Department of
Geography and
Environment

Papers in Economic Geography and Spatial Economics

High-speed Broadband, School Closures and Educational Achievements

Filippo Boeri

Paper No. 38
Geography and Environment Discussion Paper Series

February 2023

Editorial Board

Published by
Department of Geography and Environment
London School of Economics and Political Science
Houghton Street
London
WC2A 2AE

geog.comms@lse.ac.uk
www.lse.ac.uk/Geography-and-Environment

High-speed Broadband, School Closures and Educational Achievements

Abstract

In this study, I shed new light on the short-run effects of access to high-speed internet on educational disparities, before and after the pandemic shock. By following 3 million students belonging to 6 different cohorts over the period 2012-2022, I estimate the effect of the broadband infrastructure on student performance. While most previous contributions use discontinuous jumps in the available broadband connection speed across space at a given moment in time, this study exploits the actual roll-out of an infrastructural policy associated with an increase in 30 Mbit/s household broadband coverage from 40% to 80% over a 5-year period. The estimation strategy relies on a unique dataset, combining panel data on student performance with a rich set of school- and student-level information and broadband data measured at a very fine spatial scale. Results show an average null effect of high-speed broadband on 8th grade student performance in both numeracy and maths. However, this results masks substantial heterogeneity: lower performers in grade 5 and students with better backgrounds gain from internet speed, whereas the opposite is true for other students. Interestingly, the stronger effect on low-performers tends to disappear during the lockdown, suggesting a negligible mitigating role for high-speed internet during the period of school closure. On the other hand, the broadband infrastructure might have further amplified the gap between students with different socioeconomic background.

Keywords: ICT, education, economics, internet, broadband, Italy
JEL Codes: I24, H54, D83

1 Introduction

Nowadays, home computers have become an essential tool of modern education in developed countries. According to the last OECD report on ICT and education (Nusche and Minea-Pic, 2020), access to home computers is now nearly universal in most OECD countries. However, data still show significant disparities in the access to and quality of home computing. The digital divide is often related to various levels of access to high-speed internet connections, and for this reason, many countries have invested relevant amounts of public funding in order to upgrade information and communication technologies with the aim of increasing the available internet connection speeds. In Europe similar policies were supported within the broad framework of the 'Europe 2020 Strategy' ("EU2020"), which set ambitious targets for broadband development[1].

In spite of this, the actual impact of ICT on student performance, as on several other social outcomes, is still debated (Machin et al., 2007; Barrera-Osorio and Linden, 2009; Checchi et al., 2019; Cristia et al., 2017; Faber et al., 2015). Over the last 7 years, Italy, one of the lowest performers in the Pisa tests among OECD countries, was able to reduce this historical gap with other European countries in access to next generation access (NGA) broadband services (Oecd, 2014).

[Figure 1]

Fig. 1 illustrates the rapid increase in the share of Italian households with access to 30 Mbit/s internet, comparing to other European countries. While in 2015 only the largest

[1] 'Digital Agenda for Europe', one of the main initiatives promoted by EU2020, pursued the goal to reach universal access to 30 Mbps internet speed by 2020.

urban areas had access to this technology, by 2020 a large part of the Italian territory, including the poorest Southern regions, had access to 30mbps internet services (Fig. 2).

[Figure 2]

Coincidentally, the huge increase in high-speed internet supply was followed by an even stronger ’demand shock’. In March 2020, Italy was one of the first countries to be hit be the COVID-19 pandemic. In a bid to contain the number of cases, the Italian Government pioneered the severe lockdown measures that would have been introduced in most European countries in the following months. Between March 2020 and December 2021, businesses, leisure venues and even schools were closed for several months in most affected regions. These disruptions have raised concerns over possible learning losses, adverse socio-emotional effects and mental health issues among students (OECD, 2021). According to the United Nations Development Programme, 2020 was characterized by an unprecedented decrease in human development of about -0.025. In many countries, lockdown measures had a particularly severe effect on students from disadvantaged backgrounds, causing a general increase in educational inequalities (UNDP, 2020). Despite the differences across social groups, the pandemic had a negative effect on all students, regardless of the socioeconomic background (see Fig. 4).

[Figure 3]

Since the early phase of COVID-19 pandemic, many countries tried to mitigate the learning losses by implementing remote learning practices as an emergency response. For a country like Italy, characterized by a very recent infrastructure, the pandemic was the first opportunity to exploit the full potential of the new technology.

[Figure 4]

This paper sets out to estimate the causal effect of upgrades to the available internet speed on educational achievements, before and after the pandemic outbreak. The identification strategy relies on the specific features of a policy implemented by the Italian Government in 2015; the 'National Ultra-Broadband Plan' (NUBP) is a national plan aimed at ensuring 100% coverage at 30 Mbit/s and 85% coverage at 100 Mbit/s by 2020.

This study exploits the staggered roll-out of the plan. In order to cover the whole territory in a relatively short period, whilst minimizing public spending, the NUBP was implemented progressively in adjacent territories. As a result, the timing of the implementation can hardly be related to variables associated with educational outcomes[2]. Furthermore, the analysis is conducted following six cohorts of students over time. This way, I am able to control for possible time-invariant student, school and municipality characteristics and province-level shocks[3] that might be correlated with the ability of local council to 'capture' the central legislator and obtain a preferential treatment in the roll-out of the infrastructure. In addition, I focus on the broadband supply measure, rather than measuring its actual consumption. This allows me to bypass the endogeneity that may characterize the

[2]In the Appendix, table A8, I show that the rollout was not correlated with the political coalition that controlled the local government during the rollout of the infrastructure. In Fig. A3, I show that the increase in high-speed broadband coverage mostly took place through neighbouring municipalities. I also replicate the main results using only so-called 'market failure areas', where the concerns about the endogeneity of the rollout are mitigated, and excluding students that change municipality during the period considered.

[3]Italian provinces (with the exception of the current Sardinian provinces) correspond to Eurostat NUTS 3 regions. The Italian territory is currently divided in 110 autonomous provinces.

internet usage measures frequently used in the literature. Prior student's performance and a rich set of time-variant student characteristics allow me to identify the policy effect and to investigate its heterogeneity across different student types. Finally, by studying the way the broadband effect evolved in the aftermath of the pandemic shock, I shed new lights on the actual mechanisms at play.

Results show a null effect of 30 Mbit/s broadband on student performance in both numeracy and literacy subjects. However, when heterogeneity in social background is accounted for, this impact becomes positive and significant for low-performers in the previous grade and for students with a particularly rich socioeconomic background. Overall, students' backgrounds seem to play a relevant role in the heterogeneous policy outcome. The pandemic significantly affected student performance, regardless of previous performance and family background. In this context, access to high-speed broadband failed to mitigate the overall learning loss. On the other hand, the shock might have further reduced the positive effect on low-performers observed during the introduction of the digital infrastructure.

This study contributes to a growing body of literature on the socio-economic consequences of broadband, such as employment and productivity (Akerman et al., 2015), electoral outcomes (Falck, Gold, et al., 2014; Campante et al., 2018), marriage rates (Bellou, 2015) and housing prices (Ahlfeldt et al., 2017). This paper belongs to a narrow literature investigating the educational effects of high-speed broadband. Several studies analyse the effects of ICT in school settings (Angrist and Lavy, 2002; Rouse and Krueger, 2004; Goolsbee and

Guryan, 2006; Machin et al., 2007; Belo et al., 2014; Falck, Mang, et al., 2018). These studies generally find mixed results, and a direct comparison is limited by difficulty in measuring the actual use of ICT across different schools in different environments. A different strand of the literature analyses the impact of policies promoting the diffusion of education software, finding evidence of a positive effect on both maths and reading (Banerjee et al., 2007; Barrow et al., 2009; Barrera-Osorio and Linden, 2009; Muralidharan et al., 2019). This paper is different since it focuses on the effects of broadband access at home.

A few studies have focused on the relationship between home computer technology and student achievement. For example, Beltran et al. (1997) use a large panel dataset of US students to explore the causal relationship between computer ownership and various educational attainment levels, with a specific focus on high school graduation. Fairlie et al. (2010) conducted a large field study involving almost 8,000 students enrolled in grades 6-10 in 15 different middle and high schools in the United States. Fiorini (2010) use data from the Longitudinal Study of Australian Children (LSAC) to analyse the causal relationship between computer usage and children's cognitive and non-cognitive skills. They are able to test individual skills at two moments in time, instrumenting computer usage with their parents' previous computer ownership. Results generally exhibit computer usage to have a positive effect on cognitive skills, whereas the results are mixed for non-cognitive skills. This study takes a different angle, focusing instead on the exogenous supply of high-speed internet.

This paper does not focus on hardware and instead contributes to a narrow and relatively recent literature investigating the effect of home access to high-speed internet on educational outcomes. Malamud et al. (2019) implements a randomized control trial comparing a group of students who are randomly assigned a laptop with high-speed internet with two groups receiving only the laptop or the access to internet. They find no significant effects on maths and reading scores. Faber et al. (2015) exploit randomly placed jumps in the available ICT across neighbouring residences to investigate the causal effect of a sensible increase in available internet speeds on educational outcomes in the United Kingdom. The paper implements a fuzzy spatial regression discontinuity design across telephone exchange catchment areas and finds a positive and significant effect of an increase in broadband on student test scores.

This study contributes to this literature in different ways. First, I study the roll-out of an advanced fibre technology, rather than old DSL connections that still relied on copper line. Second, I exploit the features of a large infrastructure policy, bypassing the endogeneity issues that usually affect similar studies. Instead of exploiting 'jumps' in the in distance to the closest telephone local exchange station across catchment area boundaries, as in Faber et al. (2015) and Falck, Gold, et al. (2014), it directly exploits the variation over time in the staggered roll-out of the broadband. This is possible thanks to the specific characteristics of the National Ultra-Broadband Plan (NUBP), that determined an increase in 30 Mbit/s household broadband coverage from 40% to 80% over a 5-year period. In other words, this paper compares the performance of students located in areas that obtained access to the broadband over the period considered to students located in areas that remained disconnected from the network. A recent work, Cambini et al. (2021), studies the effect of the NUBP on student performance. The study focuses on the period 2012-2017 only and investigates the effect of the policy for grades II, V, and VIII. This study differs from their work in several ways. First, I only focus on grade VIII students, for whom digital technologies might be particularly relevant, and I follow students over time to control for previous performance in grade V. This allows me to control for the unobservable skills that characterize each student. Moreover, I control for both grade V and grade VIII school fixed effects, to absorb potential long-run effects on students' learning trajectories. Finally, I focus on a longer period, analysing the full introduction of the policy and how the effect was reshaped by the outbreak of the pandemic.

This study also contributes to a recent literature on the detrimental effects of Covid-19 pandemic on pupils' educational outcomes and inequalities. Other studies mostly focus on the heterogeneous effect with respect to student characteristics (Contini et al., 2021; Agostinelli et al., 2022) or school-level endowments (Gavosto and Romano, 2020), while overlooking the importance of local broadband infrastructures. This study contributes to this literature by directly exploring the nexus between access to high-internet broadband and the heterogeneous effect of the pandemic on student performance. The paper is organized as follows. Section 2 provides a general background of the Italian School systems and describes the main features of the NUBP. Section 3 presents a simple theoretical framework. Section 4 describes the different data sources used, and the procedure implemented to define school catchment areas. Section 5 describes the empirical strategy. In section 6, 7 and 8, I present the results and discuss the main policy implications.

2 Institutional Background

2.1 The National Ultra-Broadband Plan

In 2014, the Italian Government set up the 'National Ultra-Broadband Plan' (Piano Nazionale Banda Ultra-Larga - NUBP), a massive program aiming to ensure 100% coverage at 30 Mbit/s and 85% coverage at 100 Mbit/s by 2020. The plan was developed in accordance with the 'European Broadband Guidelines', which set out how the EU State aid rules apply to public funding for the roll-out of broadband networks. The national territory was classified into three different areas according to existing or expected future broadband infrastructure deployment:

1. White areas: areas where no provider of broadband services was currently operating and where no such provider was expected to enter the market in the coming three years.

2. Grey areas: areas where one (infrastructure-based) provider was already active, but another network was unlikely to be developed in the next three years.

3. Black areas: areas where there were or there would have been in the following three years at least two basic broadband networks of different operators.

The NUBP is based on four main pillars. First, the State guarantees administrative simplification and a reduction in burdens for all of the target regions. Second, private investments are encouraged in black and grey areas through the creation of tax exemption tools for infrastructure operations. Grey areas also benefit from various measures to facilitate the access to financial resources, the establishment of a guarantee fund and access to credit at subsidized rates. Finally, in white areas (commonly defined as 'market failure areas') the Public Sector intervenes directly to realize the infrastructures.

