High-speed Broadband and Educational Achievements

Abstract

This study sheds new light on the short-run effects of access to high-speed internet on educational disparities. 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 gradual rollout of a national infrastructural policy associated with an increase in 30 Mbit/s household broadband coverage from 40% to 80% over a 6-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 literacy and maths. However, these results mask substantial heterogeneity: low performers in grade 5 and students with more advantaged backgrounds benefit from access to high-speed broadband, whereas the opposite is true for other students. Overall, the findings suggest that access to broadband widened performance disparities across students with different socioeconomic backgrounds.

Keywords: ICT, education, economics, internet, broadband, inequalities
JEL Codes: I24, H54, O18

1. Introduction

Nowadays, home computers have become an indispensable educational tool in developed economies. According to the latest 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 access to and quality of home computing. This digital divide is often related to the varying levels of access to high-speed internet connections. To close this gap, many governments have invested heavily in broadband infrastructure.[1]

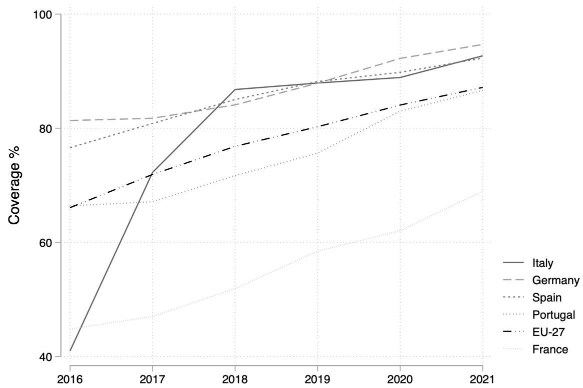
Despite the widespread commitment to accelerate the digital transition, the actual impact of ICT on student performance is still debated (Machin et al., 2007; Barrera-Osorio and Linden, 2009; Checchi et al., 2019; Cristia et al., 2017; Sanchis-Guarner et al., 2021). In particular, there is still no consensus on the effects of access to high-speed internet broadband on learning outcomes. Broadband penetration can affect educational outcomes through multiple, and potentially offsetting, channels. From a human capital accumulation perspective, high-speed internet relaxes constraints on access to educational technologies—such as online platforms, interactive learning software, and computer-assisted instruction—that complement in-person learning, improve learning productivity, and may yield disproportionate gains for lower-achieving students. Conversely, broadband also facilitates access to non-educational digital content, including streaming media and online gaming, which may displace time allocated to cognitively enriching activities. The net effect is therefore contingent on heterogeneous patterns of use, the quality of digital inputs consumed, and the extent of parental or institutional oversight.

This paper aims to estimate the causal impact of high-speed internet availability on student performance, leveraging the staggered rollout of Italy's National Ultra-Broadband Plan (NUBP). Launched in 2015, the programme aimed to provide universal access to 30 Mbit/s

[1] Within Europe, such policies were promoted under the "Europe 2020 Strategy" and its "Digital Agenda for Europe," which targeted universal access to 30 Mbit/s connections by 2020.

connections by 2020. Although this target was not fully met, over the implementation period Italy—historically among the lowest performers in the PISA assessments within the OECD (Oecd, 2014)—significantly narrowed its longstanding gap with other European countries in access to next-generation broadband services (Figure 1–2).

Figure 1: Broadband coverage (30Mbit/s)



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

Figure 2: Broadband roll-out

(a) 2015

(b) 2021

Notes: The figure illustrates the spatial evolution of broadband coverage between 2015 and 2021.

This paper makes three main contributions. First, it leverages the staggered rollout of a large-scale infrastructure policy to address the endogeneity concerns that typically affect similar studies. The NUBP expanded coverage progressively across adjacent municipalities to minimise costs while ensuring rapid nationwide implementation,[2] making timing plausibly unrelated to factors influencing educational outcomes.

Second, it combines highly granular broadband data at the census tract level with a rich longitudinal student dataset. These data allow for precise measurement of actual internet access around each school and help mitigate concerns about confounding factors and sorting. The analysis tracks six student cohorts over time (in grades 5 and 8), employing an empirical design that absorbs time-invariant school and municipality characteristics, as well as time-varying shocks at the travel-to-work area level. Conditioning on grade 5 baseline performance further adjusts for persistent individual differences in ability. To sharpen identification, the sample excludes students without any internet access three years prior to treatment[3] and those already covered by high-speed broadband at the beginning of the study.

Finally, the inclusion of prior achievement and rich student-level covariates allows for the investigation of heterogeneous policy effects. Examining differences by family background and prior performance highlights the distributional consequences of broadband expansion and its implications for education policy.

The paper is organised as follows. Section 2 discusses the relevant literature. Section 3 provides a general background of the Italian school systems and describes the main features of the NUBP. Section 4 describes the different data sources used and the procedure implemented to define school catchment areas. Section 5 describes the empirical strategy. In sections 6, 7, and 8, I present the results and discuss the main policy implications.

[2] Figure A2 shows that most of the expansion in high-speed broadband occurred across neighbouring municipalities.

[3] Only 4% of students reported lacking internet access at home; an additional 8% with missing information are also excluded.

2. Literature

This paper relates to a growing literature on the socio-economic consequences of broadband expansion. Previous research documents effects on employment and productivity (Akerman et al., 2015; Duvivier and Bussière, 2022), business location (McCoy et al., 2018), electoral outcomes (Falck et al., 2014; Campante et al., 2018), marriage rates (Bellou, 2015), and housing prices (Ahlfeldt et al., 2017; Wolf and Irwin, 2024). Within this broader field, this study focuses on the relationship between ICT and educational outcomes.

Access to digital technologies has traditionally varied with financial resources, geography, public support, and parental choice. This "first-order" digital divide has narrowed in most advanced economies with the diffusion of broadband infrastructure, mobile technology, and greater recognition of ICT's role in education (Tondeur et al., 2011; Van Deursen and Van Dijk, 2019; Bulman and Fairlie, 2016; Escueta et al., 2020). Nevertheless, inequalities remain in how technologies are used—the so-called "second-order" or digital use divide—reflecting heterogeneity in user characteristics, patterns of use, and contextual factors (Brotman, 2016; Camerini et al., 2018; Falck et al., 2018; Vedechkina and Borgonovi, 2021).

The ambiguous and context-dependent effects of ICT on student performance reflect the multiple, and potentially offsetting, channels through which digital technologies influence children's wellbeing, cognitive development, and learning outcomes. On the one hand, broadband facilitates access to educational technology resources (Ed-tech). Online courses and educational videos, once primarily used in higher education, are increasingly integrated into secondary schooling and complement classroom instruction (Jackson and Makarin, 2018). Computer-assisted learning (CAL) tools allow students to practise skills with immediate feedback, and several studies have documented significant improvements in student performance, particularly in mathematics, when such tools are deployed (Borman et al., 2009; Barrow et al., 2009; Roschelle et al., 2010, 2016; Karam et al., 2017). Evidence also suggests that these resources may disproportionately benefit lower-performing students (Banerjee et al., 2007; Barrow et al., 2009; Muralidharan et al., 2019).[4]

On the other hand, a substantial and varied body of literature highlights several challenges associated with increased access to digital technologies. For example, online gaming and other digital entertainment may increase the time spent on leisure activities (Kirschner and Karpinski, 2010). Even when the total amount of leisure time remains constant, screen time can reduce the time spent on more enriching activities (Mutz et al., 1993; Linebarger et al., 2014). However, recent research suggests that this relationship is largely dependent on content quality[5] and the level of parental involvement (Barr et al., 2008). Yet concerns remain. Moreover, several studies link excessive screen use to behavioural problems (Sharif et al., 2010), ADHD (David Acevedo-Polakovich et al., 2007), and sleep difficulties (Cain and Gradisar, 2010; Hale and Guan, 2015). Taken together, these findings suggest that, under certain conditions, ICT can hinder student performance.

Building on these insights, this paper contributes to the recent literature that seeks to identify, within a causal framework, the heterogeneous and context-specific effects of high-speed internet on student performance (Sanchis-Guarner et al., 2021; Cambini et al., 2021). Specifically, the paper advances three hypotheses: (i) high-speed internet may enhance achievement by facilitating access to effective educational tools; (ii) it may also reduce achievement by crowding out study time and hindering cognitive development, particularly under limited parental supervision; (iii) the balance of these effects likely depends on students' prior performance and family background. Appendix C provides a simple framework linking these hypotheses to the empirical design.

Empirical work has investigated the ICT–education nexus using a range of strategies.

[4]The extent to which students with varying levels of prior performance benefit may also depend on the grading system. Under relative grading, where performance is evaluated against peers, the impact of additional resources is ambiguous. Under absolute grading, such as in Italian lower secondary schools, where assessment is based on fixed objectives, high-performing students face limited room for improvement, increasing the likelihood that low-performing students benefit more.

[5]See Fisch et al. (2005); Schmidt and Anderson (2009); Kostyrka-Allchorne et al. (2017) for television and Przybylski and Wang (2016), Ventura et al. (2013) Feng and Spence (2018) and Adachi and Willoughby (2013) for gaming.

