | Italy |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Type of ICT | | | | |
| - desktop/tablet/laptop | 0.919 | 0.948 | -0.028*** | 0.933 |
| | (0.007) | (0.006) | [0.003] | (0.005) |
| - desktop computer | 0.640 | 0.561 | 0.080*** | 0.587 |
| | (0.013) | (0.013) | [0.000] | (0.009) |
| - portable laptop | 0.733 | 0.773 | -0.041** | 0.763 |
| | (0.012) | (0.011) | [0.011] | (0.008) |
| - tablet computer | 0.578 | 0.619 | -0.041** | 0.562 |
| | (0.013) | (0.012) | [0.022] | (0.009) |
| - internet connection | 0.945 | 0.966 | -0.020*** | 0.950 |
| | (0.006) | (0.005) | [0.010] | (0.004) |
| - cell phone with internet | 0.945 | 0.954 | -0.009 | 0.945 |
| | (0.006) | (0.006) | [0.313] | (0.004) |

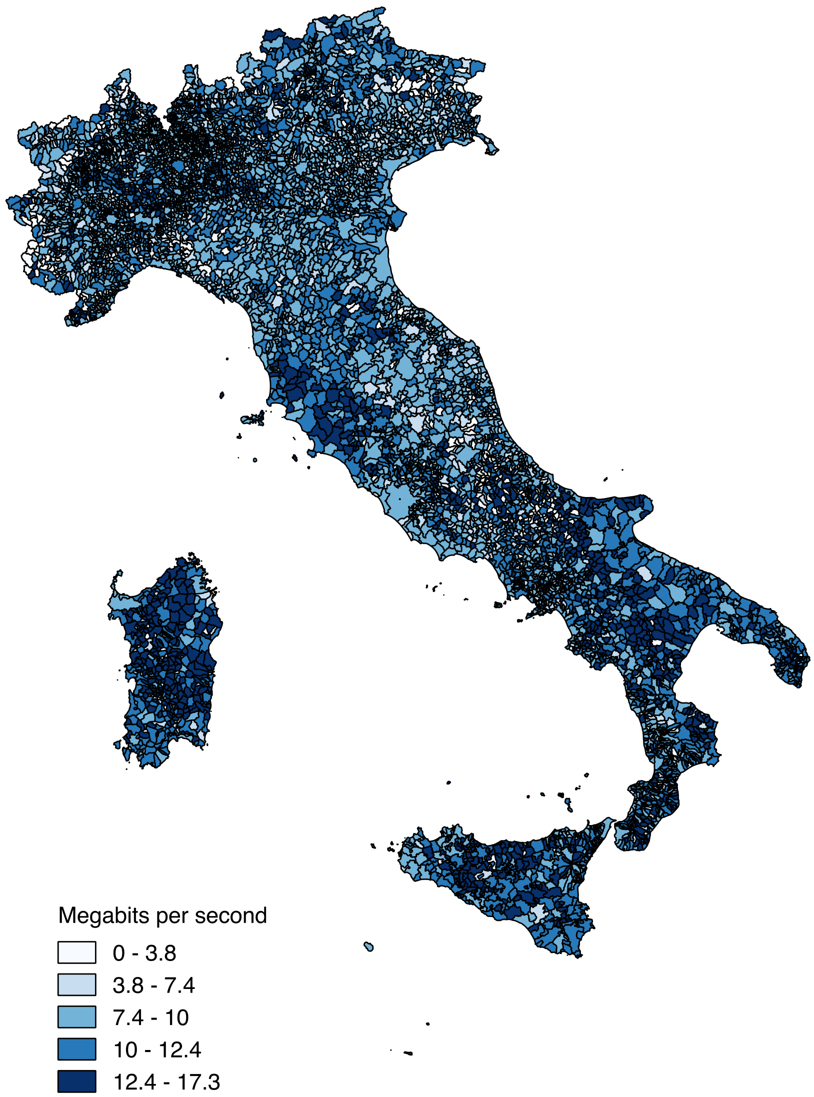
The data comes from PISA 2015 ICT Familiarity Questionnaire. Columns 1,2, and 4 report the proportion of students that answered that the reported device was available for them *and they used it*. The metric is equal to 0 if the device was available at home for them but they did not use it or if it was not available. The 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 brackets. All averages are weighted by the PISA final trimmed non-response adjusted student weights.