In 2014, most Italian households did not have access to modern fibre internet technologies and mostly relied on old infrastructures, offering an average broadband speed below 2 Mbit/s. After an initial delay, due to a number of legal disputes, in 2015 the program started to produce positive effects. Between 2015 and 2020 Italy managed to significantly reduce the historical gap with the other large European countries, by doubling the share of household with access to the infrastructure (see Figure 1). The specific characteristics of the broadband technology and the way the policy was implemented are such that these results are generally driven by an increase in high-speed broadband penetration in individual municipalities from 0% to 80%-100%. These significant results were made possible by the availability of EU structural funds (the European Regional Development Fund, European Agricultural Fund for rural development and the Development and Cohesion Fund), which complemented public funds.

2.2 The Italian School System

The poor performance of Italian students in PISA tests is often associated with an old educational system, characterized by antiquated protocols, extremely limited school autonomy and outdated curricula. The Italian compulsory education system consists of 4 stages: nursery school (children between 3 and 6 years of age), primary education (children between 6 and 11), first grade (lower) secondary school (between 11 and 14 years of age) and second grade (upper) secondary school (from 14 to 19 years of age). Once these stages have been successfully completed, students can access the higher education offered by universities, institutes for Higher Education in Art and Music as well as Higher Technical Institutes. Education is compulsory for ten years, between the ages of 6 and 16. As a result, all students are expected to gain at least a ’Licenza media’ (lower secondary school diploma). Over the last decade, the country experienced a reasonable decrease in the number of high school early dropouts. In 2014, only 1.6% (mostly first-generation foreigners) of the population in the 16-19 year-old cohorts did not hold a lower secondary school diploma. Primary and lower secondary school together form the first cycle of education, lasting 8 years. According to the new ministerial guidelines, the general aim of lower-secondary education is ’the harmonious and comprehensive development of the individual, according to the principles of the Italian Constitution and European cultural tradition, to be achieved through the promotion of knowledge, respect for individual diversity and the active involvement of students and their families’ (Framework for Key Competences for Lifelong Learning set up by the European Parliament and the Council of the European Union through the Recommendation). The subjects taught in this stage are: Italian, English, a second foreign language, mathematics, science, technology, geography, history, music, art, sports science and Catholic religious education (optional). Schools are expected to provide 30 hours of teaching per week (990 hours per year), allocated according to a common timetable.

School Councils can offer to some or all classes an ’Extended timetable’ (from 36 to 40 hours per week). In this case, the mandatory education goals remain the same, but students are expected to allocate less time to at-home study. At the end of the three-year program, students need to pass a uniform national examination in order to obtain a diploma and to access the following stage. The examination consists of a national written test set by INVALSI (also used by the Institute as a national assessment for grade 8) and four written tests set by a mixed (internal-external) committee. The subjects covered in the tests are Italian language, mathematics, science, informatics and two foreign languages. In contrast to the following two cycles (upper secondary school and tertiary education), the primary and lower-secondary schools are characterized by a very low, if not entirely absent, degree of autonomy (Ichino and Tabellini, 2014). First, individual schools have almost no influence on the design of the education programs. The national Ministry designs the course contents, defines the number of hours to allocate to each subject, and authorizes a limited number of textbooks for each field. Single institutions are only allowed to use a limited budget to set up laboratories and extra-program activities, such as optional courses. Teachers are allowed to choose among a certain number of authorized textbooks for each subject and can design their classes based on the national program, but need to report each semester to the Ministry. Teaching methods and contents must be consistent with each school's educational offer plan, which in turn must be consistent with the educational goals established at the national level. Second, the homogeneity in the service is also guaranteed by a rigid financial system. Between 97% and 100% of the school budget depends on transfers from the Central Government. Every three years, the Ministry allocates resources based on specific criteria (i.e. number of students, number of disabled students, specific needs, etc.). The uniformity of the system is also guaranteed by the human resource management, which is primarily conducted at the national level. Teachers apply to province-level lists and are assigned a ranking. Vacancies in each school are covered on the basis of teachers' preferences and rankings, with little to no involvement of the school directors. Salaries and career development are defined by national agreements. Recent reforms have tried to introduce some timid performance pay scheme, but the strong opposition by the powerful school Unions have managed to keep the status quo ante.

Another relevant feature of the lower secondary education system is the limited competition among schools. Classes can have between 15 (down to 10 in remote areas) and 26 students. Classes that exceed this number are split using the limited extra budget allocated by the Central Government. When the number of students applying exceeds the available places, schools are allowed to select entrants according to various criteria, but are expected to take into account distance as the main criterion. All of these characteristics together enforce a high degree of homogeneity among different schools. Even though national data still show disparities in teaching standards among different regions, mostly based on the quality of buildings and teachers' self-selection, there is strong evidence of a generally uniform service quality within provinces (Nuts 3 areas). For the purpose of this study, the high degree of within-province homogeneity that characterizes the Italian education system significantly mitigates the potential confounding issues that researcher often face while evaluating infrastructural policies in other countries.

3 Theoretical framework

This section presents a basic model to guide the empirical analysis. I study the effect of changes in access to high-speed internet on learning outcomes using a simple production function. Following Faber et al. (2015) I distinguish two main mechanisms:

1. ICT improvement can change the productivity associated with a given amount of time spent studying (MOOC effect).

2. High-speed internet can affect the supply of time spent studying relative to leisure activities (online-gaming effect).

This model simply extends the one proposed by Faber et al. (2015) by taking into account any potential background-biased effect of ICT access on students' productivity. In this framework, students with a better family background are expected to maximize productivity gains, whereas disadvantaged students may be less likely to offset the negative 'gaming effect'. This assumption is built on the rich - although mostly qualitative - literature investigating the factors related to the relationship between student performance and their access to ICT. Consistently, student i's knowledge production function is given by:

$$H_{i2} = A_{i2} L_{i2}^{\alpha} H_{i1} \tag{1}$$

where H_{it} is the educational achievement at the entrance to (t=1) and exit from (t=2) a given school cycle, A_{it} is an individual learning productivity shifter, L_i is the time spent studying and $\alpha > 0$ is the elasticity of learning outcomes with respect to time spent studying. I assume both productivity and individual labour supply to be functions of student-specific characteristics ($\lambda_i^A$and λ_i^L), and school characteristics (μ_s^A and μ_s^L). Access to high-speed internet (S) affects both the student productivity shifter and the time spent studying:

$$A_{i2} = S^{\delta(B)} \lambda_i^A \mu_S^A \varepsilon^{E^A} \tag{2}$$

$$L_{i2} = S^{\eta} \lambda_i^L \mu_S^L \varepsilon^{E^L} \tag{3}$$

Following a basic labour supply equation, η captures a relative price effect, since S affects the relative attractiveness of studying compared to leisure activities, online or offline. On the other hand, δ, depending on students' background (B), captures the effect on individual productivity. Substituting (2), (3) into (1) and taking logs, I obtain the following estimation equation:

$$\ln H_{i2} = [\alpha\eta + \delta(B)] \ln S + \lambda_i + \mu_s + \ln H_{i1} + \varepsilon \tag{4}$$

Where $\mu_s = \ln \mu_s^A + \alpha \ln \mu_S^L$, $\lambda_i = \ln \lambda_i^A + \alpha \ln \lambda_i^L$, $\varepsilon = \varepsilon^A + \alpha\varepsilon^L + \gamma\varepsilon^P$.

The main hypothesis tested is that the interplay between $\alpha\eta$ and $\delta(B)$ will determine different effects depending on B, a 'learning multiplier' linked to household characteristics (parents' education, occupational status, etc...). To simplify the empirical analysis, I rewrite the previous equation as:

$$\ln H_{i2} = \beta(B) \ln S + \lambda_i + \mu_s + \ln H_{i1} + \varepsilon_{i2} \tag{5}$$

where $\beta = \alpha\eta + \delta(B)$.

4 Data

4.1 Main Sources

In this study, I create a unique dataset, linking microdata on student achievements to spatial data relating to internet broadband coverage. In this way, I build up a comprehensive pupil-level dataset, enriched with regional, town and school-level data, covering the period 2012-2022.

Student level data are retrieved from the Italian National Institute for the Evaluation of the Education System (INVALSI), a public research institution which is responsible for the annual assessment of the competencies of Italian students in both reading and mathematics. Tests are taken at a number of given grades (2, 5, 6, 8 and 10) and at a national level. Every year, the Institute publishes anonymized microdata on student performance. Since 2008, individual marks have been matched with a rich set of individual information, allowing for control of personal, family and school characteristics. I have information on the test results for the whole population of students in the 2012/2013, 2013/2014, 2014/2015, 2015/2016, 2016/17, 2017/2018, 2018/2019, 2020/2021 and 2021/2022 school years at grades 5 and 8 and have retrieved from the dataset individual information for five cohorts of students: 578,000 students born in 2000, who completed grade 5 in 2013 and grade 8 in 2016; 573,000 students born in 2001, who completed grade 5 in 2014 and grade 8 in 2017; 584,000 students born in 2002, who completed grade 5 in 2015 and grade 8 in 2018; 579,000 students born in 2003, who completed grade 5 in 2016 and grade 8 in 2019: 573,000 students born in

2005, who completed grade 5 in 2018 and grade 8 in 2021[4], 569,000 students born in 2006, who completed grade 5 in 2019 and grade 8 in 2022.

Tests are administered in each school by local teachers, under the supervision of external examiners. Students are also asked to fill in a short questionnaire, providing individual information such as the country of origin, the year of arrival in Italy, family background and the conditions and facilities for studying at home. The results, along with information provided by school secretaries, are matched with student scores and included in a dataset available for research purposes. In addition to rich student-level information, I have access to the level of education (according to the ISCED scale) of both parents, as well as to their occupational status, recorded in the Socio-Economic Index of occupational status (ISEI) and an index based on individual-level economic, social and cultural information (ESCS). I also have access to data on home computer ownership and availability of an internet connection before and after the policy roll-out.

School level data are provided by the Ministry of Education (Miur). In 2011, in accordance with the Community Guideline on public access to information held by public authorities, the Miur started to publish data on each State-recognized school, for any grade. Since 2012, all schools provide, on a yearly basis, information concerning the number of students enrolled in each grade by gender and nationality, number of classes, number of teachers, the school basic budget and a self-evaluation document, to be sent to the Ministry at the end of the academic year. Moreover, 80% of schools provide information on staff (age, type of

[4] The 2020 standardized test was supposed to take place on May 2020, but was cancelled due to the pandemic outbreak

contract, level of training), environment (desktops, ICT technologies, Wi-Fi coverage) and other relevant features. Moreover, each school provides the full address of each building (plesso) belonging to the school.

Since I do not have access to student addresses, I compute catchment areas around each school, by exploiting the 2011 Italian Census. The Census is a large survey conducted every ten years by the National Institute for Statistics (Istat)[5]. The primary objective of this survey is to update and review personal data, calculate the legal population level and gather information on the number and structural characteristics of houses and other buildings. Since the 1991 census, the collected microdata have been linked to a complete digital database in ArcInfo format at a scale of 1:25.000, integrating remote sensing images, IGMI maps and technical maps at regional level with information relating to the municipality. This advanced methodology allowed the Istat to produce detailed geocoded data on the Italian territory, which is divided into 402,000 areas. On average, each section hosts 142 people and for each one, the Istat releases information concerning the number of people living in each division, by gender and age class. Furthermore, the dataset can be matched to information on wage, occupational status and other social features on the basis of a 5-10% population sample living in each division.

Data on broadband are retrieved primarily from Infratel, the public-run company in charge

[5]The survey is divided in three main sections. The first section, the Agricultural General Census, provides complete information relating to the structure of the agricultural system on a national, regional and local level. The Industry and Service Census focuses instead on the production system, providing the most detailed source of information available. Both censuses are used to develop statistical strategies to conduct any sample-based surveys during the following decade. The third and most relevant survey is the Population and Housing census, which covers the whole population residing in the country at the census date.

of the implementation of the infrastructural program. The Infratel dataset contains information gathered through the monitoring process of the Italian Ultra-Broadband Plan. Since 2015, once a year, all internet providers are requested to provide information on the availability of broadband telecommunication infrastructures and on the private investment plans for the following three years. The data are collected for each house number and then aggregated at the Census area-level. Specifically, the dataset provides historical data on:

1. Share of buildings with coverage to at least 100 Mbit/s.
2. Share of buildings with coverage to at least 30 Mbit/s.

The dataset can be linked to a spatial dataset, providing geocoded boundaries for each section. In this way, the information gathered can be easily matched with other spatial datasets.

In the final dataset, information relating to individual students is matched with school data provided by the Ministry of Education and, via the school coordinates, to the data associated with the territory in which each school is located. Thus, for each student, I have the results in the national exams, a rich set of individual and family characteristics, various information about the school attended and the town in which they live, and the weighted broadband coverage measured in the proximity to the school. A complete description of the variables used in the analysis is reported in Table 1.

[Table 1]

4.2 Broadband Measure and Catchment Areas

Infratel provides information on the share of buildings with access to fibre-to-the-node (FTTN) and fibre-to-the-home (FTTH), technologies, ensuring, respectively, a minimum of 30 Mbps and 100 Mbps internet speed. Since I am interested in the simple access to fibre internet technologies, my variable of interest will be constructed based on the share of families with access to at least 30mbps internet speed: $BA\_vt = X^{FTTH} + X^{FTTN}$.