A first strand evaluates the introduction 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; Karam et al., 2017; Falck et al., 2018).[6] and policies promoting educational software (Banerjee et al., 2007; Barrow et al., 2009; Barrera-Osorio and Linden, 2009; Leuven et al., 2007; Muralidharan et al., 2019).[7] Other studies focus instead on the introduction of home computer technologies (Beltran et al., 2006; Fairlie et al., 2010; Fiorini, 2010; Malamud and Pop-Eleches, 2011; Fairlie and Robinson, 2013; Vigdor et al., 2014; Cristia et al., 2017).[8]
A second strand exploits long longitudinal surveys (Kirschner and Karpinski, 2010; Tondeur et al., 2011; Fairlie, 2016; Camerini et al., 2018; Van Deursen and Van Dijk, 2019). These studies provide interesting insights about the heterogeneous effect of internet on student performance, but the use of self-reported measures of internet access and the limited ability to control for confounding factors limit their capacity to identify a causal link.
A third strand employs randomised and quasi-experimental approaches. RCTs provide internally valid evidence but limited external validity (Jackson and Makarin, 2018; Borman et al., 2009; Roschelle et al., 2010, 2016). For example, Malamud et al. (2019) randomly assigned students laptops and/or internet access and found no significant effects on maths or reading. More recently, quasi-experimental studies have exploited large-scale rollouts of broadband infrastructure. Cambini et al. (2021) study broadband rollout and student performance using Italian municipality-level data, while Sanchis-Guarner et al. (2021) exploit discontinuities across telephone exchange boundaries in the UK.

While related to these contributions, this paper deviates in several respects. First, rather than exploiting geographical discontinuities in distance to telephone exchanges as in Sanchis-Guarner et al. (2021) and Falck et al. (2014), I exploit variation over time in the staggered rollout of the Italian National Ultra-Broadband Plan. This programme expanded household

[6]Findings are mixed, reflecting both contextual heterogeneity and difficulties in measuring actual ICT use.

[7]These studies generally find positive effects on maths and reading outcomes.

[8]These papers typically find positive effects on cognitive skills and outcomes directly related to computer access, but more limited or null effects on academic achievement.

coverage at 30 Mbit/s from 40% to 80% over five years, enabling comparisons between students in areas gaining access during the period and those remaining disconnected. Second, unlike Cambini et al. (2021), who exploit discrete municipality-level data,[9] this paper exploits census-tract-level coverage (share of buildings with coverage of at least 30/100 Mbit/s). This allows for a more precise identification of exposure at the school catchment level, which is crucial since many students attend schools outside their municipality of residence.[10] Moreover, this work focuses only on grade 8 students, for whom digital technologies are especially relevant, and it conditions on prior performance in grade 5. This strategy helps to control for persistent individual ability and allows us to examine heterogeneity by student background. Finally, the empirical design incorporates both grade 5 and grade 8 school fixed effects, accounting for potential long-term school influences on student trajectories.

3. Institutional Background

3.1. The National Ultra-Broadband Plan

In 2014, the Italian Government set up the *'National Ultra-Broadband Plan'* (Piano Nazionale Banda Ultra-Larga - NUBP), a large-scale programme aimed at ensuring universal coverage of 30 Mbit/s and 85% coverage of 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 divided into three areas based on existing or expected infrastructure availability:

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

[9]The primary independent variable used is a simple binary indicator denoting whether the student's school is located in a municipality outfitted with high-speed broadband. The criteria for classifying a full municipality - even a large city - as having adequate broadband coverage remain unspecified.

[10]According to a recent report produced by the National Institute of Statistics ISTAT (2021), 30% of Italian students attended a school located in a different municipality. This share is likely to be much higher in small towns, where most of the variation in broadband coverage over time originates.

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 where there would have been in the following three years, at least two basic broadband networks of different operators.

The NUBP rested on four main pillars. First, the State guarantees administrative simplification and a reduction in burdens for all 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 access to financial resources, the establishment of a guarantee fund and access to credit at subsidised rates. Finally, in white areas (commonly defined as *'market failure areas'*), the Public Sector intervenes directly to realise the infrastructures.[11]

In 2014, most Italian households did not have access to modern fibre internet technologies and mainly relied on old infrastructures, offering an average broadband speed below 2 Mbit/s. Following delays due to legal disputes, implementation began in 2015 and quickly produced measurable effects. Between 2015 and 2020 Italy, while falling short of its 100% target, managed to significantly reduce the historical gap with the other large European countries, by doubling the share of households 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*, the *European Agricultural Fund for Rural Development* and the *Development and Cohesion Fund*), which complemented public funds.

[11] In this study, I focus on the whole Italian territory. However, in Section 7, I replicate the main analyses focusing exclusively on white areas, obtaining similar results.

3.2. The Italian School System

The poor performance of Italian students in PISA tests is often linked to an education system characterised by outdated curricula, highly centralised governance, and limited autonomy for schools.

Compulsory education covers ten years, from ages 6 to 16,[12] and includes primary (ages 6–11) and lower secondary school (ages 11–14), both of which culminate in a national examination required to progress to the next stage. Examinations combine a uniform written test designed by INVALSI[13] (covering Italian, mathematics, science, informatics, and foreign languages) with additional written tests set by a mixed committee. Primary and lower-secondary schools are characterised by a very low, if not entirely absent, degree of autonomy (Ichino and Tabellini, 2014). Weekly instruction is fixed at 30 hours[14], and the Ministry of Education defines both the content and the number of hours allocated to each subject, authorises a limited number of textbooks, and requires schools to align their educational plans with national goals.[15] Schools may use a limited budget for laboratories or optional activities, but 97—100% of their resources come from central government transfers.[16]

The uniformity of the system is also guaranteed by human resource management, which is primarily conducted at the national level. Teachers apply to province-level lists and are allocated to schools strictly by rank and preference, with minimal discretion for principals. Salaries and career progression are set by national agreements, leaving little scope for local variation.[17] Another important feature of the lower secondary education system is the lim-

[12]Over the last decades, the country has experienced a reasonable decrease in the number of early high school 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.

[13]Italian National Institute for the Evaluation of the Education System

[14]School Councils can offer 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.

[15]Teaching methods and content 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.

[16]Resources are allocated every three years based on enrolments, disability incidence, and other criteria.

[17]Recent reforms have tried to introduce modest performance pay schemes, but strong opposition from powerful school unions has preserved the status quo.

ited competition among schools. Class sizes typically range from 15 (down to 10 in remote areas) to 26 students, with adjustments financed by the central government. When applications exceed available places, schools are expected to prioritise geographic proximity in admissions, which further limits differentiation across institutions.

Taken together, these features produce a high degree of homogeneity across schools. While regional disparities remain—reflecting differences in infrastructure and teacher composition—the evidence suggests that service quality is broadly uniform within provinces (NUTS-3 areas). For this study, such within-province uniformity is crucial, as it mitigates the influence of confounders that typically complicate evaluations of large-scale infrastructure policies in other settings.

4. Data

4.1. Main Sources

In this study, I construct a unique dataset that links microdata on student achievements to spatial data on internet broadband coverage. The resulting pupil-level dataset is further enriched with regional, municipal, and school-level information, 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 responsible for the annual assessment of the competencies of Italian students in both reading and mathematics. Tests are administered nationally at grades 2, 5, 6, 8, and 10. Each year, the institute releases anonymised microdata on student performance. Since 2008, individual marks have been linked with a rich set of individual information, enabling control for personal, family, and school characteristics. I observe test results for the full population of students in the 2012/2013 through 2021/2022 school years,[18] at grades 5 and 8, and extract information for six student cohorts: (i) 522,000 students born in 2000, who completed grade 5 in 2013

[18]The 2020 standardised test, scheduled for May 2020, was cancelled due to the pandemic.

and grade 8 in 2016; (ii) 529,000 students born in 2001, completing grade 5 in 2014 and grade 8 in 2017; (iii) 554,000 students born in 2002, completing grade 5 in 2015 and grade 8 in 2018; (iv) 546,000 students born in 2003, completing grade 5 in 2016 and grade 8 in 2019; (v) 527,000 students born in 2005, completing grade 5 in 2018 and grade 8 in 2021; and (vi) 544,000 students born in 2006, completing grade 5 in 2019 and grade 8 in 2022.

Our analysis focuses on students who did not have broadband access prior to the policy and can therefore be regarded as 'yet-to-be-treated' at the start of the period. As Figure 2 shows, fibre was initially available only in the largest city centres. These students represent about 26% of the sample and, on average, come from more advantaged family backgrounds (see Figure A1). From the full dataset, I also exclude all observations with missing values in at least one of the key covariates (see Table 1).[19] Furthermore, I exclude students reporting no internet connection at home three years before the treatment. While only a very small number of students did not have an internet connection in the pre-treatment period, their inclusion would complicate the interpretation of the results, since access to the broadband can provide an incentive to adopt home internet (regardless of the connection speed).[20]

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, reporting personal information such as the country of origin, the year of arrival in Italy, family background, and the conditions and facilities for studying at home. Notably, the questionnaire also provides information about early childcare, which has been shown to strongly affect children's cognitive outcomes, especially among migrants (Corazzini et al., 2021).

The results, along with information provided by school secretaries, are matched to student scores and included in a dataset available for research purposes. In addition to rich student-

[19]In Table A11, I estimate the baseline model using the full population while excluding covariates with missing values. The estimates are consistent with the baseline results.