Figure A1: Comparison between number of active Gmail users and Google Trends Index for Gmail



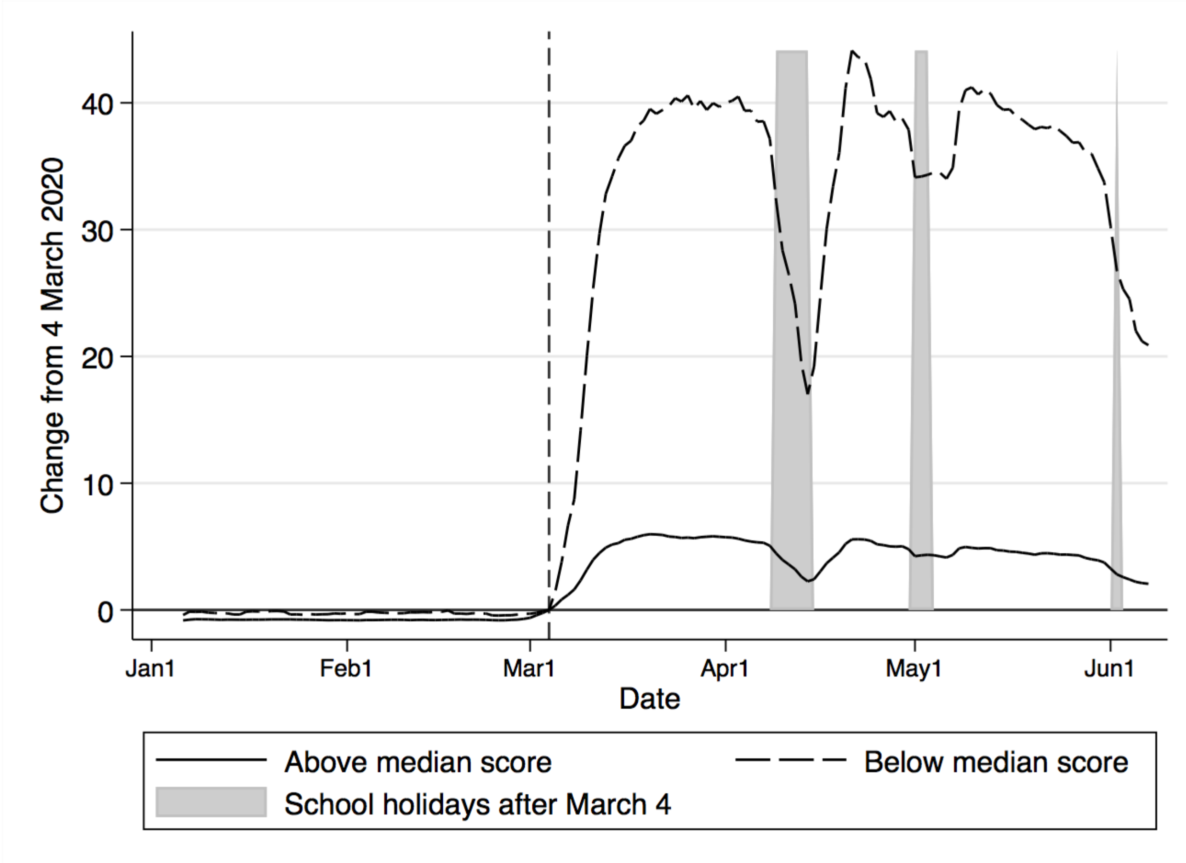

Note: This figure plots the average monthly number of active users of Gmail, provided by AirnowData, and the average monthly Google Trends index for Gmail, between May 2018 and May 2020. Both series are rescaled relative to the peak in May 2020.