This simple variable is easy to interpret and looks particularly relevant for the purpose of this study. However, it could be the case that different broadband technologies have different effects on student performance.

Ideally, in order to correctly identify the effect of the policy, I would assign students' homes to treated and control groups. As I do not have information on student addresses, I perform this task by defining geographical catchment areas around each school. Following De Simone (2013), I implement a method to identify an area within which most resident students would be attending a specific school.

The strategy relies on two main features of the Italian lower secondary school system. First, as discussed in Section 2.2, schools are characterized by a high degree of homogeneity, especially at the province level. Second, the enrolment follows rigid geographic criteria. Each school is assigned limited funds and a maximum number of students. When applying students exceed the available places, schools are allowed to select students according to a number of criteria, but are expected to take into account home-to-school distance as the main criterion. In summary, parents have little voice in the choice of the school among different institutions that accept students primarily on a distance basis. When choice is possible, the high homogeneity between institutions still guarantees allocation based mainly on geographical criteria. As a result, the design of the catchment areas, especially in rural areas, appears to be a suitable method with which to assign each student to a specific broadband area.

This strategy exploits the Italian census 2011, which provides information on population by age at a very low spatial scale (402,000 census areas). Specifically, for each school j, the association procedure consists of the following steps:

1. identify the school type (primary, lower secondary or upper secondary) and, consequently, the relevant student population in the census area (population aged 5–9 years, 10–14 years or 15–19 years respectively);
2. compute the distance between school j and the nearest census areas up to 30 km distance;
3. for each school, neighbouring census areas v are sorted by distance (in ascending order);
4. compute the areas' cumulative relevant population;
5. select the closest N areas so that the cumulative relevant population contains a multiple k of the number of students enrolled in school j.
6. For each school, I create two different catchment areas: a small catchment area ($k = 1$), where at least 80% of the enrolled students live and a large one (k = 3), where all

enrolled students (and many non-enrolled) live.

Once the data have been extracted and the catchment areas defined, I can build a proxy for broadband coverage, BA_{ct}, obtained as the weighted average of the broadband coverage measured in each catchment area c. The weights used are the share of students living in each Census area v to the total catchment area.

$$BA_{ct} = \sum_{v} BA_{cvt} \left(\frac{n_{vc,t-l}}{N_c, t-l} \right)$$

where BA_{cvt} is the share of households with 30mbps internet broadband in the census area v, belonging to catchment area c , n_{vct} is the number of students living in the census area v and N_{ct} is the total number of students living in catchment area c. Finally, in order to simplify the interpretation of the results, I produce a simple dummy variable, BA^d_{ct} that equals 1 when a catchment area records a weighted coverage above 75%. This approximation does not involve a relevant loss of information, since, as shown in Fig. X in the Appendix, the values of broadband coverage are concentrated at the two extremes of the distributions. This is due to the fact, that the policy involves the progressive roll-out over single areas covering at least one full medium-size municipality. As a result, most treated catchment areas move quickly from a coverage close to 0% to a coverage close to 100%.

5 Empirical Strategy

5.1 School-level analysis

First, I focus on the effect of high-speed internet broadband on school-level performance. In order to identify the effect of the infrastructure, I need to control for changes in school composition. To do so, I first estimate the following model:

$$y_{ict} = \gamma X_{ict} + n_u + \nu_{ct} + \varepsilon_{ict} \tag{6}$$

where y_{ict} represents student's i achievement in school c at time t, X_{ict} is a vector of individual characteristics, n_u is a Grade V school fixed-effect. The residual year-on-year change in the school specific score is the estimated as school-year fixed effect ν_{ct}.

I then estimate the following specification:

$$\tilde{\nu}_{ct} = \beta BA^d_{ct} + \rho S_{ct} + \theta W_{ct} + \mu_c + \lambda_{pt} + \varepsilon_{ct} \tag{7}$$

where BA^d_{ct} is treatment dummy for catchment area c, S_{ct} is a vector of school-level time-variant characteristics and W_{ct} is a vector of average socio-economic characteristics of the school's catchment area. The specification includes grade VIII school and year-time-province fixed-effects, respectively μ_c and λ_{pt}

5.2 Student-level analysis

The school-level specification exploiting residualized average test scores does not fully account for the possible interplay between worker-, school- and catchment area-level characteristics.

Thus, following the stylized model presented in Section 3, I estimate a comprehensive student-level specification. In order to avoid a problematic comparison across observations treated at different moments in times, I stack the dataset in 2x2 subsamples satisfying the following criteria:

- All treated units have the same adoption date t_d^τ
- All units fall inside the sub-experiment's event window $(\tau_d - 1, \tau_d)$

I end up with 4 subexperiments covering the academic years 2015/16-2016/17, 2016/17-2017/18, 2017/18-2018/19, 2020/21-2021/22 [6].

This way, in each subexperiments d, treated observations are observed just before and after treatment and compared to observations yet to be treated or never treated. Then I append each dataset created for each sub-experiment and estimate the following specification:

$$y_{ict} = \beta BA_{ct}^d + \gamma X_{ict} + \rho S_{ct} + \theta W_{ct} + \mu_c \# g_d + n_u + \lambda_{pt} \# g_d + \varepsilon_{ict} \quad (8)$$

where y_{ict} represent student's i achievement in school c at time t, BA_{ct}^d is the weighted share of families with access to high-speed broadband in catchment area c, X_{ict} is a vector of time-

[6] The subsamples 2018/19-2019/20 and 2019/20-2020/21 are excluded since no exam took place in the academic year 2019/20 due to the pandemic.

variant individual characteristics, S_{ct} is a vector of school-level characteristics and W_{ct} is a vector of average socio-economic characteristics of the school's catchment area. The unobservable school quality is accounted by grade 8 school building fixed effects and grade 5 school fixed effects, respectively μ_c, and n_u, whereas λ_{pt} controls for province-specific shocks. Time and school-level fixed effects are interacted with the group d dummy.

In order to identify the causal effect of high-speed broadband on student achievement, I need to address various sources of bias. Many studies in the literature exploit data on internet usage, which often represents the only available information. The choice of this variable is problematic for a number of reasons. First, survey data, especially when covering technical information, are affected by attrition and measurement error. Moreover, internet usage is likely to be highly correlated with several variables related to the tested outcome, such as social background, profession, individual network and income. In this case, the treatment would not be orthogonal with respect to the unobservable individual characteristics. For all these reasons, I choose to focus on the broadband supply, measured as the average internet speed guaranteed by the broadband technology, weighted by the share of students located in each census area.

Broadband access measures are not exempted from endogeneity concerns. First, the empirical strategy must address possible selection bias. This may occur because broadband access can determine a self-selection of different groups across regions with different levels of coverage. Moreover, even excluding an active sorting based on internet speed, there are several reasons to believe that broadband roll-out may be far from random. Exchange stations are normally located close to central locations, where they can guarantee the best connectivity to high income households and offices. Faber et al. (2015) address this problem by adopting a neighbouring discontinuity design, able to guarantee a high degree of homogeneity between the treated and control group. This strategy is particularly effective in addressing endogeneity concerns, but it inevitably requires the analysis to focus on a small sample of the available data. Moreover, it is possible to perform a neighbouring discontinuity design only when the treatment and outcome variables share the same level of geographical detail. This is not the case for Italian student data, which can only be linked to school catchment areas. Data on distance from the closest local exchange station are exploited by Campante et al. (2018), who study the diffusion of access to high-speed internet using Italian municipal data from 1996 to 2013. The strategy is based on the assumption that the cost of providing ADSL-based broadband services varies depending on its relative position in the pre-existing voice telecommunications infrastructure. Since the pre-existing infrastructure was not randomly distributed, the authors implicitly assume that the correlation between distance and unobserved municipal characteristics has not changed during the period considered, other than through the introduction of high-speed internet. In other words, firms and households may differ in terms of time-invariant unobservables, but are assumed to have, for instance, the same wage/productivity growth. This assumption appears to be relatively strong and would require strong supporting evidence to justify. In fact, the regional economics literature provides rich evidence of rising regional disparities in most developed countries, including Italy (A'Hearn and Venables, 2013).

Instead of relying on the existing infrastructure, this paper exploits the specific design of the Italian 'National Ultra-Broadband Plan', which guarantees the timing of the roll-out to be exogenous. Even though some geographical characteristics associated with the lack of coverage of any previous infrastructure may still influence the implementation costs, the policy aims to cover 100% of municipalities within 5 years. As a consequence, implementation costs can hardly be correlated with the roll-out timing. On the other hand, an efficient implementation would imply a progressive geographical coverage. Moreover, the timing and the universal target mitigate the risk of a spatial sorting of households based on treatment. These assumptions alone do not allow political bias in the implementation phase to be excluded. Local administrators may lobby to obtain full coverage before neighbouring municipalities, which would result in a selection bias. However, this issue does not appear to be particularly relevant in this context, since the program has been designed by the national Government and Italian political system and does not have a relevant representation at the local authority level. Moreover, Mayors in small towns lack the political power to deliver relevant changes to a national plan, especially when this kind of change would involve higher costs for the whole project. This issue is further addressed by using school fixed effects, that can absorb the ability of local authorities to capture resources from the central government and province-year fixed effects, that absorb cross-province variation in the speed of the roll-out.

5.2.1 Value added model

The rich set of student level-characteristics may still not be sufficient to predict student performance. For this reason, I extend the model, controlling for student performance at the end of the previous educational stage (grade 5). Building on the basic framework, student educational attainment can be described through a value-added model (VAM)

$$y_{ist} = \alpha y_{ist-1} + \beta BA^{d}_{ct} + \gamma X_{ist} + \rho S_{st} + \theta W_{ct} + \mu_c + v_u + \lambda_{pt} + \varepsilon_{ist} \tag{9}$$

where y_{ict} and y_{ict-1} are, respectively, student i's achievement in school c at time t and $t-1$, BA^{d}_{ct} is the weighted share of families with access to high-speed broadband in the catchment area, X_{ict} is the vector of time-variant individual characteristics, S_{ct} is a vector of time-variant school-level characteristics, $W_c t$ is a vector of average socio-economic characteristics of the school catchment area, μ_c and n_u are school fixed effects, λ_{pt} are cohort-province fixed effects and ε_{ist} is the unobserved error term. Controlling for previous performance, I partially take into account potential time-invariant differences between treated and control students. Moreover, even assuming the treatment to be uncorrelated with the performance at $t-1$, the autoregressive specification allows me to investigate the heterogeneity of the treatment across different social groups. In particular, I can test the impact of the access to high-speed broadband on student performance, as depending on prior achievement, ethnicity, socio-economic background and nursery attendance.

In the economics of education literature, VAMs have often been used to measure the importance of productivity inputs (such as teacher quality or peer effects) on student performance. Recently, a number of concerns have been raised regarding the opportunity of using VAM models to estimate teacher quality (Kane and Staiger, 2008; Hanushek and

Rivkin, 2010; Kane and Staiger, 2008; Chetty et al., 2014; Condie et al., 2014). Most of these studies have discussed the possible bias resulting from student sorting and the reliability of standardized tests scores as a proxy for student achievement. These issues do not appear to be relevant in the framework of this study. First, as explained in Section 2.2, student sorting across schools appears to be a negligible phenomenon in Italy. On the other hand, although unobservable characteristics could, in principle, influence the cheating-corrected test scores, they are unlikely to be correlated with the roll-out of the broadband infrastructure. To further address these concerns, in Section 7 I show that results are consistent when I remove from the sample students that attended grade 8 in a different municipality from the one where they attended grade 5. If student performance exhibits a mean reverting pattern at the tails of the previous performance distribution, the value-added model might fail to describe the learning process. For this reason, in Section 7 I also estimate a quadratic specification.

5.2.2 Covid effect

Traditionally, Invalsi provided test scores standardized over the whole population of each cohort and subject. Starting with the academic year 2017-2018, they began providing a 5-point score that is comparable across cohorts. In order to test the learning loss experienced by the Covid cohort, I exploit this second measure, controlling for the previous performance in grade 5, standardized within cohorts. The identifying takes the following form:

$$\bar{y}_{ict} = \alpha y_{ict-1} + Covid_d + \gamma X_{ict} + \rho S_{ct} + \theta W_{ct} + \mu_c + n_u + \varepsilon_{ist} \tag{10}$$

where $\bar{y}_{ist}$ is the unstandardized score recorded by student i in grade 8, school c and time t, y_{ict-1} is student i's standardized achievement in school c at time t and $t-1$, X_{ict} is the vector of time-variant individual characteristics, S_{ct} is a vector of time-variant school-level characteristics, W_c is a vector of average socio-economic characteristics of the school catchment area, μ_c and n_u are school fixed effects.