[20]Only 4% of students reported not having any access to internet at home, but I exclude a further 8% for whom this information was missing. In Table A10, I replicate the analysis including these students, obtaining similar results.

level information, I have access to the level of education (according to the ISCED scale) of both parents, as well as their occupational status, recorded by the Socio-Economic Index of Occupational Status (ISEI) and an index based on individual-level Economic, Social, and Cultural Status (ESCS). I also have access to data on home computer ownership and availability of an internet connection before and after the policy rollout.

School-level data come from the Ministry of Education (MIUR). In 2011, in accordance with the Community Guideline on public access to information held by public authorities, the Miur began publishing data on each State-recognised school at 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's basic budget, and a self-evaluation document submitted to the Ministry at the end of the academic year. Furthermore, each school provides the full address of each building (plesso) associated with it.

Since INVALSI does not provide the students' addresses, I compute catchment areas around each school, by exploiting the 2011 Italian Census. The Census is a universal survey conducted every ten years by the National Institute for Statistics (ISTAT)[21]. 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

[21]The survey is divided into 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.

data on the Italian territory, which is divided into 402,000 areas. On average, each census tract hosts 142 people and, for each one, ISTAT releases information concerning the number of residents and the breakdown by gender and age class. Furthermore, the dataset can be matched to information on wages, occupational status, and other social features on the basis of a 5-10% population sample living in each division.

Broadband data are obtained from Infratel, the state-owned company responsible for implementing the broadband plan. Since 2015, Infratel has conducted annual surveys of all internet providers, which are required to report the availability of broadband infrastructure and investment plans for the following three years. The data are collected at the house-number level and aggregated to the census-area level. Specifically, Infratel provides:

1. The share of buildings covered by at least 100 Mbit/s;
2. The share of buildings covered by at least 30 Mbit/s.

These data can be linked to spatial datasets providing census-area boundaries, enabling direct matching with student and school information.

In the final dataset, information relating to individual students is matched to 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 of 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 proximity to the school. A complete description of the variables used in the analysis is reported in 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, which guarantee minimum speeds of

Table 1: Summary Statistics

	Source	Obs. (Nb)	Mean	Sd	Min	Max
Internet access						
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	100
Broadband dummy	Infratel	2,600,248	0.653	0.476	0	1
Test scores						
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
Student characteristics						
Occupational status of the father (BFMJ)	Invalsi	2,600,248	3.753	0.986	1	5
Occupational status of the mother (BMMJ)	Invalsi	2,600,248	3.167	1.164	1	5
Highest occ. status of the parents (HISEI)	Invalsi	2,600,248	3.884	0.906	1	5
Educational level of the father (FISCED)	Invalsi	2,600,248	3.29	1.188	1	5
Educational level of the mother (MISCED)	Invalsi	2,600,248	3.178	1.203	1	5
Highest parental educational level (HISCED)	Invalsi	2,600,248	3.541	1.167	1	5
Family background (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
Class characteristics						
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
Municipality characteristics						
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 earner 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

30 Mbit/s and 100 Mbit/s, respectively. Since I am interested in simple access to fibre internet technologies, my variable of interest is constructed from the share of families with access to at least 30 Mbit/s internet speed: $BA\_vt = X^{FTTH} + X^{FTTN}$.

Ideally, 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 attend a specific school.

The strategy relies on the specific features of the Italian lower secondary school system. As discussed in Section 3.2, school enrolment follows rigid geographic criteria, which leave parents little discretion in the choice of the school. 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 appears to be a suitable method to link students to broadband availability.

This strategy exploits the 2011 Italian Census, 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 c and the nearest census areas;

3. for each school, neighbouring census areas v are sorted by distance (in ascending order);

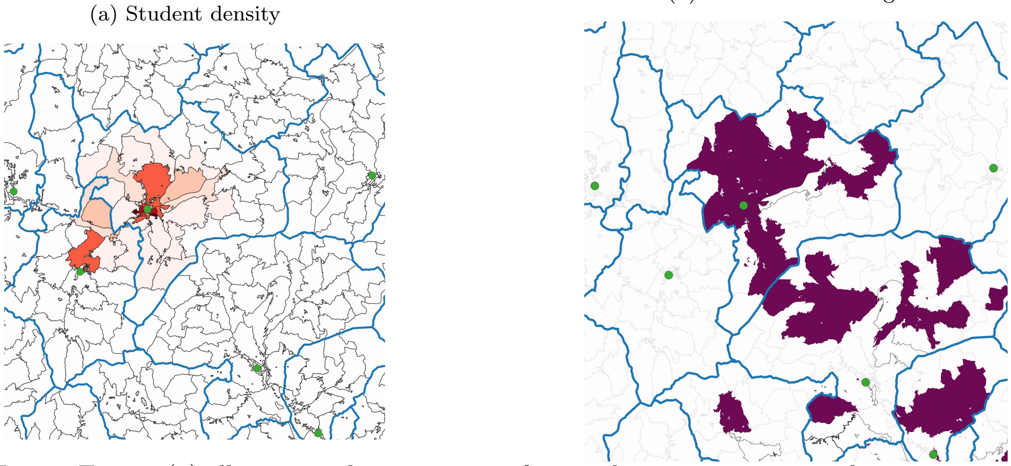
4. compute the areas' cumulative relevant population;

5. select the closest N areas so that the cumulative relevant population contains a multiple k of the number of students enrolled in school c.

Figure 3: School catchment areas

(a) Student density

(b) Broadband coverage



Notes: Figure (a) illustrates the extension of a catchment area across the territory of two municipalities (red areas). Figure (b) highlights the census areas where more than 75% of households have access to high-speed broadband. Blue lines define the municipality borders.

Once the data have been extracted and the catchment areas have been defined, I can build a proxy for broadband coverage, BA_{ct}, obtained as the weighted average of broadband coverage measured in the catchment area of school c. The weights used are the share of students living in each census area v over 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 access to (at least) a 30 Mbit/s internet broadband in the census area v, belonging to school c's catchment area, n_{vct} is the number of students living in the census area v, and N_{ct} is the total number of students living in the catchment area. Figure 3 illustrates the extent of a school's catchment area across contiguous census areas, alongside the corresponding broadband coverage recorded in the same territory. Finally, in order to simplify the interpretation of the results, a simple dummy variable BA^{d}_{ct} is produced that equals 1 when a catchment area records a weighted coverage above 75%. This approximation does not involve a relevant loss of information, since over the period considered, 96% of catchment areas recorded a weighted broadband coverage above 75% or below 5%. This reflects the rollout design, whereby coverage was extended progressively to entire municipalities at once, quickly shifting most catchment areas from near-zero to near-complete coverage.

4.3. Homogeneous Travel-to-work Areas (HTTWA)

As discussed in Section 5.3.1, a robust identification of the effect of broadband over time requires a certain degree of homogeneity between treated and control catchment areas. For this purpose, the study exploits a revised version of the Italian *"Sistemi Locali del Lavoro"* (travel-to-work areas, TTWA). ISTAT divided the Italian territory into 611 TTWAs, based on an algorithm capable of identifying areas where the majority of the labour force lives and works within their boundaries. This methodology, consistent with the criteria developed by the *Eurostat Task Force on the TTWA algorithm*, ensures substantial spatial homogeneity in socioeconomic and institutional characteristics. By assessing the nature of the municipalities belonging to the same TTWA, it is reassuring to discover that 523 areas cover exclusively municipalities belonging to the same urban class 'cities' or 'small towns and rural areas' (ISTAT). However, 88 TTWAs contain municipalities from two different classes, potentially undermining within-area comparability. To address this, I interact the TTWA boundaries with the municipality-level urban density classification map, thereby producing a more granular partition. The final spatial classification yields 699 *homogeneous travel-to-work-areas* (HTTWAs), each designated as either urban or peripheral/rural (See Figure A3). On average, HTTWAs cover 429 km^2 and include 11 municipalities.[22]

4.4. Additional Data Sources

In addition to the data sources used in the main analysis, I also draw on two additional data sources to conduct robustness exercises.

The *Indagine Multiscopo sulle Famiglie* (Multipurpose Household Survey) is a nationally representative survey that collects detailed information on demographics, education, labour

[22] The 95th percentile of the two distributions are, respectively, 1,044 km^2 and 39 municipalities. As a comparison, the city of Rome has a total area of 1,283 km2

market status, family background, health, and social participation. The sampling frame is stratified to ensure representativeness across regions and population groups, and interviews are conducted face-to-face with all household members. Its breadth allows for a descriptive exploration of the relationship between household internet access and children's use of ICT.

The *Autovalutazione della scuola* dataset captures schools' self-assessment of their internal practices as reported in the Self-Evaluation Report (RAV). These evaluations are based on collective reflection involving school leaders and teachers, who contribute by interpreting student performance data, assessing teaching and organisational practices, and identifying priorities for improvement. While leadership teams draft the official report, individual teachers play a key role in providing classroom-level insights, participating in departmental discussions, and implementing the actions defined in the School Improvement Plan.
I construct a teaching quality index as the average of schools' self-assessed performance across seven process areas reported in the RAV: (i) Curriculum, planning, and assessment; (ii) Learning environment; (iii) Inclusion and differentiation; (iv) Continuity and guidance; (v) Strategic orientation and school organisation (vi) Development and enhancement of human resources; (vii) Integration with the local community and relations with families.[23] This index reflects a broad definition of teaching quality, encompassing both pedagogical practices and organisational capacities that may directly or indirectly affect student learning.