Figure A2: Geographic Distribution of ADSL Download Speed in 2018




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

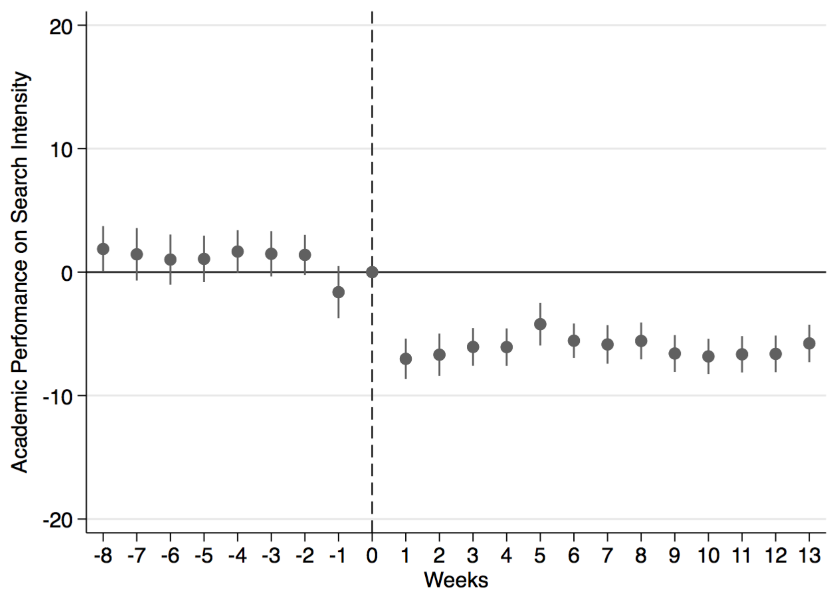
Figure A3: Google Trends Search Index for Google Classroom
by Academic Performane



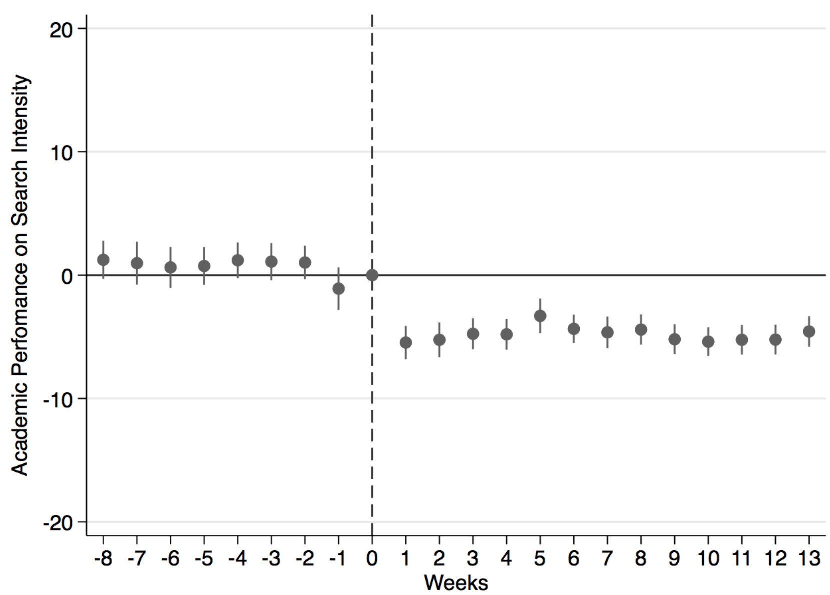

Note: This figure plots smooth daily changes of the Google Trends search index for the term *Google Classroom* in two groups of regions relative to March 4, 2020. The smoothing technique corresponds to the 7 day moving average computed by taking the average of each daily value together with the values from the preceding 6 days. Search index represented under below (above) median score contain the population weighted mean of the search index for the regions with a mean score in Italian below (above) the national median. Regional mean scores in Italian are extracted from the 2018 INVALSI report corresponding to Grade 10 students. Regional population shares used for the weights correspond to 2019 and are extracted from ISTAT. Shaded areas mark the official school holidays including the adjacent weekends that happened after the national closing of schools on March 4, date represented by the dashed line. See the footnote of Figure 1 for the precise days tha have been considered as school holidays.