6 Results

6.1 School-level analysis

I start the investigation by studying whether the broadband infrastructural roll-out affected average student performance. In Table 2, I report the regression estimates for equation 7. Column (1) illustrates results for maths and literacy when the dependent variable is the average school performance and the explanatory variable is the share of houses with access to high-speed broadband in the catchment area. Estimates show a small negative effect for maths and a small positive effect for literacy, but both coefficients are non-significant. In column (2) I move to the simple dummy equal to one when the weighted share of broadband coverage recorded in the catchment area exceeds 75%. For both maths and literacy, I find a null result. The pattern is confirmed in columns (3), where the dependent variable is a school-year fixed effect estimated from a student-level specification (see Eq. 7), in column (4), when I adopt a stacked design and in column (5), where I estimate a staggered DID model based on the methodology proposed by Callaway and Sant'Anna (2021). Overall, results point to a null effect of broadband on average student performance.

[Table 2]

6.2 Student-level analysis

In Table 3, I move to a student-level specification (see Eq. 8), progressively including a rich set of school-, municipality- and student-level covariates, as well as grade V and grade VIII school fixed-effects and province-time-year fixed effects. All specification are estimated using a stacked design.

[Table 3]

Columns (1) and (2) illustrate, respectively, results for maths and literacy standardized tests when no covariates are considered in the specification. High-speed internet broadband seems to have a small but significant effect on student performance in literacy. Access to broadband raises average student scores in literacy by 3% of a standard deviation. No significant effect is found for maths. In columns (3) and (4), with the inclusion of student-level and school-level variables, the coefficient for both test scores become non-significant and slightly negative. The further inclusion of peer effects (mean class values for a set of student level characteristics) in columns (5) and (6) and time-varying municipality variables in columns (7) and (8) do not significantly affect the point estimates. Overall, the estimates confirm a null average effect of the policy on student performance.

In Table 4, I extend the specification, by including student performance in grade V (see Eq. 9). The new specification can be interpreted as a value-added model of cognitive achievement, that makes it possible to investigate the effects of the staggered roll-out of the broadband infrastructure on students' learning trajectories.

[Table 4]

In columns (1) and (2), the policy is found to negligible and non-significant effect on student performance in both subjects. When I include student-level and school-level covariates in column (3) and (4), both coefficients turn negative. The inclusion of class composition and municipality-level variables in columns (5)-(8) slightly impact the significance of the broadband coefficient for maths scores, but mostly confirm previous results. Overall, the value-added model confirms a perfect zero effect of the policy on both maths and literacy scores. These results are in contrast with Vigdor et al. (2014), that find internet access in North Carolina to be associated with a 2.7% standard deviation decrease in numeracy scores and a 1% decrease for literacy. These results can also be analysed with respect to other policies aimed at improving ICT use in education.

6.3 Heterogeneity analysis

Table 5 sheds further light on these findings by looking at the heterogeneous effect of the policy on test scores, controlling for student performance in grade 5 and their socio-economic background.

[Table 5]

In columns (1) and (2), I interact the main regressor with student performance in grade V. Results suggest high-speed broadband to foster a mean reverting pattern. Low-performers in grade V relatively benefit from the policy with respect to their peers who obtained better grades.

In columns (3) and (4) I further investigate this relationship, interacting the policy dummies with the quartile of the previous test score distribution. Students in the lowest quartile of the performance distribution in grade V record a 1.8% increase in maths test scores, although the coefficient is non statistically significant. Similarly, they record a significant 2.9% increase in literacy scores. For both subjects, the effect of the policy progressively decline moving towards the higher quartiles and becomes negative for students above the median.

In columns (5) and (6) I focus on the second relevant source of heterogeneity, namely the parental background. In this case, I find a positive and significant contribution of ESCS index on the policy outcome. Students with an average ESCS index are not affected by the policy, while the ones with a standard deviation ESCS score above the mean benefit from a 2.1% standard deviation increase in maths scores and 3.1% gain in literacy. The pattern is further investigated in columns (7) and (8), where I focus on ESCS quartiles.The broadband infrastructure reduces students performance in the first quartile of the ESCS distribution by 5% standard deviations in both maths and reading. The negative effect progressively declines in magnitude for the upper quartiles and turns positive for students with a family background above the median.

Overall, results highlight an important nexus between the effectiveness of the policy and students' socioeconomic background. Average performers whose parents are sufficiently educated might benefit the most from the introduction of new information technologies.

In Table 6, I further investigate this dynamic, considering together student heterogeneity with respect to parental background and performance in grade V. To do so, I interact the main explanatory variables with two dummies that take the value of one when a student has an ESCS score or a grade in class V above the median of his cohort.

[Table 6]

Columns (1) and (2) I report the baseline model. Among disadvantaged students who performed poorly in grade V, the rollout of the infrastructure leads to a non-significant 0.5% standard deviation decrease in maths scores and 0.8% loss in literacy.

In contrast, low-performers with a good family background record a 2.6% standard deviation increase in numeracy and a 4.6% increase in literacy.

With respect to high performers, disadvantaged students record a 7.2% standard deviation drop in both subjects. On the other hand, high-performers slightly benefit from the policy, with a 1.1% standard deviation gain in maths and a 0.6% gain in literacy.

Moving to columns (3) and (4), I include in the model individual, school and municipality-level characteristics. Results are generally confirmed, although I observe a lower distance between disadvantaged students with different performance in grade V. High-speed broadband is found to have a non-significant effect on test scores of disadvantaged low performers. In contrast, low performers with a good parental background gain a 2.5% of a standard deviation in numeracy scores and 4.3% of a standard deviation in literacy. Disadvantaged high-performers record a 7.1% standard deviation decline in maths and a 6.7% loss in literacy, while advantaged ones slightly benefit from the policy.

Overall, I find that parental background significantly impacts the effect of high-speed broadband on student performance. At the same time, the infrastructure is found to reduce the gap between low and high performers. Low-performers with a good family background are found to benefit the most, while disadvantaged high-performers are negatively affected.

6.4 School closures

In Table 7, I shift the focus of the analysis to the impact of the lockdown on student performance. As shown in Fig. 3, Italy was among the countries that recorded the longest duration of school closures. Descriptive statistics suggest the lockdown to have had a relevant negative impact on all students, regardless of the socioeconomic background (see Fig. 4). To further test this conclusion, I implement the estimating Equation (10), where the dependent variable is an absolute score, comparable across cohorts.

[Table 7]

In columns (1) and (2) I test the effect of school closures on student performance. Consistently with the recent literature, I find a significant drop in both numeracy and literacy score, with Covid-cohort student experiencing a 10.4% standard deviation score loss in maths and a 7.3% score loss in literacy. In columns (3) and (4) I investigate the heterogeneity of the school closure effect with respect to the student socioeconomic background. Interestingly, I find a limited negative relation between socioeconomic background and magnitude of the Covid shock.

Students above the average ESCS do not experience a less severe Covid-shock than students below the average in maths and experience only a 15% reduction in magnitude for reading. In columns (5) and (6), I further investigate the relation between parental background and Covid performance. Students in the lowest quartile of the ESCS distribution experience an average drop of 9.6% standard deviation scores in maths and 7.6% standard deviation scores in reading. For both maths and reading, the effect looks more severe for students in the second quartile and for literacy only we observe some evidence of a lower magnitude for the higher quartiles. In any case, even students in the highest ESCS quartile record a 9% standard deviation loss in maths and a 5% loss in literacy.

Overall, I find that negative effects of the pandemic were quite evenly distributed across students with different socioeconomic background.

6.5 Broadband effect during the lockdown

In Table 8, I investigate how the covid pandemic affected the compositional effects of high-speed broadband on student performance. In this case, I am not able to implement a stacked design, due to the lack of data for year 2019/2020.

[Table 8]

In columns (1) and (2) I distinguish between high- and low-performers in grade V. In non-pandemic years, Low performers benefit from a 3.7% standard deviation higher scores in maths and 1.91% higher scores in literacy. On the other hand, high-performers experience a 2.9% standard deviation loss in maths and a 0.8% loss in literacy. During covid year, the positive effect on low-performers significantly declines, with just a 2.3% standard deviation gain in maths and a 0.7% gain in literacy. High-performers do slightly better in maths and slightly worse in literacy, but in both case the policy effect is closer to the one experienced by low-performers than in non-pandemic years. In columns (3) and (4) I focus on the ESCS distribution. Students with below median scores of the index record a 0.5% standard deviation loss in maths and a 0.8% standard deviation loss in literacy. In contrast, students with a score above the median gain 1.6% in maths and 1.9% in literacy. During the pandemic, both groups perform worse, but the gap in treatment effect slightly increases. In columns (5) and (6) we analyse the two dimensions together, confirming the reduction in the distance between high- and low-performers and a limited increase in the gap between socioeconomic groups, at least for maths. Overall, the pandemic reduced the hetereogeneity in treatment effect with respect to previous performance and further widened the gap among students with different socioeconomic background.

7 Robustness Checks

Thus far, I expressed performance in grade VIII as a linear function of performance in grade V. In order to account for mean reverting patterns at the lower tale of the test score distribution, in Table A1, I add a quadratic component to the estimation. The new model does not lead to significant changes with respect to the basic VAM, causing only a negligible reduction in the magnitude of the coefficients of interest.

In Table 7, I illustrate the way school closures affected on student performance, assuming a homogeneous treatment. In Table A2, I replicate the same analysis, but this time I allow for heterogeneity in the effect of the pandemic, by exploiting regional differences in the duration of school closures. This heterogeneity is only partially explained by the actual propagation of the virus and was based on arbitrary decisions taken by the local governments. Exploiting this fact, I study how the duration of school closures affected student performance. In columns (1) and (2), I find each week of school closures to be associated with a 0.96% standard deviation decline in performance in maths and a 0.59% decline in literacy. Columns (3) and (4) confirm this effect to be only marginally influenced by student socioeconomic background. Finally, columns (5) and (6) show the effect to be particularly severe for students in the second quartile of the ESCS distribution.

In the main analysis, I estimate stacked regressions where the variable of interest is a dummy that equals one if the catchment area records a weighted coverage of the broadband above 75%. This threshold might look arbitrary and mask some relevant information.

In order to rule out this hypothesis, in the Appendix I re-estimate each specification with a simple high-dimensional fixed-effect regression framework, where the explanatory variable is the actual broadband coverage reported in each area. In Table A5 I estimate the basic value-added model. In contrast with Table 4, columns (1) and (2) show a positive and significant effect of the policy on student performance. A 10% increase in broadband coverage leads to an increase of 0.46% of a standard deviation in maths and 0.51% in literacy. However, the results disappear as soon as I include school-level covariates in the specification and in the following columns progressively converge towards results in line with my baseline results.

In Table A6 I consider the heterogeneous effect of the policy with respect to previous performance and family background. Once again, the effect is consistent with my baseline. High-speed broadband tends to have a negative effect on high-performers in grade V and disadvantaged students. Finally, in Table A7 I replicate the empirical model estimated in Table 6. Results largely confirm the baseline predictions, but the intensity varies across groups and subjects. First, disadvantaged low-performers record a positive, although non-significant, effect in maths, where the baseline model shows a non-significant negative effect. The negative effect is instead confirmed in literacy. Low performers with a good family background still benefit the most, but in this case the magnitude is higher for maths score, whereas in the baseline the result is more sizeable for literacy. Similarly, the gap between low and high performers seems less relevant in literacy, whereas in the baseline results, the two subjects show similar patterns. Overall, despite some minor differences, the choice of the specification does not affect the main findings.

As documented in Section 5, the identification strategies rely on the assumption that within provinces the broadband rollout can be considered exogenous with respect to the variable of interest.As a further robustness check, in Table A3 I re-estimate the main specification focusing only the so-called 'market failure areas', where no provider of broadband services was willing to invest at the beginning of the period and the broadband roll-out during the period is driven only by the public intervention (see Sec. 2.1). In columns (1) and (2) I re-estimate the specification presented in column (7) and (8) of Table 4. When I focus only on 'market failure areas', obtain a small positive -although non-significant - effect on average performance in both maths and literacy. In columns (3) and (4) I re-estimate the specification presented in columns (3) and (4) of Table 6. Once again, results are mostly consistent with the pattern obtained before. However, I record a more relevant relation between the policy effect in maths and previous performance, whereas the effect on literacy scores is particularly amplified by the socioeconomic background. Overall, the table shows that if we focus on 'market-failure areas', the main patterns become even more clear-cut.

The policy outcome could simply result from sorting over space of high-performers. Advantaged families might choose to move to a neighbouring town to get access to high speed broadband. Alternatively, they might simply face higher commuting time in order to have their children attend a school with better access to ICT technologies. As suggested before, this appears rather unlikely, since the policy was designed to cover the entire national territory within a relatively short period. As a result, it is hard to imagine a family moving in order to access a service they would access anyway in a short period of time. Moreover, Italy is known to be one of the countries with the lowest mobility rate in Europe.