[23] In the RAV, the areas are: "Curriculum, planning, and assessment" concerns how the school designs, delivers, and evaluates its educational offer; "Learning environment" refers to the quality of the classroom climate, teaching methods, and resources; "Inclusion and differentiation" captures the strategies adopted to promote equity and support diverse learning needs; "Continuity and guidance" relates to measures that assist students in school transitions and in making future educational or career choices; "Strategic orientation and school organisation" reflects the leadership's capacity to set priorities and manage school processes effectively; "Development and enhancement of human resources" covers policies for teacher training, professional development, and staff management; and "Integration with the local community and relations with families" describes the school's engagement with parents, institutions, and local stakeholders.

5. Empirical Strategy

5.1. Student-level analysis

The baseline model is a student-level specification, based on the stylised model presented in the Annex (Section C).

$$y_{ict} = \beta BA_{ct}^{d} + \gamma X_{ict} + \rho S_{ct} + \theta W_{ct} + \mu_c + n_u + \lambda_{pt} + \varepsilon_{ict} \tag{1}$$

where y_{ict} represents student i's achievement in school c at time t, BA_{ct}^{d} is a dummy variable that equals 1 when the catchment area of school c records a weighted broadband coverage above 75%, X_{ict} is a vector of time-varying 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. Unobservable school quality is accounted for by grade 8 school building fixed effects and grade 5 school fixed effects, respectively μ_c, and n_u, whereas HTTWA-year fixed effects, λ_{pt}, absorb all cross-HTTWA variation within a given year.

To avoid a problematic comparison across observations treated at different points in time, I stack the dataset in 2×2 subsamples satisfying the following criteria:

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

This procedure yields four sub-experiments covering the academic years 2015/16-2016/17, 2016/17-2017/18, 2017/18-2018/19, and 2020/21-2021/22 [24].

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

$$y_{ict} = \beta BA_{ct}^{d} + \gamma X_{ict} + \rho S_{ct} + \theta W_{ct} + \mu_c \times g_d + n_u + \lambda_{pt} \times g_d + \varepsilon_{ict} \tag{2}$$

where HTTWA-year and school-level fixed effects are interacted with the group d dummy, g_d.

[24] 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.

5.2. 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 by 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).

$$y_{ict} = \alpha y_{ist-3} + \beta BA^d_{ct} + \gamma X_{ict} + \rho S_{ct} + \theta W_{ct} + \mu_c \times g_d + n_u + \lambda_{pt} \times g_d + \varepsilon_{ict} \tag{3}$$

where y_{ict} and y_{ict-3} denote, respectively, student i's achievement in school c at time t and at $t-3$. 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 performance at $t-3$, the autoregressive specification allows me to investigate the heterogeneity of the treatment across different social groups. In particular, I test the impact of access to high-speed broadband on student performance, as a function of prior achievement, ethnicity, socio-economic background, and nursery attendance.

5.3. Estimation issues

To correctly identify the effect of the rollout of the broadband infrastructure on student performance, it is critical to address important identification concerns regarding the exogeneity of the rollout and possible sorting dynamics.

5.3.1. Treatment exogeneity

A first concern regards the exogeneity of the treatment variable. While other studies exploit measures of internet usage, this study adopts a supply variable, measured as the average internet speed guaranteed by the broadband technology, weighted by the share of students located in each census area.[25] This choice helps overcome attrition and measurement error

[25] In Section 6.3, I provide evidence on the relationship between internet usage and use of ICT for educational purposes, exploiting a representative survey conducted by Istat.

that generally affect measures of internet usage and, mitigates the concerns about omitted variable bias that typically affect analyses which exploit survey data.

Nevertheless, broadband access measures are themselves not exempt from endogeneity concerns. A first concern is that the rollout might be partly endogenous, driven by local demand, and possibly even ‘capturable’ by local stakeholders. Some studies try to address these issues by absorbing enough variation to account for unobservable confounders. For instance, Sanchis-Guarner et al. (2021) address this problem by adopting a neighbouring discontinuity design, ensuring 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, reducing the external validity of the results. 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.

Other studies try to map the data-generating process behind the rollout and isolate the shock driven by the distance from the local exchange station. For instance, Campante et al. (2018) 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 remained stable 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 is a strong assumption, particularly given the extensive evidence in the regional economics literature of rising regional disparities in 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'*, combined with a solid empirical model and a rich set of robustness checks.

There are good reasons to believe that the NUBP can be characterised as exogenous. Even though some geographical characteristics associated with the local infrastructural endowments may have influenced the implementation costs, the policy aimed to cover 100% of municipalities within 5 years. Since an efficient implementation required a progressive geographical coverage, these local characteristics should not have significantly affected the rollout timing.[26]

These assumptions alone do not fully rule out political bias in the implementation phase. 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 programme was designed and implemented by the national government, through a top-down process, leaving little role for local authorities. Moreover, Mayors in small towns lacked the political power to deliver relevant changes to a national plan, especially when this kind of change would have involved higher costs for the whole project.

The claim of exogenous treatment is further supported by a robust empirical model. All specifications include school fixed effects, which absorb the time-invariant factors, including the ability of local authorities to capture resources from the central government. In addition, they include HTTWA-year fixed effects, that absorb the cross-HTTWA variation in the speed of the rollout. As a result, the model focuses on the evolution of broadband that takes place within the 699 homogeneous TTWAs, as characterised in Section 4.3.

[26]In Figure A2 in the appendix I show the relation between treatment in t and distance from the closest treatment municipality in $t-1$.

Still, the timing of the rollout could be partially correlated with some key geographical factors, such as the distance from a main urban centre or the density recorded in the area. This is a realistic possibility, since the broadband was rolled out sequentially starting from 35 nodes located in city centres around Italy. Even within HTTWAs, municipalities located closer to the main urban centre might be more likely to get access to the fibre, and at the same time they could record slightly different student performance trends. If school building and HTTWA-year fixed effects cannot fully absorb the confounding factors, our results would be biased. To further address these endogeneity concerns, I run a series of robustness checks.

In Section 7, I address the issue of non-random broadband access by introducing a novel metric for expected access to the infrastructure. This measure is developed using the approach proposed by Borusyak and Hull (2023), which involves constructing a hypothetical broadband network modelled on the characteristics of the actual one. The rationale behind this is that the sequential deployment of broadband depends on the proximity to the 35 central nodes (see Figure A4) and on additional factors such as terrain and institutional effectiveness. Each year, I generate a hypothetical broadband network by varying the decay parameter at each node while keeping constant the geographic distances between nodes and census areas and the overall network size (i.e., the number of census areas covered that year).[27] The expected broadband proxy is intended to capture a census area's potential exposure to high-speed broadband services. As shown in Figure A5, the constructed network closely mirrors actual broadband availability. Including this hypothetical broadband variable in the baseline model allows me to isolate the influence of actual broadband availability, and thereby obtain an unbiased estimate of its effect on student outcomes. To further address concerns related to the timing of the rollout, I directly assess whether the results are driven by spatial confounders. Specifically, I divide the sample into three density-based groups (see Figure A6)[28] and re-estimate the model separately for each one. In both exer-

[27]More details about the procedure are provided in Appendix B.

[28]The first group includes HTTWAs with less than 93 people per square kilometres, the second group includes areas with a density between 92 and 202, the third group includes large urban areas, with a density between 203 and 6,701.

cises, I retain the full specification, with both HTTWA-year and school building fixed effects.

Moreover, I assess whether the results are driven by unobservable characteristics of the municipalities that influence the timing of the rollout, by re-estimating the model on *'white areas'* only, thereby excluding all larger municipalities, that are the only ones potentially capable of influencing the central government. As reported in Section 7, all these exercises broadly confirm the baseline results.

5.3.2. Sorting dynamics

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.[29] Recently, several concerns have been raised regarding the opportunity of using value added models to assess 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 focused mainly on the possible bias arising from student sorting dynamics and the reliability of standardised tests scores as a proxy for student achievement. These issues do not appear to be relevant in the framework of this study. First, student sorting across schools appears to be a negligible phenomenon in Italy. Between 2009 and 2011, Italy recorded the second lowest internal mobility rate among OECD countries (OECD, 2013). In 2015, only 2% of the residents moved to a different region, and most of them moved from the South to the North. Short-distance migration patterns appear to be a quite negligible phenomenon in the Italian context.

Furthermore, while unobservable characteristics might theoretically influence test scores, they are unlikely to exhibit any correlation with the deployment of broadband infrastructure. Nevertheless, in an effort to comprehensively address these concerns, in Section 7 I show

[29] In cases where student performance demonstrates a mean-reverting pattern at the tails of the previous performance distribution, there is a risk that the Value-Added Model may fail to adequately capture the dynamics of the learning process. As a result, in Section 7, I also estimate a quadratic specification to address this issue.

that the main results remain consistent when I restrict the sample to students who attended grade 8 in the same municipality as in grade 5.

5.3.3. Heterogeneity in Teacher Quality

A further potential source of bias in the estimation strategy is the role of teachers in shaping the effects of the policy. In principle, teachers could influence outcomes through two channels.