Figure A4: Difference in E-learning Adoption by Academic Performance

(a) Italian



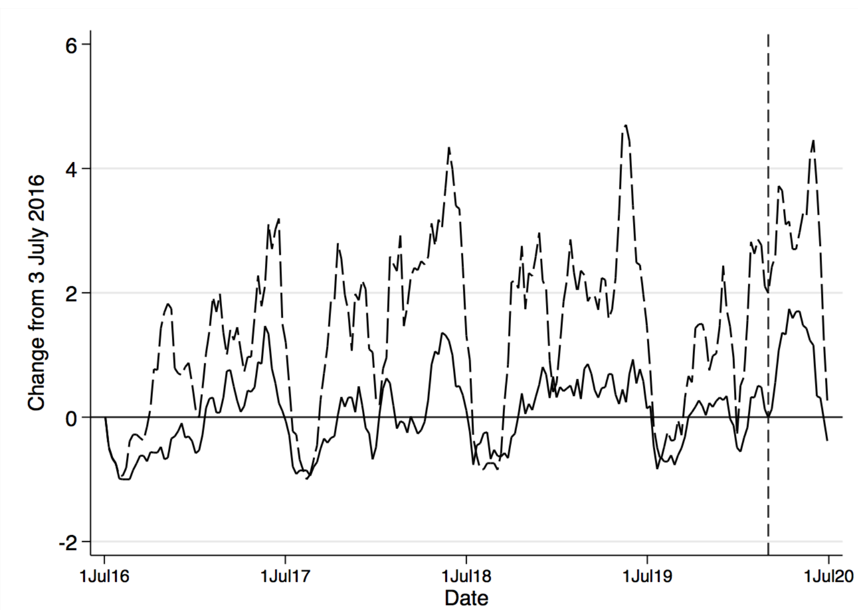

(b) Mathematics



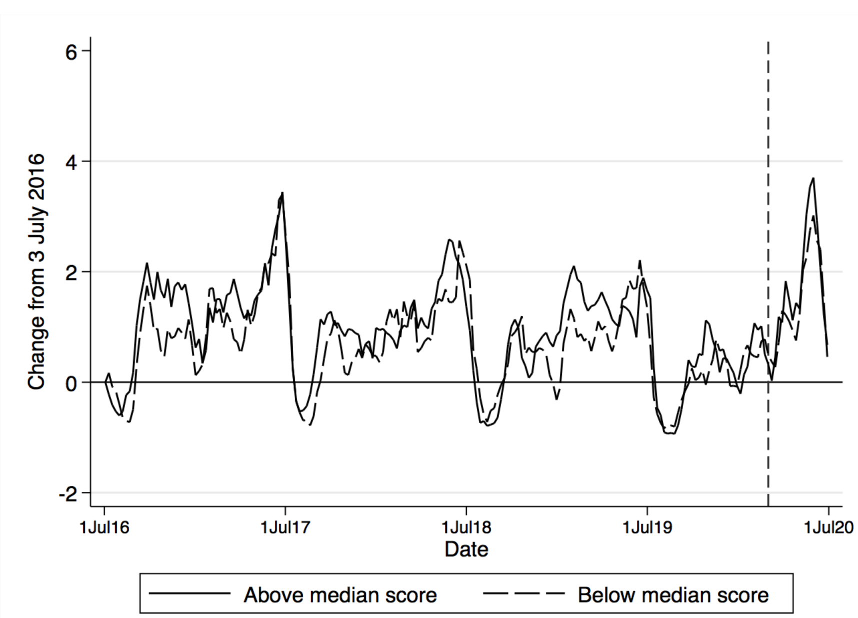

Note: This figure plots selected coefficients resulting from estimating by ordinary least squares equation 3. Panel (a) reads as one week after schools closed, Italian regions with a 1% higher average academic performance in Italian language are associated with a 8% on average lower level of internet searches of the two selected online learning platforms. The coefficient plotted are those in front of the interaction term between the natural logarithm of *INVALSI Score* and *AfterSchoolClosure*. Included controls and fixed effects are the same one as those detailed in column 2 of Table 2. The sample is comprised of weekly values on the Search Intensity Index from January 6 to June 7 2020 for three selected terms "*Google Classroom*" and "*WeSchool*". Vertical solid lines represent 95% confidence intervals computed using heteroskedasticity robust standard errors. Vertical dashed line marks the week of March 2 to 8 when all schools had to close. The regression is weighted by each region's population.