Few people move and when this happens, it is to move to the richer North, rather than to a neighbouring town. Nevertheless, in Table A4 I re-estimate the main specifications excluding students that for different reasons ended up attending a school in grade VIII located in a different municipality than the one where they attended grade V. This strategy goes well beyond the purpose of the robustness, since it excludes all students whose closest primary school was located in a different municipality than the closest lower secondary school. However, results are generally confirmed. In columns (1) and (2) I re-estimate the specification presented in column (7) and (8) of Table 4. By excluding movers, the magnitude of the coefficient slightly increases while remaining non-significant for both maths and numeracy scores. Columns (3) and (4), where I re-estimate the specification presented in columns (3) and (4) of Table 6, generally confirms the heterogeneous effect recorded in previous tables. Once again, similar results are found for literacy.

Finally, I consider the possibility that municipalities controlled by the same party as national governments could receive a preferential treatment. This is quite unlikely, for at least three reasons. First, since the policy aims to cover the entire territory in a relatively short period, the possible gains would be limited. Second, any change to the incremental rollout over adjacent territories would be extremely inefficient. Third, considering all largest cities already had access to the broadband at the beginning of the period, we would need to assume that mayors of medium-small towns had the power to influence the central government. Nevertheless, I identify Italian municipalities that during the period were run by either centre-left or a centre-right coalition[7] and assess whether being controlled by the same party as national governments was correlated with access to the broadband

[7] I exclude the large majority of the municipality, that were run by independent mayors

infrastructure. Results are reported in Table A8. In column (1), where I run a simple regression, I find that municipality run by the same political party as the government are 9.8% less likely to receive the policy. This result is likely to depend on specific characteristics of each territory. Thus, in the following columns I progressively move to the combinations of fixed effects used in the main analysis. In column (2), where I add municipality fixed-effects, the results declines to just 3.6%. The introduction of year-fixed effects in column (3) does not significantly change the result. In column (4) I introduce year-time-region fixed effects[8]. In this case, the magnitude of the coefficient shrinks by 80% and becomes non-significant. Finally, in column (5) I adopt the even stricter set of fixed-effects used in the main analysis, namely year-time-province fixed-effects. The results become negligible and non-significant. Overall, this exercise suggests that, at least within provinces, the rollout of the policy does not depend on political affiliations. Finally, in Fig. A3, I report a bin scatter illustrating the likelihood of treatment in t for municipalities that were not treated in $t-1$ with respect to the distance from the nearest municipality with access to the infrastructure in $t-1$, absorbing time fixed effects. The chart shows a clear negative relation between the two variables, confirming that the rollout of the infrastructure advanced mostly through neighbouring - if not adjacent - municipalities.

[8] Italy has 22 regions, corresponding to Eurostat NUTS1 administrative level.

8 Conclusions

In this study, I tested the impact of access to high-speed internet broadband on educational achievements. To this aim, I exploited a large infrastructural program implemented by the Italian government. Available data allowed me to investigate the heterogeneous effect of the treatment with respect to students' performance in previous grades and their family background. Overall, broadband access is found to have a negligible effect on average educational achievements. However, this result masks a significant heterogeneity with respect to parental socio-economic background and prior performance. Low performers in grade V, when benefiting from a rich cultural background, might take advantage of the new learning devices available thanks to the new infrastructure in order to reduce the achievement gap with their peers. On the other hand, no productivity gain is recorded for pupils with a poor background. This might be due to the fact that, without parental supervision, the online-gaming effect offsets any possible positive effect on learning productivity. Alternatively, it might simply reflect a financial constraint, with poorer families not able to provide their children with the hardware necessary to benefit from the new infrastructure. Interestingly, the heterogeneous effect with respect to and prior performance tends to disappear after the outbreak of the pandemic, while the opposite is true for the socioeconomic background. On average, the detrimental effect of the lockdown on student performance is not compensated by access to high-speed internet, nor it is its effect on performance inequality. Although more work is required to better understand the underlying mechanisms, this empirical evidence may have relevant policy implications. ICT upgrading programs can be beneficial for students, but need to be accompanied by training programs and other policies aimed at allowing disadvantaged students to access the benefits. Moreover, they cannot substitute in person interaction and become useless when students are not able to benefit from the direct guidance of teachers, as it happened during the Covid pandemic.

References

A'Hearn, B. and A. J. Venables (2013). "Regional disparities: internal geography and external trade". In: *The Oxford handbook of the Italian economy since unification*.

Agostinelli, F., M. Doepke, G. Sorrenti, and F. Zilibotti (2022). "When the great equalizer shuts down: Schools, peers, and parents in pandemic times". *Journal of public economics* 206, p. 104574.

Ahlfeldt, G., P. Koutroumpis, and T. Valletti (2017). "Speed 2.0: Evaluating access to universal digital highways". *Journal of the European Economic Association* 15.3, pp. 586–625.

Akerman, A., I. Gaarder, and M. Mogstad (2015). "The skill complementarity of broadband internet". *The Quarterly Journal of Economics* 130.4, pp. 1781–1824.

Angrist, J. and V. Lavy (2002). "New evidence on classroom computers and pupil learning". *The Economic Journal* 112.482, pp. 735–765.

Banerjee, A. V., S. Cole, E. Duflo, and L. Linden (2007). "Remedying education: Evidence from two randomized experiments in India". *The Quarterly Journal of Economics* 122.3, pp. 1235–1264.

Barrera-Osorio, F. and L. L. Linden (2009). "The use and misuse of computers in education: evidence from a randomized experiment in Colombia". *World Bank Policy Research Working Paper* 4836.

Barrow, L., L. Markman, and C. E. Rouse (2009). "Technology's edge: The educational benefits of computer-aided instruction". *American Economic Journal: Economic Policy* 1.1, pp. 52–74.

Bellou, A. (2015). "The impact of Internet diffusion on marriage rates: evidence from the broadband market". *Journal of Population Economics* 28.2, pp. 265–297.

Belo, R., P. Ferreira, and R. Telang (2014). "Broadband in school: Impact on student performance". *Management Science* 60.2, pp. 265–282.

Beltran, D., K. Das, and R. W. Fairlie (1997). "Do home computers improve educational outcomes? Evidence from matched current population surveys and the national longitudinal survey of youth 1997". *Evidence from matched current population surveys and the national longitudinal survey of youth*.

Callaway, B. and P. H. Sant'Anna (2021). "Difference-in-differences with multiple time periods". *Journal of Econometrics* 225.2, pp. 200–230.

Cambini, C., L. Sabatino, and S. Zaccagni (2021). *The Faster the Better? The Effect of Ultra-Fast Broadband on Students' Performance*. Tech. rep. CEBI Working Paper Series.

Campante, F., R. Durante, and F. Sobbrio (2018). "Politics 2.0: The multifaceted effect of broadband internet on political participation". *Journal of the European Economic Association* 16.4, pp. 1094–1136.

Checchi, D., E. Rettore, and S. Girardi (2019). "IC technology and learning: an impact evaluation of Cl@ ssi 2.0". *Education Economics* 27.3, pp. 241–264.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates". *American economic review* 104.9, pp. 2593–2632.

Condie, S., L. Lefgren, and D. Sims (2014). "Teacher heterogeneity, value-added and education policy". *Economics of Education Review* 40, pp. 76–92.

Contini, D. et al. (2021). *The COVID-19 pandemic and school closure: learning loss in mathematics in primary education*. Tech. rep. IZA Discussion Papers.

Cristia, J. et al. (2017). "Technology and child development: Evidence from the one laptop per child program". *American Economic Journal: Applied Economics* 9.3, pp. 295–320.

De Simone, G. (2013). "Render unto primary the things which are primary's: Inherited and fresh learning divides in Italian lower secondary education". *Economics of Education Review* 35, pp. 12–23.

Faber, B., R. Sanchis-Guarner, and F. Weinhardt (2015). *ICT and education: evidence from student home addresses*. Tech. rep. National Bureau of Economic Research.

Fairlie, R. W., D. O. Beltran, and K. K. Das (2010). "Home computers and educational outcomes: Evidence from the NLSY97 and CPS". *Economic inquiry* 48.3, pp. 771–792.

Falck, O., R. Gold, and S. Heblich (2014). "E-lections: Voting Behavior and the Internet". *American Economic Review* 104.7, pp. 2238–65.

Falck, O., C. Mang, and L. Woessmann (2018). "Virtually no effect? Different uses of classroom computers and their effect on student achievement". *Oxford Bulletin of Economics and Statistics* 80.1, pp. 1–38.

Fiorini, M. (2010). “The effect of home computer use on children’s cognitive and non-cognitive skills”. *Economics of Education review* 29.1, pp. 55–72.

Gavosto, A. and B. Romano (2020). “The impact of the COVID-19 pandemic on Italian schools and universities: the challenge of distance learning”. *Short-term Costs and Long-term Adjustments*, p. 103.

Goolsbee, A. and J. Guryan (2006). “The impact of Internet subsidies in public schools”. *The Review of Economics and Statistics* 88.2, pp. 336–347.

Hanushek, E. A. and S. G. Rivkin (2010). “Generalizations about using value-added measures of teacher quality”. *American economic review* 100.2, pp. 267–71.

Ichino, A. and G. Tabellini (2014). “Freeing the Italian school system”. *Labour Economics* 30, pp. 113–128.

Kane, T. J. and D. O. Staiger (2008). “Are teacher-level value-added estimates biased? An experimental validation of non-experimental estimates”. In: *National Conference on Value-Added Modeling, Madison, WI*. Vol. 47. Citeseer, pp. 106–118.

Machin, S., S. McNally, and O. Silva (2007). “New technology in schools: Is there a payoff?” *The Economic Journal* 117.522, pp. 1145–1167.

Malamud, O., S. Cueto, J. Cristia, and D. W. Beuermann (2019). “Do children benefit from internet access? Experimental evidence from Peru”. *Journal of Development Economics* 138, pp. 41–56.

Muralidharan, K., A. Singh, and A. J. Ganimian (2019). “Disrupting education? Experimental evidence on technology-aided instruction in India”. *American Economic Review* 109.4, pp. 1426–60.

Nusche, D. and A. Minea-Pic (2020). *ICT resources in school education: What do we know from OECD work*.

OECD (2021). *The State of School Education*, p. 51. DOI: https://doi.org/https://doi.org/10.1787/201dde84-en. URL: https://www.oecd-ilibrary.org/content/publication/201dde84-en.

Oecd (2014). *PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know*. OECD Paris.

Rouse, C. E. and A. B. Krueger (2004). “Putting computerized instruction to the test: A randomized evaluation of a ?scientifically based? reading program”. *Economics of Education Review* 23.4, pp. 323–338.

UNDP (2020). “COVID-19 and human development: Assessing the crisis, envisioning the recovery”. *2020 Human Development Perspectives*, pp. 1–30.

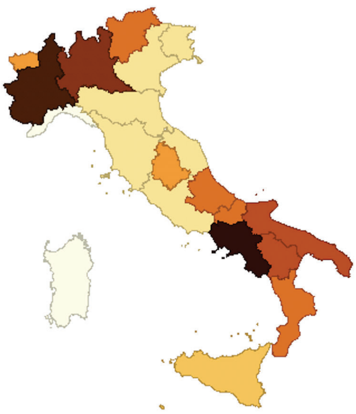
Vigdor, J. L., H. F. Ladd, and E. Martinez (2014). “Scaling the digital divide: Home computer technology and student achievement”. *Economic Inquiry* 52.3, pp. 1103–1119.

Appendices

This appendix presents additional text, tables and figures that complement the main paper.

A Charts

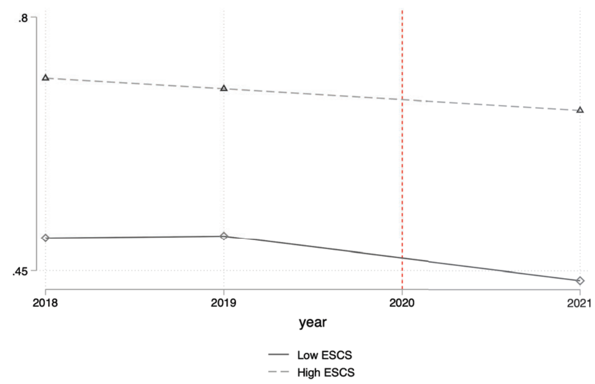
Figure A1: Weeks of school closure, 2020/21



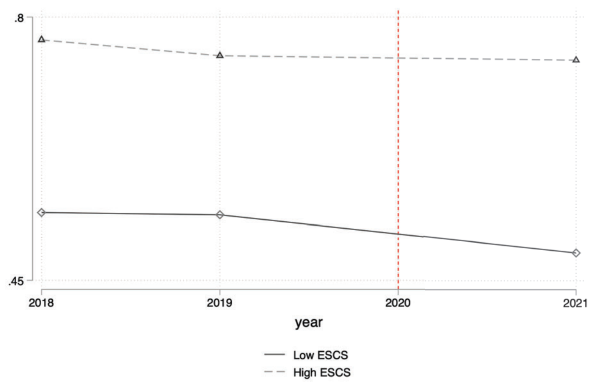
Notes: The figure illustrates the heterogeneity in the duration of school closures by region.