First, teachers could leverage the internet to improve the quality of their teaching materials. In that case, part of the estimated impact on student performance could reflect higher teacher productivity rather than students' direct use of ICT. This hypothesis appears less relevant in the Italian context. As explained in Section 3.2, in the Italian school system—particularly in primary and lower secondary schools—teachers and schools have very limited autonomy in setting curricula and teaching methods. The Ministry of Education sets course content, approves a restricted set of textbooks, and requires teachers to report periodically on compliance with national standards. Teaching methods and materials must also be consistent with each school's educational plan, which, in turn, must align with national objectives. As a result, teachers have limited autonomy in the adoption of ICT for instructional purposes. Moreover, any heterogeneity across schools in the educational plan is absorbed in the baseline specification through school fixed effects.

Second, one might worry about sorting of teachers into treated areas. If higher-quality teachers were able to move to schools that benefit from broadband rollout, the results could be biased. Again, this concern is unlikely to be relevant in Italy. Human resource management is conducted primarily at the national level: teachers apply to province-level lists, are ranked, and then assigned to schools based on vacancies and preferences, with limited discretion for school principals. The baseline model also absorbs time-varying shocks at the HTTWA level. The remaining within-HTTWA variation is unlikely to reflect sorting dynamics, since teachers cannot apply directly to specific schools, but only to provinces through the national recruitment system.

Nevertheless, it would be desirable to directly control for teacher quality and classroom practices. Because the INVALSI data lack this information, I partially address the concern in Appendix A.3.2 by drawing on school-level indicators of teaching quality from each school's Self-Evaluation Report (RAV) (see Section 4.4). Controlling for these proxies does not alter the main findings, suggesting that heterogeneity in teaching quality is unlikely to drive the results. However, the available measures do not allow me to fully isolate teacher-specific effects, and therefore the estimates should be interpreted as reflecting the overall impact of broadband availability within the broader school-student environment.

6. Results

6.1. Baseline analysis

I start the investigation by examining whether the broadband infrastructural rollout affected average student performance.

In Table 2, I report the regression estimates for Equation (2), progressively adding a rich set of school, municipality, and student-level covariates, as well as grade 5 and grade 8 school fixed-effects and HTTWA×year fixed effects. All specifications are estimated using a stacked design.

Table 2: Student-level analysis - Baseline

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband dummy	0.00311	0.0222***	-0.0153	-0.00827	-0.0160	-0.00894	-0.0162	-0.00921
	(0.00964)	(0.00801)	(0.0108)	(0.00918)	(0.0108)	(0.00921)	(0.0109)	(0.00924)
Observations	937,261	937,314	936,526	936,578	936,526	936,578	936,526	936,578
R-squared	0.053	0.031	0.179	0.184	0.180	0.185	0.180	0.185
Year×HTTWA FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 5 School FE	-	-	✓	✓	✓	✓	✓	✓
School variables	-	-	✓	✓	✓	✓	✓	✓
Peer effects	-	-	-	-	✓	✓	✓	✓
Municipality variables	-	-	-	-	-	-	✓	✓

Notes: This table reports regression results of the model in Equation (2). The dependent variables are students' standardised numeracy and literacy scores in grade 8. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors are clustered at the school level and reported in parentheses. ***, **, and * indicate significance at the 0.01, 0.05, and 0.1 levels, respectively.

Columns (1) and (2) illustrate, respectively, the results of standardised maths and literacy tests when no covariates are considered in the specification. High-speed internet broadband appears to have a small but significant effect on student performance in literacy. Access to broadband increases the average student literacy scores by 2.2% of a standard deviation. No significant effect is found for maths. In columns (3) and (4), after including school-level variables, the coefficients for both test scores become statistically insignificant 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) does not significantly affect the point estimates. Overall, the estimates confirm a null average effect of the policy on student performance.

In Table 3, I extend the specification by including student performance in grade 5 (see Equation (3)). The new specification can be interpreted as a value-added model of cognitive achievement, which makes it possible to investigate the effects of the staggered rollout of the broadband infrastructure on students' learning trajectories.

Table 3: Value added model

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband dummy	0.00136	0.0104	0.00188	0.00896	0.00156	0.00856	0.000596	0.00802
	(0.0103)	(0.00856)	(0.0120)	(0.0101)	(0.0120)	(0.0101)	(0.0120)	(0.0101)
Test score (Grade 5)	0.590***	0.592***	0.646***	0.606***	0.646***	0.606***	0.646***	0.606***
	(0.00407)	(0.00284)	(0.00313)	(0.00217)	(0.00313)	(0.00217)	(0.00313)	(0.00217)
Observations	936,331	910,413	935,598	909,672	935,598	909,672	935,598	909,672
R-squared	0.379	0.380	0.484	0.477	0.484	0.478	0.484	0.478
Year×HTTWA FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 5 School FE	-	-	✓	✓	✓	✓	✓	✓
School variables	-	-	✓	✓	✓	✓	✓	✓
Peer effects	-	-	-	-	✓	✓	✓	✓
Municipality variables	-	-	-	-	-	-	✓	✓

Notes: This table reports regression results of the value-added model in Equation (3). The dependent variables are students' standardised numeracy and literacy scores in grade 8. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors are clustered at the school level and reported in parentheses. ***, **, and * indicate significance at the 0.01, 0.05, and 0.1 levels, respectively.

In columns (1) and (2), the policy is found to have no effect on student performance in both subjects. When I include school-level covariates in columns (3) and (4), both coefficients remain insignificant. The inclusion of class composition and municipality-level variables in columns (5)–(8) further reduces the magnitude of the coefficient for literacy scores, but largely confirms previous results. Overall, the value-added model confirms a perfect zero effect of the policy on both maths and literacy scores.

6.2. Heterogeneity analysis

Table 4 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 socioeconomic background. In columns (1) and (2), I interact the main regressor with a standardised measure of student performance in grade 5. The results suggest that high-speed broadband fosters a mean-reverting pattern. While the effect is null for students with an average performance in grade 5, it turns negative for high-performers.

In columns (3) and (4), I further investigate this relationship, interacting the policy variable with a set of dummies identifying the quartiles of the previous test score distribution. Students in the lowest quartile of the performance distribution in grade 5 record a 2% standard deviation increase in maths test scores, although the coefficient is only statistically significant at the 10% level. Similarly, they record a significant 3.6% increase in literacy scores. For both subjects, the effect of the policy progressively declines moving toward the higher quartiles and becomes negative for students above the median. High performers in grade 5 who obtain access to high-speed internet record a 5.5% standard deviation decline in maths scores and a 4.1% decline in literacy. This result aligns with the hypothesis discussed in Section 2. While the interplay between positive effects (such as access to ed-tech resources) and negative ones (such as increase in leisure time and excessive screen time) produces an average null effect, low-performing students are more likely to get a net benefit from the new learning tools.

In columns (5) and (6), I focus on the second relevant source of heterogeneity, namely the family background (ESCS index). In this case, I find a positive and significant contribution of the standardised ESCS index on the policy outcome. Students with an average family background are not affected by the policy, while those with one standard deviation ESCS score above the mean record a 2.8% standard deviation increase in maths scores and a 3% gain in literacy. The pattern is further investigated in columns (7) and (8), where I focus on the ESCS quartiles. The broadband infrastructure reduces student performance in the first quartile of the family background distribution by 4.7% standard deviations in maths and by 4.1% standard deviations in reading.

Table 4: Interactions (1)

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband coverage	0.00268	0.0107			-0.00155	0.00560		
	(0.0120)	(0.0101)			(0.0120)	(0.0102)		
Test score (Grade 5)	0.652***	0.612***			0.646***	0.606***	0.649***	0.610***
	(0.00315)	(0.00225)			(0.00313)	(0.00217)	(0.00312)	(0.00216)
Broadband × Test score (Grade 5)	-0.0392***	-0.0368***						
	(0.00558)	(0.00429)						
Test score (Grade 5) - II			0.534***	0.557***				
			(0.00338)	(0.00319)				
Test score (Grade 5) - III			0.986***	0.969***				
			(0.00468)	(0.00350)				
Test score (Grade 5) - IV			1.589***	1.466***				
			(0.00669)	(0.00462)				
Broadband coverage × Test score (Grade 5) - I			0.0207*	0.0360***				
			(0.0124)	(0.0106)				
Broadband coverage × Test score (Grade 5) - II			0.0206*	0.0192*				
			(0.0120)	(0.0103)				
Broadband coverage × Test score (Grade 5) - III			-0.00152	-0.00460				
			(0.0121)	(0.0103)				
Broadband coverage × Test score (Grade 5) - IV			-0.0555***	-0.0415***				
			(0.0143)	(0.0116)				
Family Background	0.112***	0.118***	0.126***	0.132***	0.108***	0.113***		
	(0.00135)	(0.00134)	(0.00133)	(0.00134)	(0.00146)	(0.00144)		
Broadband coverage × Family Background					0.0280***	0.0297***		
					(0.00312)	(0.00305)		
Family Background - II							0.104***	0.112***
							(0.00312)	(0.00304)
Family Background - III							0.154***	0.162***
							(0.00339)	(0.00323)
Family Background - IV							0.258***	0.271***
							(0.00378)	(0.00376)
Broadband coverage × Family Background - I							-0.0468***	-0.0410***
							(0.0137)	(0.0118)
Broadband coverage × Family Background - II							-0.00375	-0.000343
							(0.0127)	(0.0109)
Broadband coverage × Family Background - III							0.0126	0.0245**
							(0.0126)	(0.0107)
Broadband coverage × Family Background - IV							0.0310**	0.0391***
							(0.0125)	(0.0107)
Observations	935,616	909,694	935,616	909,694	935,616	909,694	935,598	909,672
R-squared	0.477	0.473	0.449	0.445	0.477	0.473	0.482	0.476
Year×HTTWA FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓	✓	✓	✓	✓
School variables	✓	✓	✓	✓	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table reports regression results of the value-added model in Equation (3). The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics, and is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 8 school fixed effects, and HTTWA×year fixed effects. Columns (3)–(4) and (7)–(8) report interactions of the variable of interest with the quartiles of grade 8 performance and the ESCS index, respectively. Robust standard errors are clustered at the school level and reported in parentheses. ***, **, and * indicate significance at the 0.01, 0.05, and 0.1 levels, respectively.