Figure A5: Google Trends Search Index for two E-learning Platforms by Academic Performance

(a) Scuola.net



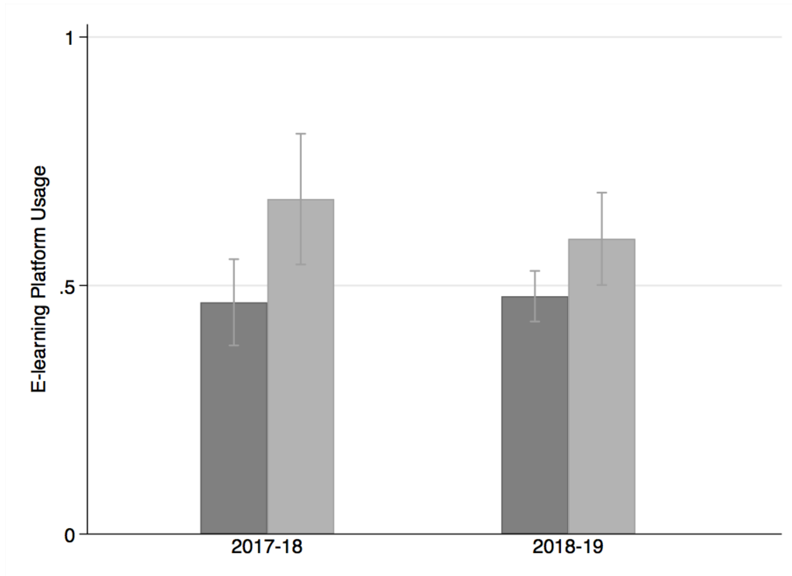

(b) Studenti.it



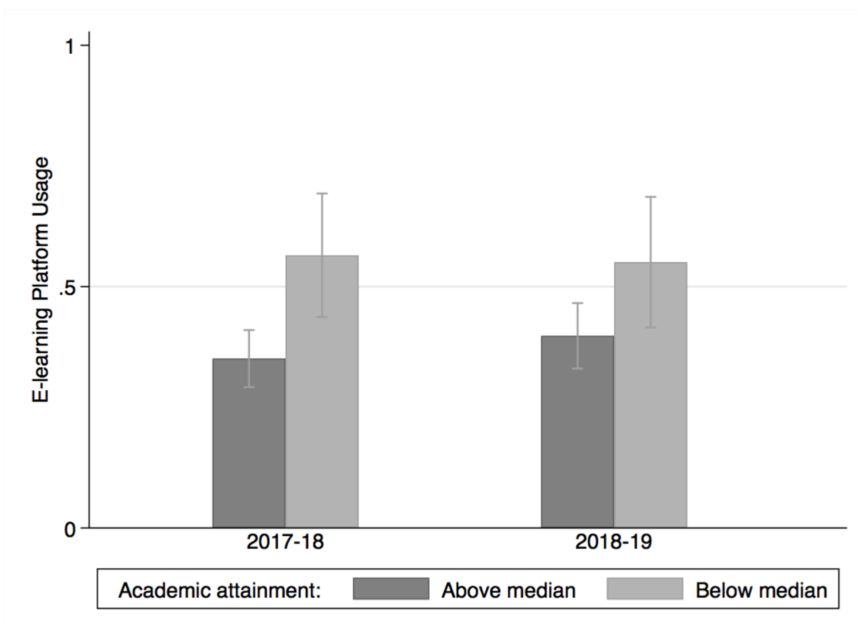

Note: This figure plots smooth weekly values of the Google Trends search index for the terms *Scuola.net* and *Studenti.it* in two groups of regions relative to July 3, 2016. The smoothing technique corresponds to the 4 week moving average computed by taking the average of each weekly value together with the values from the preceding 3 weeks. Search index represented under below (above) median score contain the population weighted mean of the search index for the regions with a mean score in Italian below (above) the national median. Regional mean scores in Italian are extracted from the 2018 INVALSI report corresponding to Grade 10 students. Regional population shares used for the weights correspond to 2019 and are extracted from ISTAT.