Figure A2: Share of students passing the exam

(a) Math

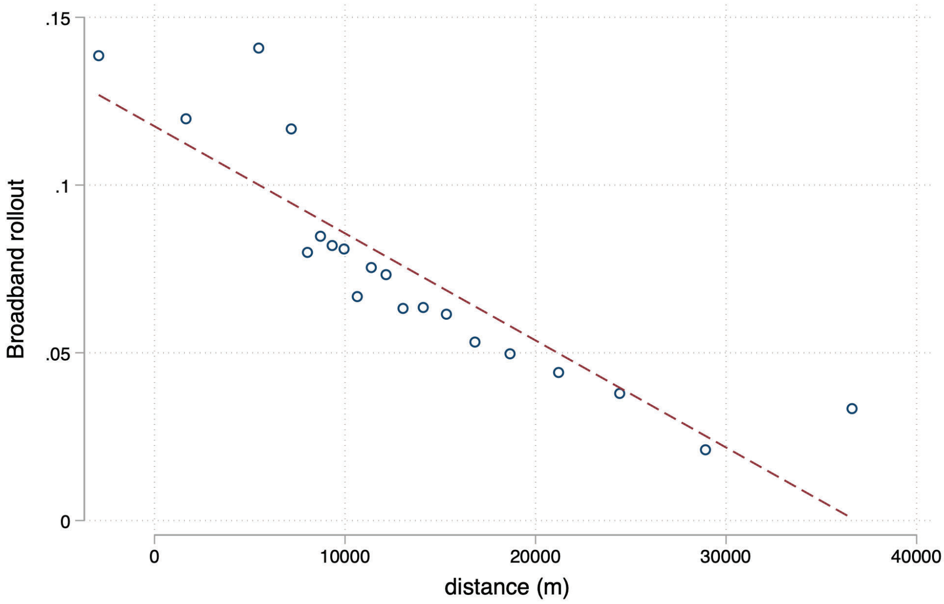


(b) Numeracy



Notes: The figure illustrates the evolution of absolute scores in the standardized tests before and after the outbreak of the pandemic.

Figure A3: Broadband rollout and distance from the infrastructure in t-1



Notes: The figure illustrates the relation between the distance from the closest treated municipality in $t-1$ and the likelihood for a municipality not-yet-treated to get access to the broadband infrastructure in t. Both variables have been residualized with respect to year fixed-effects, in order to allow a comparison across different years.

B Tables

Table A1: Quadratic value added model

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband	-0.000478	0.00892	-0.00484	-0.000652	-0.00530	-0.00124	-0.00630	-0.00170
	(0.00960)	(0.00842)	(0.0115)	(0.00921)	(0.0115)	(0.00921)	(0.0115)	(0.00923)
Test score (Grade 5)	0.592***	0.590***	0.644***	0.601***	0.644***	0.601***	0.644***	0.601***
	(0.00360)	(0.00274)	(0.00295)	(0.00212)	(0.00296)	(0.00212)	(0.00296)	(0.00212)
(Test score - Grade 5)^2	-0.0436***	-0.0358***	-0.0227***	-0.0176***	-0.0227***	-0.0176***	-0.0227***	-0.0176***
	(0.00199)	(0.00178)	(0.00139)	(0.00130)	(0.00139)	(0.00130)	(0.00139)	(0.00130)
Observations	1,091,993	1,075,547	1,091,309	1,074,847	1,091,309	1,074,847	1,091,309	1,074,847
R-squared	0.363	0.369	0.478	0.475	0.478	0.476	0.478	0.476
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year # Province FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents regression results of the quadratic version of the value-added model presented in Eq. 9. The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table A2: Heterogenous Covid effect

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.
School closures (weeks)	-0.00958***	-0.00595***	-0.00958***	-0.00594***	-0.00853***	-0.00571***
	(0.000406)	(0.000370)	(0.000406)	(0.000371)	(0.000565)	(0.000520)
Test score - Grade 5	0.566***	0.142***	0.566***	0.141***	0.570***	0.567***
	(0.00234)	(0.00111)	(0.00234)	(0.00118)	(0.00233)	(0.00182)
ESCS	0.136***	0.562***	0.136***	0.562***		
	(0.00112)	(0.00183)	(0.00119)	(0.00183)		
Covid cohort # ESCS			0.000107	0.000804***		
			(0.000251)	(0.000224)		
ESCS - II					0.139***	0.142***
					(0.00245)	(0.00243)
ESCS - III					0.214***	0.221***
					(0.00267)	(0.00264)
ESCS - IV					0.330***	0.339***
					(0.00308)	(0.00304)
Covid cohort # ESCS - II					-0.00230***	-0.00164***
					(0.000534)	(0.000508)
Covid cohort # ESCS - III					-0.000673	5.46e-05
					(0.000549)	(0.000525)
Covid cohort # ESCS - IV					0.000343	0.00225***
					(0.000662)	(0.000583)
Observations	1,230,028	1,209,851	1,230,028	1,209,851	1,230,028	1,209,851
R-squared	0.458	0.456	0.458	0.456	0.457	0.454
V grade School FE	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES
School variables	NO	NO	NO	NO	NO	NO
Peer effects	NO	NO	NO	NO	NO	NO
Municipality variables	NO	NO	NO	NO	NO	NO

Notes: This table presents regression results of the model presented in Eq. 10. The dependent variables are the unstandardised scores in math and literacy recorded in grade 8. While Table 7 analyses the homogeneous lockdown shock, this specification accounts for heterogeneity in the duration of school closures across different regions. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table A3: White areas only

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	0.0151	0.0168	0.0348**	0.00888
	(0.0135)	(0.0115)	(0.0136)	(0.0115)
Test score (Grade 5)	0.637***	0.601***		
	(0.00360)	(0.00260)		
ESCS	0.114***	0.120***		
	(0.00152)	(0.00148)		
Test score (Grade 5) > mdn.			0.950***	0.906***
			(0.00565)	(0.00425)
ESCS > mdn.			0.172***	0.191***
			(0.00347)	(0.00341)
Broadband # (Test score (Grade 5) > mdn.			-0.109***	-0.0679***
			(0.0121)	(0.0107)
Broadband # (ESCS > avg.)			0.0273***	0.0579***
			(0.00829)	(0.00848)
(ESCS > mdn) # (Test score (Grade 5) > mdn)			0.0573***	0.0314***
			(0.00541)	(0.00495)
Broadband # (ESCS > mdn) # (Test score (Grade 5) > mdn)			0.0696***	0.0277**
			(0.0125)	(0.0125)
Observations	791,131	780,063	791,131	780,063
R-squared	0.482	0.479	0.370	0.368
V grade School FE	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES
Year # Province FE	YES	YES	YES	YES
School variables	YES	YES	YES	YES
Peer effects	YES	YES	YES	YES
Municipality variables	YES	YES	YES	YES

Notes: This table presents regression results of the value-added model presented in Eq. 9, focusing only on the so-called 'white areas' (see Section 2.1). The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table A4: Stayers only

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	-0.0182	0.0101	-0.0142	-0.00162
	(0.0130)	(0.0104)	(0.0126)	(0.0103)
Test score (Grade 5)	0.656***	0.613***		
	(0.00354)	(0.00244)		
ESCS	0.113***	0.116***		
	(0.00151)	(0.00147)		
Test score (Grade 5) > mdn.			0.973***	0.920***
			(0.00555)	(0.00431)
ESCS > avg.			0.178***	0.195***
			(0.00362)	(0.00351)
Broadband # (ESCS > mdn.)			0.0306***	0.0567***
			(0.00755)	(0.00773)
Broadband # (Test score (Grade 5) > mdn.			-0.0764***	-0.0559***
			(0.0110)	(0.00980)
(ESCS > mdn) # (Test score (Grade 5) > mdn)			0.0495***	0.0220***
			(0.00560)	(0.00524)
Broadband # (ESCS > mdn) # (Test score (Grade 5) > mdn)			0.0560***	0.0175
			(0.0118)	(0.0114)
Observations	770,146	759,475	770,146	759,475
R-squared	0.484	0.480	0.366	0.364
V grade School FE	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES
Year # Province FE	YES	YES	YES	YES
School variables	YES	YES	YES	YES
Peer effects	YES	YES	YES	YES
Municipality variables	YES	YES	YES	YES

Notes: This table presents regression results of the value-added model presented in Eq. 9, excluding students that attended Grade VIII in a different municipality than the one where they attended Grade V. The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table A5: Broadband coverage - value added model

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
broadband coverage	0.0462***	0.0511***	0.000365	-0.00278	0.000226	-0.00291	4.40e-05	-0.00228
	(0.00561)	(0.00494)	(0.00693)	(0.00629)	(0.00694)	(0.00630)	(0.00697)	(0.00637)
Test score (Grade 5)	0.560***	0.588***	0.570***	0.561***	0.569***	0.561***	0.569***	0.561***
	(0.00320)	(0.00225)	(0.00278)	(0.00192)	(0.00278)	(0.00193)	(0.00278)	(0.00193)
Observations	1,506,626	1,454,204	1,506,563	1,454,144	1,506,563	1,454,144	1,506,563	1,454,144
R-squared	0.335	0.354	0.429	0.438	0.429	0.438	0.429	0.438
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year#Province FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents regression results of the value-added model presented in section 5. The specification coincides with the one estimated in Table 4, with the only exception of the main variable of interest, that in this case is the actual weighted average of broadband coverage measured at the level of the catchment area (see Section 4.2). The dependent variables are students' standardised numeracy and literacy scores in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table A6: Broadband coverage - interactions (1)

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband coverage	0.00134	-0.00108	0.0332***	0.0183***	-0.00140	-0.00373	-0.0293***	-0.0384***
	(0.00699)	(0.00637)	(0.00752)	(0.00699)	(0.00696)	(0.00636)	(0.00791)	(0.00736)
Test score (Grade 5)	0.591***	0.575***			0.569***	0.561***	0.573***	0.565***
	(0.00353)	(0.00252)			(0.00278)	(0.00193)	(0.00278)	(0.00192)
Broadband # Test score (Grade 5)	-0.0334***	-0.0210***						
	(0.00367)	(0.00278)						
Test score (Grade 5) - II			0.508***	0.544***				
			(0.00341)	(0.00301)				
Test score (Grade 5) - III			0.924***	0.945***				
			(0.00503)	(0.00356)				
Test score (Grade 5) - IV			1.472***	1.409***				
			(0.00756)	(0.00502)				
Broadband coverage # Test score (Grade 5) - II			-0.00738*	-0.0227***				
			(0.00410)	(0.00378)				
Broadband coverage # Test score (Grade 5) - III			-0.0549***	-0.0335***				
			(0.00556)	(0.00435)				
Broadband coverage # Test score (Grade 5) - IV			-0.0824***	-0.0309***				
			(0.00806)	(0.00587)				
ESCS	0.129***	0.135***	0.142***	0.146***	0.115***	0.119***		
	(0.00105)	(0.000990)	(0.000998)	(0.000944)	(0.00150)	(0.00139)		
Broadband coverage # ESCS					0.0222***	0.0260***		
					(0.00194)	(0.00172)		
ESCS - II							0.112***	0.116***
							(0.00306)	(0.00289)
ESCS - III							0.168***	0.171***
							(0.00334)	(0.00307)
ESCS - IV							0.277***	0.285***
							(0.00377)	(0.00358)
Broadband coverage # ESCS - II							0.0189***	0.0255***
							(0.00395)	(0.00380)
Broadband coverage # ESCS - III							0.0411***	0.0513***
							(0.00435)	(0.00400)
Broadband coverage # ESCS - IV							0.0562***	0.0665***
							(0.00495)	(0.00451)
Observations	1,845,269	1,823,496	1,845,269	1,823,496	1,845,269	1,823,496	1,845,269	1,823,496
R-squared	0.442	0.452	0.416	0.428	0.442	0.452	0.440	0.450
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year # Province FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents regression results of the value-added model presented in Eq. 9. The specification coincides with the one estimated in Table 5, with the only exception of the main variable of interest, that in this case is the actual weighted average of broadband coverage measured at the level of the catchment area (see Section 4.2). The dependent variables are the students' standardised numeracy and literacy scores recorded in grade 8. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Columns (3)-(4) and (7)-(8) report the interaction of the variable of interest with the quartile of the distribution of, respectively, student performance recorded in Grade 5 and the ESCS index. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table A7: Broadband coverage - interactions (2)