The negative effect progressively declines in magnitude for the upper quartiles and turns positive for students with a family background above the median.
Students in the top quartile who obtain access to high-speed internet record a 3.1% standard deviation increase in maths scores and a 3.9% increase in literacy.
This result is noteworthy given that the analysis focuses on areas without broadband access in 2015, thus excluding students from the most advantaged family backgrounds. Consequently, our estimates may underestimate the overall gap with the most advantaged students, who had already obtained access prior to the policy.
The findings are once again consistent with the hypothesis presented in Section 2. The higher degree of parental supervision generally guaranteed by advantaged families might help students make the best of the available technology and minimise the harmful consequences of a misuse of the new digital tools. Overall, the results highlight an important nexus between the effectiveness of the policy and students' prior performance and socioeconomic background. Low performers whose parents are sufficiently educated might benefit the most from the introduction of new information technologies.

In Table 5, I further examine the heterogeneity of the results by considering the performance of the students and the family background jointly. To this end, I interact the main explanatory variable with dummies that classify students into four groups based on their grade 5 test scores and ESCS quartiles.
Columns (1) and (2) report the baseline model. Among disadvantaged students who performed poorly in grade 5, the rollout of the infrastructure has no measurable effect on either maths or literacy. By contrast, high-performers with a poor family background experience a 6.6% standard deviation decrease in numeracy scores and a 6.1% decrease in literacy.
The only group that benefits from the policy is low performers from advantaged families, who gain 2.5% of a standard deviation in maths and 4.8% in literacy. For advantaged high performers, the coefficients are positive, but not statistically significant.

Table 5: Interactions (2)

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
(Fam. Background < avg) × (Test score (Grade 5) > avg)	0.963***	0.925***	0.950***	0.902***
	(0.00521)	(0.00396)	(0.00516)	(0.00391)
(Fam. Background > avg) × (Test score (Grade 5) < avg)	0.181***	0.199***	0.170***	0.188***
	(0.00325)	(0.00322)	(0.00324)	(0.00320)
(Fam. Background > avg) × (Test score (Grade 5) > avg)	1.200***	1.154***	1.178***	1.122***
	(0.00473)	(0.00359)	(0.00471)	(0.00353)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) < avg)	-0.000892	-0.00526	-0.00164	-0.00829
	(0.0118)	(0.0104)	(0.0118)	(0.0103)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) > avg)	-0.0664***	-0.0611***	-0.0657***	-0.0591***
	(0.0136)	(0.0115)	(0.0136)	(0.0114)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) < avg)	0.0255**	0.0484***	0.0248**	0.0462***
	(0.0117)	(0.0102)	(0.0117)	(0.0101)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) > avg)	0.0110	0.00753	0.0101	0.00863
	(0.0120)	(0.0102)	(0.0120)	(0.0101)
Observations	935,598	909,672	935,598	909,672
R-squared	0.366	0.351	0.370	0.367
Year×HTTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	-	-	✓	✓
Peer effects	-	-	✓	✓
Municipality variables	-	-	✓	✓

Notes: This table reports regression results of the value-added model in Equation (3). The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics, and is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors are clustered at the school level and reported in parentheses. ***, **, and * indicate significance at the 0.01, 0.05, and 0.1 levels, respectively.

In columns (3) and (4), I add individual, school, and municipality-level controls. The results are almost identical, with just a slight reduction in the magnitude of all coefficients. Disadvantaged high performers continue to show declines of around 6.6% in maths and 5.9% in literacy, while advantaged low performers still display gains of about 2.5% in maths and 4.6% in literacy.

Overall, these findings indicate that both prior achievement and family background shape the effect of broadband access. Advantaged low performers benefit the most, while disadvantaged high performers are negatively affected. In this sense, broadband expansion both narrows gaps between low- and high-performing students and widens inequalities linked to family background.

6.3. Channels

As discussed in Section 5, the identification strategy exploits the exogenous assignment of treatment, thereby avoiding the endogeneity concerns that typically affect studies based on self-reported internet use. This approach, while strengthening the internal validity of the results, necessarily rests on assumptions regarding actual usage patterns. In particular, it requires that access to high-speed internet leads to effective use, and that students differ systematically in how they employ this technology. Given these limitations, it is informative to complement the baseline analysis with descriptive evidence on the relationship between broadband access and students' frequency and patterns of internet use.

For this purpose, I draw on the *Indagine Multiscopo sulle Famiglie* (Multipurpose Household Survey), as reported in Appendix A.2. Panel A of Table A1 examines the relationship between household access to high-speed broadband and the likelihood of daily internet use. Across the full sample of students aged 6 to 18, access is positively and significantly associated with daily use, even after controlling for extensive household characteristics. When interacted with parental education, the effect is strongest among students whose parents completed secondary schooling, but remains significant across all groups. Restricting the sample to students aged 11–15 (mainly lower secondary) produces qualitatively similar patterns.

Panels B and C focus on specific uses of ICT. Panel B shows that broadband access is strongly associated with online gaming. The effect is visible across all parental education levels and somewhat larger among students with parents educated to secondary level, though no clear gradient emerges. Results for lower-secondary students are comparable.

Panel C turns to the use of ICT for educational purposes. Here, the relationship with high-speed access is positive and significant for the full sample, but the magnitude increases monotonically with parental education. For students whose parents did not progress beyond primary schooling, the association is not statistically significant.[30] Restricting the analysis

[30] A Wald test rejects the null hypothesis that the coefficients of the interaction terms 'High-speed

to lower-secondary students yields the same pattern: the likelihood of educational use rises with parental education.[31]

These findings are descriptive and should not be given a causal interpretation. Nonetheless, they provide valuable insights. First, access to high-speed broadband is associated with more frequent internet use across the distribution of family background, with somewhat stronger associations among more advantaged households. This supports the assumption of a clear link between broadband availability and changes in internet use. Second, the type of use varies systematically: while high-speed access is positively related to both gaming and educational applications, the latter association strengthens with parental education and is not detectable among students from less-advantaged households. These results are consistent with the hypotheses outlined in Section 2, namely that ICT can enhance learning when complemented by supportive household resources, but may also crowd out time for study through gaming and leisure activities.

7. Robustness Checks

7.1. Alternative Specifications

Thus far, I have expressed grade 8 performance as a linear function of grade 5 performance. To account for mean-reverting patterns at the lower tail of the test score distribution, Table A2 adds a quadratic term. The new model does not lead to significant changes relative to the basic VAM (Table 3), causing only a negligible reduction in the magnitude of the coefficients of interest.

The baseline analysis relied on a stacked specification in which the treatment indicator equals one if a catchment area records a weighted broadband coverage above 75%. This threshold

internet×Parent with tertiary educ.' and 'High-speed internet×Parent with secondary educ.' are equal (Chi-square=38.7; p-value=0.00

[31] Again, the Wald test rejects the null hypothesis that the coefficients of the interaction terms 'High-speed internet×Parent with tertiary educ.' and 'High-speed internet×Parent with secondary educ.' are equal (Chi-square=17.69; p-value=0.00).

may appear arbitrary and conceal some meaningful information. To address this concern, in Appendix A.3.1 I re-estimate each specification with a simple high-dimensional fixed-effects regression framework, where the explanatory variable is the actual share of households covered in each area.[32] Table A3 reports results for the basic value-added model. In contrast to Table 3, I find some evidence of a negative effect of broadband on numeracy, though the coefficient is only marginally significant once peer effects and municipality-level controls are included. Literacy results remain consistently null.

Table A4 examines heterogeneity by prior performance and family background. Once again, the estimates are broadly in line with those presented in Table 4. High-speed broadband tends to have a negative effect on high-performers in grade 5 and disadvantaged students. I find a clear negative effect of broadband connection on students in the two top quartiles of the distribution of previous performance and in the two bottom quartiles of family background. However, the positive effects are less significant for the other groups. Finally, in Table A5 I replicate the empirical model estimated in Table 5. Advantaged high-performers record a negative and significant effect, while low-performers with a good family background seem to benefit from the policy. While in the baseline model, I do not find significant effects for the other two groups, in this case, I find a significant negative effect on disadvantaged low-performers and, at least in maths, on advantaged high-performers. Overall, while average effects are somewhat more negative, the pattern of heterogeneity is robust.