Figure A6: Teachers' E-Learning Platform Usage by Academic Performance

(a) Italian



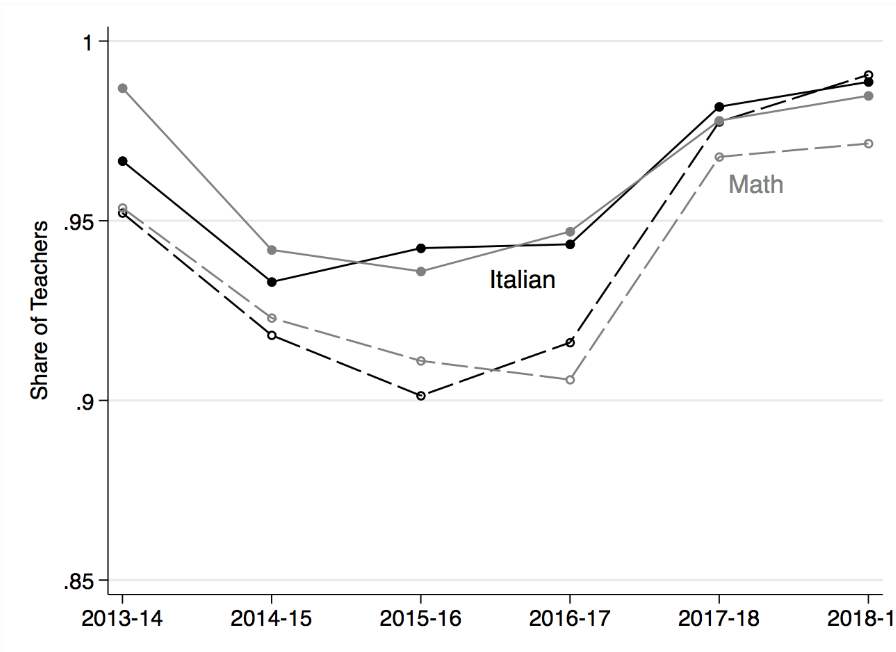

(b) Mathematics



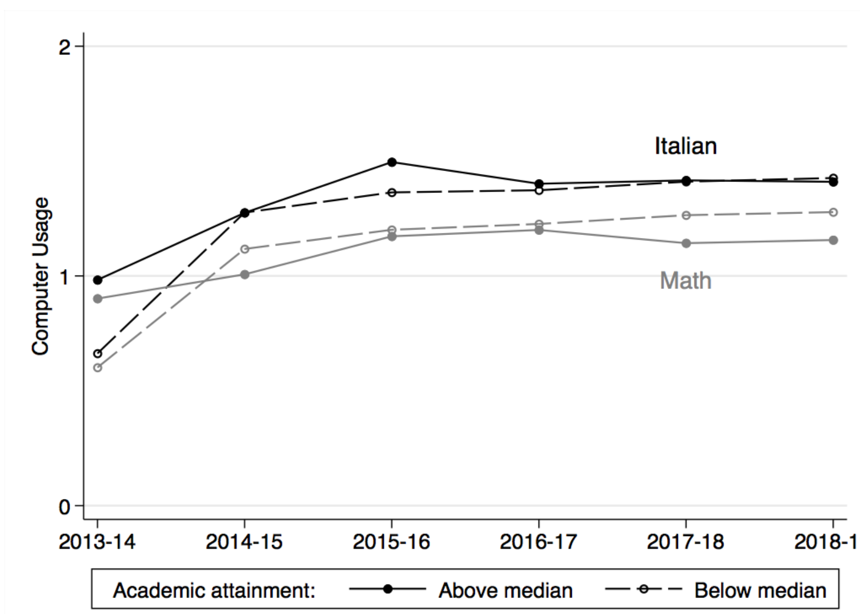

Note: This figure plots the reported usage of e-learning platforms by teachers when conducting their didactic activity in Italian and mathematics, respectively. The values are taken from the responses to question D2e administered by INVALSI to Grade 10 teachers both subjects during 2017-2018 and 2018-2019. The question states: *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 = Some times; 2 = Often; 3 = Always or almost always.* Below (above) the median contain the population weighted 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 2018 INVALSI report corresponding to Grade 10 students. Regional population shares used for the weights correspond to 2019 and are extracted from ISTAT. Vertical lines represent the 95% confidence interval of the sample mean.

Figure A7: Computer Availability and Usage by Teachers in Class by Academic Performance

(a) Computer Availability




(b) Computer Usage




Note: The figure plots the proportion of Italian 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 *Not present in School.* Panel (b) plots the group average of the following response options: *0 = I don't use it; 1 = Occasional use; 2 = Regular use.* The answers used are those from teachers responding that there was a computer available in class. Below (above) the median contain the population weighted mean of the responses in the regions with a mean score in each subject below (above) the national median, respectively.