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	0.00992	-0.0167**	0.00955	-0.0163**
	(0.00710)	(0.00665)	(0.00709)	(0.00666)
Test score (Grade 5) > mdn.	0.914***	0.915***	0.896***	0.887***
	(0.00554)	(0.00393)	(0.00549)	(0.00388)
ESCS > mdn.	0.185***	0.207***	0.165***	0.185***
	(0.00313)	(0.00307)	(0.00311)	(0.00303)
Broadband # (ESCS > mdn.)	0.0397***	0.0537***	0.0399***	0.0539***
	(0.00392)	(0.00398)	(0.00389)	(0.00392)
Broadband # (Test score (Grade 5) > mdn)	-0.0784***	-0.0262***	-0.0771***	-0.0266***
	(0.00652)	(0.00501)	(0.00643)	(0.00493)
(ESCS > mdn) # (Test score (Grade 5) > mdn)	0.0628***	0.0330***	0.0660***	0.0377***
	(0.00485)	(0.00435)	(0.00481)	(0.00430)
Broadband # (ESCS > mdn) # (Test score (Grade 5) > mdn)	0.0321***	0.00281	0.0305***	0.00251
	(0.00621)	(0.00554)	(0.00616)	(0.00546)
Observations	1,845,269	1,823,496	1,845,269	1,823,496
R-squared	0.336	0.337	0.342	0.353
V grade School FE	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES
Year # Province FE	YES	YES	YES	YES
School variables	NO	NO	YES	YES
Peer effects	NO	NO	YES	YES
Municipality variables	NO	NO	YES	YES

Notes: This table presents regression results of the value-added model presented in Eq. 9. The specification coincides with the one estimated in Table 6, with the only exception of the main variable of interest, that in this case is the actual weighted average of broadband coverage measured at the level of the catchment area (see Section 4.2). The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

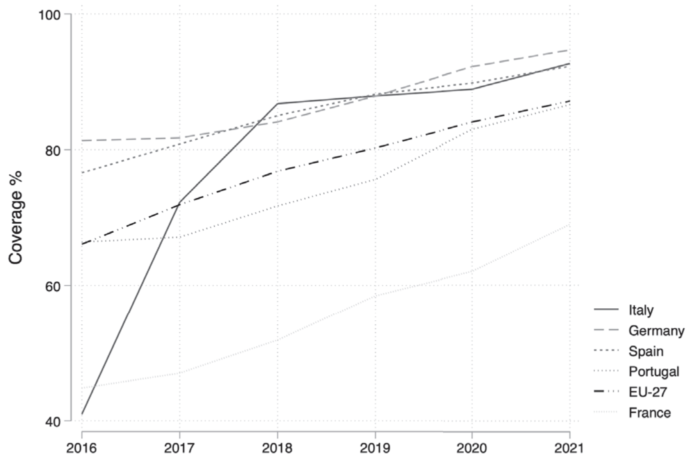
Table A8: Political influence

VARIABLES	(1) broadband	(2) broadband	(3) broadband	(4) broadband	(5) broadband
Same political party	-0.0977***	-0.0358***	-0.0383***	0.00613	0.00293
	(0.0101)	(0.00880)	(0.00870)	(0.00893)	(0.0115)
Constant	0.641***	0.609***	0.610***	0.587***	0.510***
	(0.00729)	(0.00610)	(0.00594)	(0.00572)	(0.00692)
Observations	9,378	9,378	9,363	9,361	7,808
R-squared	0.010	0.369	0.444	0.587	0.600
Municipality FE	NO	YES	YES	YES	YES
year FE	NO	NO	YES	-	-
year#region FE	NO	NO	NO	YES	-
year#province FE	NO	NO	NO	NO	YES

Notes: In this regression I test the hypothesis that local councils were able to influence the rollout of the infrastructure. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Figures and Tables in the main text

Figure 1: Broadband coverage (30Mbit/s)



Notes: The figure illustrates the evolution of broadband coverage in Italy and in other European countries.

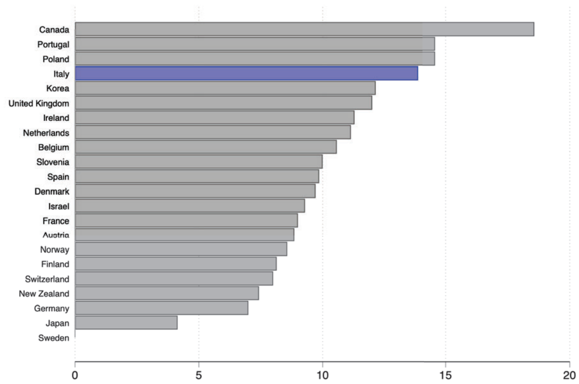
Figure 2: Broadband roll-out

(a) 2015

(b) 2020

Notes: The figure illustrates the spatial evolution of broadband coverage between 2015 and 2020.

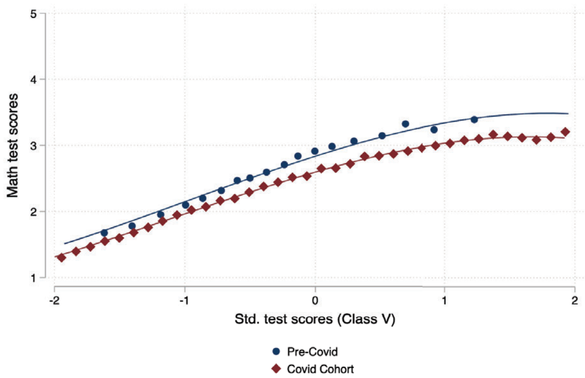
Figure 3: Duration of school closures in weeks, 2020/21



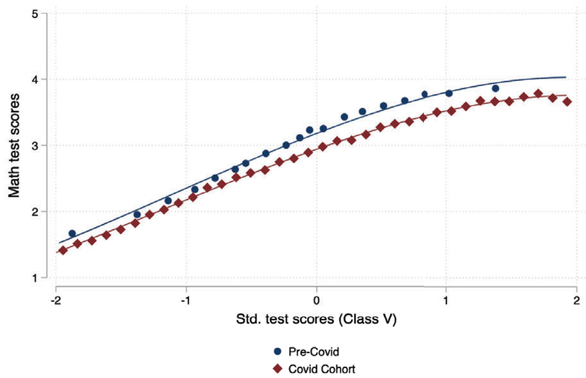
Notes: The figure illustrates the duration of school closures (in weeks) recorded by different countries over the period 2020-2021.

Figure 4: The impact of being in the Covid cohort and previous standardised scores

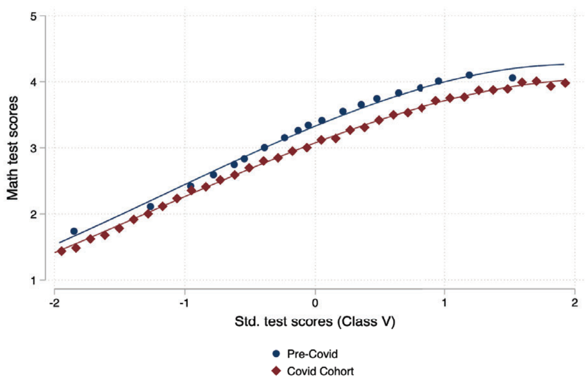
(a) I quartile of ESCS



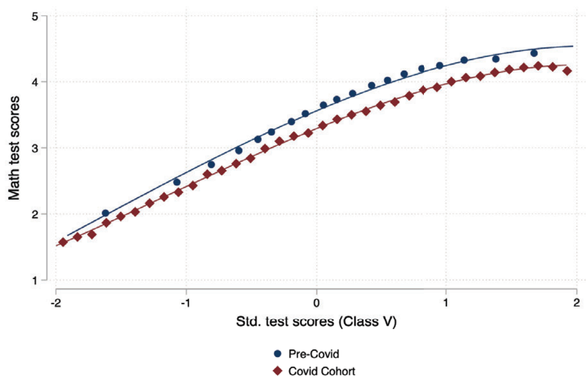
(b) II quartile of ESCS



(c) III quartile of ESCS



(d) IV quartile of ESCS



Notes: The figure illustrates impact of the pandemic on student performance in maths for students belonging to different quartiles of the ESCS distribution. The ESCS index is a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics.

Table 1: Summary Statistics

	Source	Obs. (Nb)	Mean	Sd	Min	Max
broadband coverage (30mbps)	Infratel	2,600,248	0.689	0.405	0	1
Internet speed, mbps	Infratel	2,600,248	37.74	29.91	0	1
broadband dummy	Infratel	2,600,248	0.653	0.476	0	1
Numeracy Test score	Invalsi	2,600,248	0	1	-4.537	4.472
Numeracy Test score (Grade 5)	Invalsi	2,600,248	0.0394	0.982	-5.146	4.246
Math Test score	Invalsi	2,600,248	0	1	-5.758	4.451
Math Test score (Grade 5)	Invalsi	2,600,248	0.0484	0.974	-5.594	4.344
BFMJ	Invalsi	2,600,248	3.753	0.986	1	5
BMMJ	Invalsi	2,600,248	3.167	1.164	1	5
HISEI	Invalsi	2,600,248	3.884	0.906	1	5
MISCED	Invalsi	2,600,248	3.178	1.203	1	5
FISCED	Invalsi	2,600,248	3.29	1.188	1	5
HISCED	Invalsi	2,600,248	3.541	1.167	1	5
ESCS	Invalsi	2,600,248	0	1	-3.9	2.731
Male	Invalsi	2,600,248	0.512	0.5	0	1
Full time	Invalsi	2,600,248	0.176	0.381	0	1
I Gen. migrant	Invalsi	2,600,248	0.0365	0.188	0	1
II Gen. migrant	Invalsi	2,600,248	0.07	0.255	0	1
Nursery	Invalsi	2,600,248	0.273	0.446	0	1
Preschool	Invalsi	2,600,248	0.813	0.39	0	1
Early enrolled	Invalsi	2,600,248	0.0839	0.277	0	1
Late enrolled	Invalsi	2,600,248	0.0701	0.255	0	1
Internet dummy (grade 5)	Invalsi	2,600,248	0.841	0.366	0	1
Computer Dummy (grade 5)	Invalsi	2,600,248	0.685	0.464	0	1
Male (class %)	Invalsi	2,600,248	0.512	0.0529	0	1
I Gen. migrant (class %)	Invalsi	2,600,248	0.0365	0.0396	0	1
Class size	Invalsi	2,600,248	20.7	3.897	1	67
Income per capita	Istat	2,600,248	8.755	0.303	7.682	9.805
Earners share	Istat	2,600,248	32.61	3.98	19.02	50.68
Foreign share	Istat	2,600,248	8.142	4.699	0.0763	35.21
High erner share	Istat	2,600,248	4.229	2.824	0	30.94
High income share	Istat	2,600,248	18.59	10.92	0	72.2

Notes: The table reports descriptive statistics for the variables used in the main specifications

Table 2: School-level analysis

	baseline specification			stacked specification	Callaway & Santanna (2021)
	(1)	(2)	(3)	(4)	(5)
VARIABLES	Num.	Num.	Num. (residual)	Num. (residual)	Num. (residual)
Broadband coverage	-0.0114				
	(0.00993)				
Broadband dummy		-0.00543	-0.00172	-0.00885	-0.0002
		(0.00734)	(0.00791)	(0.0129)	(0.013)
Observations	34,417	34,417	34,128	31,326	27,541
R-squared	0.605	0.605	0.453	0.694	-
	Lit.	Lit.	Lit. (residual)	Lit. (residual)	Lit. (residual)
Broadband coverage	0.00547				
	(0.00946)				
Broadband dummy		-0.000132	-0.00313	-0.00244	-0.0074
		(0.00693)	(0.00715)	(0.0108)	(0.0116)
Observations	34,394	34,394	34,038	31,148	27,476
R-squared	0.558	0.558	0.492	0.723	-
grade VIII school FE	YES	YES	YES	YES	YES
year#province FE	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO

Notes: This table presents regression results of the model presented in Eq. 7. In columns (1) and (2) the dependent variables are average students' standardised numeracy and literacy scores recorded in grade VIII. In columns (3) to (5) the dependent variable is instead the residualised school-time-year fixed effects. This variable can be interpreted as the average school performance in each subjects, net of the student skill distribution. The explanatory variable in column (1) is the weighted broadband coverage recorded in the catchment area. Other columns report a dummy that equal one when the weighted broadband coverage is above 75%. All specifications include school and year-time-province fixed effects. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 3: Student-level analysis - baseline