This paper focuses on the effect of the progressive rollout of internet broadband (fibre-to-the-node or fibre-to-the-home), which guarantees at least 30 Mbit/s internet speed. This minimum threshold was at the core of the NUBP, which aimed to guarantee this speed over the whole Italian territory by 2020. However, the rapid increase in access to 30 Mbit/s internet was accompanied by a substantial increase in fibre-to-the-home connections, which guarantee a minimum of 100 Mbit/s internet speed. Table A6 augments the specification

[32] Also, by exploiting the original dataset instead of the stacked one, I am able to significantly increase the sample size.

with the share of households covered at 100 Mbit/s. In this case, the coefficients reported for the variable 'broadband coverage' (30 mbps) can be interpreted as the effect of a 100% coverage of 30 Mbit/s broadband internet in an area with no access to 100 Mbit/s broadband. In columns (1) and (2) I re-estimate the specification presented in columns (7) and (8) of Table A3. On average, I do not find significant differences between the two technologies. In columns (3) and (4) I re-estimate the specification presented in columns (3) and (4) of Table A5. A higher share of buildings with access to 100 Mbit/s internet speed is associated with worse results for disadvantaged low-performers and better results for disadvantaged high-performers. No additional effect is found for advantaged students. These findings suggest that access to 100 Mbit/s internet speed might have a more homogeneous effect among disadvantaged students. However, the effects on the different groups appear to be broadly confirmed.

7.2. Alternative Samples

As documented in Section 5, the identification strategy relies on the assumption that, within homogeneous travel-to-work areas, the broadband rollout can be considered exogenous with respect to the variable of interest. As a further robustness check, in Table A7 I re-estimate the main specification focusing only on 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 rollout during the period was driven only the public intervention (see Section 3.1). In columns (1) and (2) I re-estimate the specification presented in columns (7) and (8) of Table 3. When I focus only on *'market failure areas'*, I record a small, although non-significant, positive 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 5. Once again, the results are broadly consistent with the baseline pattern: advantaged low-performers benefit the most, while disadvantaged high-performers are negatively affected. However, most coefficients report a higher magnitude. Overall, the table shows that if we focus on 'white areas', the main patterns become even more pronounced.

As discussed in Section 5.3.2, 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 a longer commute time 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 (OECD, 2013). Few people move, and when this happens, it is typically to move to the richer North, rather than to a neighbouring town. Nevertheless, in Table A8 I re-estimate the main specifications excluding students, who for different reasons, ended up attending a school in grade 8 located in a different municipality than the one where they attended grade 5. This strategy goes beyond the intended scope of the robustness check, as it excludes students whose closest primary and lower secondary schools were located in different municipalities. However, results are generally confirmed. In columns (1) and (2) I re-estimate the specification presented in columns (7) and (8) of Table 3. By excluding movers, I still find a null effect on maths, but the positive effect on literacy grows in magnitude and becomes significant. In columns (3) and (4), I re-estimate the specification presented in columns (3) and (4) of Table 5. Once again, the results are very similar to the ones recorded in the baseline model, with only a notable reduction in the significance of the positive effect of high-speed internet on maths scores for advantaged low-performers.

In Section 4, I discuss in detail the construction of the dataset. The estimating sample involves all students who, at the beginning of the period, had no access to high-speed internet but reported having an internet connection at home. The decision to exclude students with no connection to the internet was meant to focus the analysis on the connection to the broadband and exclude other channels, such as the incentive to invest in a personal computer. In Table A10, I replicate the baseline models including in the sample students who did not have internet at home or for whom the information was missing. The results are very similar to those in Tables 3 and 5. The magnitude of the average effect is somewhat larger for numeracy scores, but remains insignificant. The positive effect on advantaged low performers and the negative effect on disadvantaged high performers are confirmed, with slightly larger coefficients. A similar exercise is conducted in Table A11, where I include all observations dropped due to missing values in some of the key covariates (which are excluded from the model). Once again, the estimates are almost identical to the ones recorded in the baseline analysis.

Finally, in Table A12, I re-estimate the model excluding the post-COVID years. By dropping these years, I lose only 10% of the observations, due to the high share of treated units that characterise the last years of the study period. As before, I find a null average effect of high-speed internet on both numeracy and literacy scores, and I record the same heterogeneous effect across prior student performance and family background.

7.3. Exogeneity of the Rollout

The correct identification of the effect of high-speed broadband on student performance relies on the assumption that the rollout of the infrastructure was exogenous. While this might appear to be as a strong assumption per se, it is far more credible within the proposed model, which includes school fixed effects, which absorb local time-invariant factors, and HTTWA-year fixed effects, which absorb the cross-HTTWA variation in the speed of the rollout. The fact that the results are broadly confirmed when focusing on *'market failure'* areas only (Table A7) further supports this assumption. In order to provide further reassurance, in this section I provide four additional robustness exercises which test the robustness of the results with respect to possible local confounders.

In the first exercise, as discussed in Section 5.3.1, I account for the non-random exposure to the broadband access by introducing a new measure of expected access to the infrastructure, inspired by Borusyak and Hull (2023). Table A13 reports the correlation between the broadband dummy and the expected broadband dummy constructed following the approach discussed in Section B. Not surprisingly, the two variables appear to be strongly associated when I include only region-year fixed effects (see Figure A5). The coefficient decreases in magnitude when I include province-year fixed effects and becomes non-significant when I include HTTWA-year fixed effects. This result already provides some reassurance about the ability of the model to absorb the non-random component of the rollout. However, in Table A14, I also estimate the flexible model used in Tables A3 and A5, including the expected broadband variable. According to Borusyak and Hull (2023), the inclusion of this term in the model helps isolate the impact of actual broadband availability, allowing me to estimate an unbiased effect on student performance. Comparing the estimates in columns (1) and (2) with those presented in columns (7) and (8) of Table A3, I find that the average effect of the policy is not affected by the inclusion of the term in the model. Similarly, the heterogeneous results presented in columns (3) and (4) of Table A5 are robust to the inclusion of the expected broadband term, with even a slight increase in the magnitude of all coefficients.

The second exercise consists of testing the heterogeneous effect of the policy with respect to the urban density recorded in the HTTWA. In Table A15, I divide the sample into three subsamples (0-92, 93-202, and 203-6,701 people per km^2) based on the density distribution of the sample (see Figure A6). The results can be compared with those presented in Tables A3, A4. The average effect of high-speed internet on numeracy and literacy scores is negative and significant for the third group and insignificant for the others. In columns (3)–(4), (7)–(8) and (11)–(12) I find a negative and significant effect on disadvantaged high performers in all three groups (with the exception of the effect on literacy in the first group). The effect on advantaged low performers varies more in magnitude and significance across groups and subjects, but is always positive and does not seem to have a linear relationship with density. Overall, despite some variation in significance and magnitude, which can be explained by the reduction in sample size, the baseline results are broadly confirmed across all three density groups.

The third exercise tests whether the estimates vary with school-level teaching quality. In Section 5.3.3, I discuss how heterogeneity in teaching quality could bias the results and explain why this concern is unlikely to be severe in the Italian context. To support this argument, Table A9 divides the sample into two subsamples based on data from each school's Self-Evaluation Report (RAV) (see Section 4.4). Schools are grouped according to whether their score on the teaching quality index is above or below the median. The estimates reported in columns (1)–(2) and (5)–(6) can be compared to those in columns (7) and (8) of Table A3. For both high- and low-quality schools, the average effect of high-speed internet is statistically insignificant. In columns (3)–(4) and (7)–(8), I replicate the specifications from columns (3) and (4) of Table A4. The negative and significant effect on disadvantaged high performers is confirmed, with a somewhat larger magnitude in high-quality schools. The positive effect on advantaged low performers is also confirmed, but the magnitude is notably larger in low-quality schools. Overall, the baseline results are broadly consistent between the two groups, helping to mitigate concerns that variation in teaching quality drives the estimated policy effects. The fact that students in high-quality schools appear to benefit less from broadband access may seem counter-intuitive, but it is consistent with the interpretation that digital education tools are relatively more valuable in lower-quality learning environments.

Finally, I consider the possibility that municipalities controlled by the same party as the national government could receive a preferential treatment. This is unlikely, for 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 that all largest cities already had access to the broadband at the beginning of the period, I 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 a centre-left or a centre-right coalition[33] and assess whether being controlled by the same party as national governments was correlated with access to the broadband infrastructure.

[33] I exclude all municipalities that were run by independent mayors.

Results are reported in Table A16, where I separately estimate the relationship between the local governing party affiliation and both the actual broadband access and the 'excess access', obtained by subtracting the pseudo-fibre term from the broadband variable. Any correlation between political party affiliation and this term would suggest that the component of the policy that is not explained by the proximity to the node is in fact influenced by political pressures. In column (1), where I run a simple regression, I find that municipalities run by the same political party as the government are 16% 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 result declines to just 11%. 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[34]. In this case, the magnitude of the coefficient shrinks by two thirds and becomes non-significant. Finally, in columns (5) and (6) I adopt province-year and HTTWA-year 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. In the second panel, I replicate the same analysis, this time focusing on 'excess broadband', namely the difference between the recorded broadband and the expected broadband. A significant relationship would suggest that the variation of the broadband not explained by the distance from the nodes is in part driven by political influences. However, the estimates are never significant, with the only exception of column (5). These findings are consistent with Fig. A2, which illustrates 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. These results rule out any systematic link between actual broadband, 'excess broadband', and local political affiliation.