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband dummy	0.0119	0.0288***	-0.0116	-0.00857	-0.0125	-0.00963	-0.0128	-0.00969
	(0.00926)	(0.00815)	(0.0105)	(0.00840)	(0.0105)	(0.00842)	(0.0106)	(0.00843)
Male			0.136***	-0.258***	0.137***	-0.257***	0.137***	-0.257***
			(0.00263)	(0.00223)	(0.00263)	(0.00224)	(0.00263)	(0.00224)
Full time			0.123***	0.0966***	0.123***	0.0958***	0.123***	0.0958***
			(0.00715)	(0.00610)	(0.00706)	(0.00603)	(0.00706)	(0.00603)
I gen. migrant			-0.125***	-0.215***	-0.122***	-0.213***	-0.122***	-0.213***
			(0.00756)	(0.00743)	(0.00757)	(0.00743)	(0.00757)	(0.00743)
Nursery			0.0199***	0.00264	0.0200***	0.00271	0.0200***	0.00272
			(0.00283)	(0.00273)	(0.00282)	(0.00273)	(0.00282)	(0.00273)
Preschool			0.0720***	0.0755***	0.0716***	0.0751***	0.0717***	0.0752***
			(0.00621)	(0.00560)	(0.00620)	(0.00559)	(0.00620)	(0.00558)
Early enrolled			0.0271***	-0.0175***	0.0270***	-0.0176***	0.0270***	-0.0176***
			(0.00436)	(0.00427)	(0.00435)	(0.00427)	(0.00435)	(0.00427)
ESCS			0.237***	0.251***	0.236***	0.250***	0.236***	0.250***
			(0.00148)	(0.00138)	(0.00148)	(0.00137)	(0.00148)	(0.00137)
Internet dummy			0.125***	0.139***	0.124***	0.139***	0.124***	0.139***
			(0.00315)	(0.00313)	(0.00315)	(0.00312)	(0.00315)	(0.00312)
Computer dummy			-0.00412	0.00732***	-0.00400	0.00745***	-0.00401	0.00744***
			(0.00268)	(0.00259)	(0.00268)	(0.00259)	(0.00268)	(0.00259)
Constant	0.111***	0.102***	-0.144***	0.0376***	-0.140***	0.0412***	-0.144***	0.0431***
	(0.00364)	(0.00343)	(0.00663)	(0.00597)	(0.00664)	(0.00597)	(0.0154)	(0.0136)
Observations	1,108,194	1,108,102	1,107,505	1,107,408	1,107,505	1,107,408	1,107,505	1,107,408
R-squared	0.036	0.016	0.178	0.186	0.178	0.186	0.178	0.186
V grade School FE	NO	NO	YES	YES	YES	YES	YES	YES
VIII Grade School FE	NO	NO	YES	YES	YES	YES	YES	YES
Year#Province FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents regression results of the model presented in Eq. 8. The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 4: Value added model

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband dummy	-0.000678	0.00890	-0.00450	-0.000585	-0.00496	-0.00117	-0.00597	-0.00164
	(0.00973)	(0.00853)	(0.0116)	(0.00926)	(0.0116)	(0.00926)	(0.0116)	(0.00927)
Test score (Grade 5)	0.575***	0.585***	0.636***	0.599***	0.636***	0.599***	0.636***	0.599***
	(0.00424)	(0.00309)	(0.00318)	(0.00220)	(0.00318)	(0.00221)	(0.00318)	(0.00221)
Observations	1,091,993	1,075,547	1,091,309	1,074,847	1,091,309	1,074,847	1,091,309	1,074,847
R-squared	0.359	0.366	0.477	0.475	0.477	0.475	0.477	0.475
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year#Province FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents regression results of the value-added model presented in Eq. 9. The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-time-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 5: Interactions (1)

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband coverage	-0.00413	0.000505	0.0180	0.0289***	-0.00714	-0.00305	-0.0499***	-0.0544***
	(0.0115)	(0.00921)	(0.0120)	(0.00968)	(0.0116)	(0.00928)	(0.0130)	(0.0111)
Test score (Grade 5)	0.642***	0.605***			0.636***	0.599***	0.639***	0.603***
	(0.00319)	(0.00227)			(0.00318)	(0.00220)	(0.00317)	(0.00220)
Broadband # Test score (Grade 5)	-0.0407***	-0.0372***						
	(0.00561)	(0.00434)						
Test score (Grade 5) - II			0.535***	0.561***				
			(0.00323)	(0.00299)				
Test score (Grade 5) - III			0.980***	0.970***				
			(0.00462)	(0.00339)				
Test score (Grade 5) - IV			1.573***	1.459***				
			(0.00676)	(0.00460)				
Broadband coverage # Test score (Grade 5) - II			-0.00399	-0.0171***				
			(0.00649)	(0.00617)				
Broadband coverage # Test score (Grade 5) - III			-0.0248***	-0.0411***				
			(0.00774)	(0.00711)				
Broadband coverage # Test score (Grade 5) - IV			-0.0790***	-0.0801***				
			(0.0114)	(0.00915)				
ESCS	0.114***	0.119***	0.127***	0.132***	0.109***	0.114***		
	(0.00131)	(0.00127)	(0.00129)	(0.00126)	(0.00141)	(0.00136)		
Broadband coverage # ESCS					0.0280***	0.0333***		
					(0.00310)	(0.00295)		
ESCS - II							0.108***	0.113***
							(0.00293)	(0.00280)
ESCS - III							0.158***	0.164***
							(0.00317)	(0.00300)
ESCS - IV							0.263***	0.273***
							(0.00364)	(0.00357)
Broadband coverage # ESCS - II							0.0355***	0.0450***
							(0.00683)	(0.00694)
Broadband coverage # ESCS - III							0.0585***	0.0719***
							(0.00737)	(0.00727)
Broadband coverage # ESCS - IV							0.0767***	0.0882***
							(0.00826)	(0.00778)
Observations	1,091,309	1,074,847	1,091,309	1,074,847	1,091,309	1,074,847	1,091,309	1,074,847
R-squared	0.477	0.475	0.449	0.447	0.477	0.475	0.475	0.473
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year # Province FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents regression results of the value-added model presented in Eq. 9. The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. Columns (3)-(4) and (7)-(8) report the interaction of the variable of interest with the quartile of the distribution of, respectively, the ESCS index and student performance recorded in Grade V. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 6: Interactions (2)

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	-0.00571	-0.00876	-0.00690	-0.0126
	(0.0115)	(0.00949)	(0.0115)	(0.00941)
Test score (Grade 5) > mdn.	0.960***	0.933***	0.942***	0.903***
	(0.00513)	(0.00378)	(0.00506)	(0.00374)
ESCS > mdn.	0.189***	0.212***	0.171***	0.191***
	(0.00307)	(0.00305)	(0.00305)	(0.00301)
Broadband # (ESCS > mdn.)	0.0317***	0.0545***	0.0311***	0.0552***
	(0.00652)	(0.00676)	(0.00649)	(0.00666)
Broadband # (Test score (Grade 5) > mdn)	-0.0668***	-0.0637***	-0.0652***	-0.0585***
	(0.00966)	(0.00855)	(0.00954)	(0.00850)
(ESCS > mdn) # (Test score (Grade 5) > mdn)	0.0578***	0.0253***	0.0613***	0.0302***
	(0.00483)	(0.00446)	(0.00480)	(0.00440)
Broadband # (ESCS > mdn) # (Test score (Grade 5) > mdn)	0.0523***	0.0238**	0.0509***	0.0227**
	(0.0101)	(0.0100)	(0.0101)	(0.00994)
Observations	1,091,309	1,074,847	1,091,309	1,074,847
R-squared	0.360	0.347	0.366	0.365
V grade School FE	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES
Year # Province FE	YES	YES	YES	YES
School variables	NO	NO	YES	YES
Peer effects	NO	NO	YES	YES
Municipality variables	NO	NO	YES	YES

Notes: This table presents regression results of the value-added model presented in Eq. 9. The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-time-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. To simplify the interpretation, the main variable of interest is interacted with 2 dummies that are equal to one when the distribution of the performance in Grade V or the ESCS score is above the median. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 7: School closures

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.
Covid cohort	-0.104***	-0.0726***	-0.104***	-0.0737***	-0.0963***	-0.0762***
	(0.00308)	(0.00279)	(0.00310)	(0.00281)	(0.00467)	(0.00449)
Test score - Grade 5	0.567***	0.563***	0.567***	0.563***	0.570***	0.567***
	(0.00234)	(0.00184)	(0.00234)	(0.00184)	(0.00234)	(0.00183)
ESCS	0.136***	0.142***	0.135***	0.139***		
	(0.00112)	(0.00111)	(0.00123)	(0.00125)		
Covid cohort # ESCS			0.00322	0.0115***		
			(0.00204)	(0.00198)		
ESCS - II					0.138***	0.141***
					(0.00258)	(0.00260)
ESCS - III					0.213***	0.218***
					(0.00280)	(0.00280)
ESCS - IV					0.328***	0.335***
					(0.00321)	(0.00322)
Covid cohort # ESCS - II					-0.0154***	-0.0104**
					(0.00483)	(0.00475)
Covid cohort # ESCS - III					-0.00479	0.0107**
					(0.00495)	(0.00492)
Covid cohort # ESCS - IV					0.00582	0.0270***
					(0.00543)	(0.00524)
Observations	1,130,346	1,086,612	1,130,346	1,086,612	1,130,346	1,086,612
R-squared	0.454	0.449	0.455	0.449	0.453	0.447
V grade School FE	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES
School variables	NO	NO	NO	NO	NO	NO
Peer effects	NO	NO	NO	NO	NO	NO
Municipality variables	NO	NO	NO	NO	NO	NO

Notes: This table presents regression results of the model presented in Eq. 10. The dependent variables are the unstandardised scores recorded in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade 5 and grade 8 school fixed effects. The covid cohort variable is a dummy that equals one when students took the standardised exam in the academic year 2020/21. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 8: Broadband effect before and after the pandemic shock

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.
Panel A						
Broadband	0.0373***	0.0191**	-0.00505	-0.00806	0.0309***	0.00678
	(0.0104)	(0.00934)	(0.0103)	(0.00930)	(0.0106)	(0.00965)
Test score (Grade 5) > avg.	0.962***	0.923***	0.905***	0.912***	0.938***	0.910***
	(0.00565)	(0.00456)	(0.00453)	(0.00365)	(0.00715)	(0.00605)
ESCS > avg.	0.262***	0.279***	0.220***	0.221***	0.188***	0.201***
	(0.00322)	(0.00313)	(0.00380)	(0.00376)	(0.00481)	(0.00509)
Broadband # (ESCS > avg.)			0.0214***	0.0270***	0.0185***	0.0303***
			(0.00444)	(0.00439)	(0.00560)	(0.00587)
Broadband # (Test score (Grade 5) > avg)	-0.0672***	-0.0275***			-0.0890***	-0.0361***
	(0.00689)	(0.00541)			(0.00887)	(0.00732)
(ESCS > avg) # (Test score (Grade 5) > avg)					0.0534***	0.0334***
					(0.00760)	(0.00748)
Broadband # (ESCS > avg) # (Test score (Grade 5) > avg)					0.0260***	0.00358
					(0.00893)	(0.00859)
Panel B						
Broadband # Covid cohort	-0.0141*	-0.0122	-0.00889	-0.0127*	-0.0107	-0.0152*
	(0.00829)	(0.00793)	(0.00855)	(0.00758)	(0.00928)	(0.00889)
Test score (Grade 5) > avg. # Covid cohort	-0.0171*	0.00496	0.00408	-0.0104***	0.00660	0.0251**
	(0.00917)	(0.00847)	(0.00413)	(0.00386)	(0.0124)	(0.0115)
ESCS > avg. # Covid cohort	-0.0257***	-0.0375***	0.0198**	0.0281***	0.0415***	0.0460***
	(0.00365)	(0.00356)	(0.00780)	(0.00749)	(0.00974)	(0.0107)
Broadband # (ESCS > avg.) # Covid cohort			0.00608	0.0102	-0.00892	0.00694
			(0.00879)	(0.00842)	(0.0110)	(0.0119)
Broadband # (Test score (Grade 5) > avg) # Covid cohort	0.0189*	0.00797			0.00608	0.00492
	(0.0103)	(0.00946)			(0.0141)	(0.0131)
(ESCS > avg) # (Test score (Grade 5) > avg) # Covid cohort					-0.0447***	-0.0365**
					(0.0155)	(0.0153)
Broadband # (ESCS > avg) # (Test score (Grade 5) > avg) # Covid cohort					0.0268	0.00551
					(0.0173)	(0.0171)
Observations	1,230,028	1,209,851	1,230,028	1,209,851	1,230,028	1,209,851
R-squared	0.376	0.379	0.376	0.379	0.376	0.379
V grade School FE	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES
Year # Province FE	YES	YES	YES	YES	YES	YES
School variables	YES	YES	YES	YES	YES	YES
Peer effects	YES	YES	YES	YES	YES	YES
Municipality variables	YES	YES	YES	YES	YES	YES

Notes: This table presents regression results of the value-added model presented in Eq. 9. The dependent variables are the students' standardised numeracy and literacy scores recorded in grade VIII. The ESCS index provides a proxy of student individual economic, social and cultural status, based on known and unknown family characteristics. A complete list of the covariates included in the model is reported in Table 1. All regressions include grade V and grade VIII school fixed effects, as well as year-province fixed effects. Italy is divided in 110 provinces, corresponding to EU NUTS3 regions. In panel A, I report the main covariates used in the model estimated in table 6. The coefficients refer to the effect of the policy before and after the pandemic shock. Covariates in panel B, interacted with the covid cohort dummy, illustates the way the policy effect changed during the pandemic. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.