[34]Italy has 22 regions, corresponding to the Eurostat NUTS-2 administrative level.

Overall, these exercises provide consistent evidence that the broadband rollout can be treated as exogenous within the empirical framework. The tests rule out systematic relationships between broadband access and local political affiliation, teaching quality, or urban density. Moreover, the introduction of the expected broadband proxy shows that the model effectively absorbs the non-random component of the rollout. Taken together, these findings strengthen confidence in the identification strategy and the validity of the estimated effects.

8. Conclusions

This study assesses the influence of access to high-speed broadband internet on educational attainment. This investigation exploits a large infrastructural programme implemented by the Italian government. The available dataset facilitated an examination of the heterogeneous effects of the policy with respect to students' performance in previous grades and their family background. The overall findings indicate that the presence of broadband internet access has a negligible impact on average educational achievements. However, this overarching result masks a significant heterogeneity with respect to parental socioeconomic background and prior performance.

Notably, low-performing students in the fifth grade who come from more advantaged backgrounds show positive gains from improved access to educational resources. This pattern is consistent with the hypothesis that a culturally enriched family environment enables these students to benefit disproportionately, thereby helping to narrow the achievement gap. In contrast, a detrimental effect is observed for students from less privileged family backgrounds. This might be due to the fact that, without parental supervision, the online-gaming effect and the harmful consequences of excessive screen time offset any possible positive effect on learning productivity. Alternatively, this could stem from financial constraints, with economically disadvantaged families unable to provide their children with the necessary hardware to capitalise on the newly available infrastructure.

The empirical strategy employed in this study maximises the internal validity of the results but does not permit a precise identification of the mechanisms at work. Descriptive evidence from the ISTAT Multipurpose Household Survey, however, provides useful guidance. While all children—regardless of family background—increase their internet use once high-speed internet becomes available, the rise in the use of educational tools is concentrated among more advantaged students. This pattern suggests that parental supervision and motivation, rather than economic barriers to broadband access, are likely to play a central role in explaining the heterogeneous effects. By contrast, there is no evidence that disadvantaged students allocate more time to online gaming (if anything, the opposite). In this case, the concern appears to lie less with the type of use than with the intensity of use.

Taken together, these results suggest that access to high-speed internet may have widened the performance gap between advantaged and disadvantaged students, while at the same time reducing the performance gap between students with similar socioeconomic background. The magnitude of these effects is likely to be underestimated, since students in large urban centres—who typically come from more advantaged families and already had broadband access—are excluded from the analysis, which may have attenuated the estimated impact.

Although more work is required to better understand the underlying mechanisms, these results carry important policy implications. Broadband expansion by itself is not sufficient to level the playing field. Rather, its benefits accrue disproportionately to those students with access to the complementary resources—both material and cultural—necessary to translate digital opportunities into improved learning outcomes. Effective ICT upgrading programmes must therefore be accompanied by policies that ensure disadvantaged families can access appropriate hardware and that teachers and parents are supported in guiding students' use of digital technologies. Without such complementary measures, investments in broadband may unintentionally reinforce existing educational inequalities.

Future research should pursue two complementary avenues. On the one hand, future studies should investigate whether the impact of high-speed internet on student performance varies depending on the digital infrastructure of schools (e.g., ICT technologies) and teachers' ability to harness the new technologies to improve teaching outcomes. On the other hand, by exploiting longer time-periods, it should be possible to study how student performance varies over time in the years following the introduction of the new technology.

References

Adachi, P.J., Willoughby, T., 2013. More than just fun and games: The longitudinal relationships between strategic video games, self-reported problem solving skills, and academic grades. Journal of youth and adolescence 42, 1041–1052.

A'Hearn, B., Venables, A.J., 2013. Regional disparities: internal geography and external trade, in: The Oxford handbook of the Italian economy since unification. Oxford University Press Oxford, UK.

Ahlfeldt, G., Koutroumpis, P., Valletti, T., 2017. Speed 2.0: Evaluating access to universal digital highways. Journal of the European Economic Association 15, 586–625.

Akerman, A., Gaarder, I., Mogstad, M., 2015. The skill complementarity of broadband internet. The Quarterly Journal of Economics 130, 1781–1824.

Angrist, J., Lavy, V., 2002. New evidence on classroom computers and pupil learning. The Economic Journal 112, 735–765.

Banerjee, A.V., Cole, S., Duflo, E., Linden, L., 2007. Remedying education: Evidence from two randomized experiments in india. The quarterly journal of economics 122, 1235–1264.

Barr, R., Zack, E., Garcia, A., Muentener, P., 2008. Infants' attention and responsiveness to television increases with prior exposure and parental interaction. Infancy 13, 30–56.

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

Barrow, L., Markman, L., Rouse, C.E., 2009. Technology's edge: The educational benefits of computer-aided instruction. American Economic Journal: Economic Policy 1, 52–74.

Bellou, A., 2015. The impact of internet diffusion on marriage rates: evidence from the broadband market. Journal of Population Economics 28, 265–297.

Belo, R., Ferreira, P., Telang, R., 2014. Broadband in school: Impact on student performance. Management Science 60, 265–282.

Beltran, D.O., Das, K.K., Fairlie, R.W., 2006. Do home computers improve educational outcomes? Evidence from matched current population surveys and the national longitudinal survey of youth 1997. Technical Report. IZA Discussion Papers.

Borman, G.D., Benson, J.G., Overman, L., 2009. A randomized field trial of the fast forword language computer-based training program. Educational Evaluation and Policy Analysis 31, 82–106.

Borusyak, K., Hull, P., 2023. Nonrandom exposure to exogenous shocks. Econometrica 91, 2155–2185.

Brotman, S.N., 2016. The real digital divide in educational technology. Brookings TechTank .

Bulman, G., Fairlie, R.W., 2016. Technology and education: Computers, software, and the internet, in: Handbook of the Economics of Education. Elsevier. volume 5, pp. 239–280.

Cain, N., Gradisar, M., 2010. Electronic media use and sleep in school-aged children and adolescents: A review. Sleep medicine 11, 735–742.

Cambini, C., Sabatino, L., Zaccagni, S., 2021. The Faster the Better? The Effect of Ultra-Fast Broadband on Students' Performance. Technical Report. CEBI Working Paper Series.

Camerini, A.L., Schulz, P.J., Jeannet, A.M., 2018. The social inequalities of internet access, its use, and the impact on children's academic performance: Evidence from a longitudinal study in switzerland. New Media & Society 20, 2489–2508.

Campante, F., Durante, R., Sobbrio, F., 2018. Politics 2.0: The multifaceted effect of broadband internet on political participation. Journal of the European Economic Association 16, 1094–1136.

Checchi, D., Rettore, E., Girardi, S., 2019. Ic technology and learning: an impact evaluation of cl@ ssi 2.0. Education Economics 27, 241–264.

Chetty, R., Friedman, J.N., Rockoff, J.E., 2014. Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. American economic review 104, 2593–2632.

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

Corazzini, L., Meschi, E., Pavese, C., 2021. Impact of early childcare on immigrant children?s educational performance. Economics of Education Review 85, 102181.

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

David Acevedo-Polakovich, I., Lorch, E.P., Milich, R., 2007. Comparing television use and reading in children with adhd and non-referred children across two age groups. Media Psychology 9, 447–472.

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, 12–23.

Duvivier, C., Bussière, C., 2022. The contingent nature of broadband as an engine for business startups in rural areas. Journal of Regional Science 62, 1329–1357.

Escueta, M., Nickow, A.J., Oreopoulos, P., Quan, V., 2020. Upgrading education with technology: Insights from experimental research. Journal of Economic Literature 58, 897–996.

Fairlie, R.W., 2016. Do boys and girls use computers differently, and does it contribute to why boys do worse in school than girls? The BE Journal of Economic Analysis & Policy 16, 59–96.

Fairlie, R.W., Beltran, D.O., Das, K.K., 2010. Home computers and educational outcomes: Evidence from the nlsy97 and cps. Economic inquiry 48, 771–792.

Fairlie, R.W., Robinson, J., 2013. Experimental evidence on the effects of home computers on academic achievement among schoolchildren. American Economic Journal: Applied Economics 5, 211–240.

Falck, O., Gold, R., Heblich, S., 2014. E-lections: Voting behavior and the internet. American Economic

Review 104, 2238–65.

Falck, O., Mang, C., Woessmann, L., 2018. Virtually no effect? different uses of classroom computers and their effect on student achievement. Oxford Bulletin of Economics and Statistics 80, 1–38.

Feng, J., Spence, I., 2018. Playing action video games boosts visual attention. Video game influences on aggression, cognition, and attention , 93–104.

Fiorini, M., 2010. The effect of home computer use on children's cognitive and non-cognitive skills. Economics of Education review 29, 55–72.

Fisch, S.M., Kirkorian, H., Anderson, D., 2005. Transfer of learning in informal education. Transfer of learning from a modern multidisciplinary perspective , 371–393.

Goolsbee, A., Guryan, J., 2006. The impact of internet subsidies in public schools. The Review of Economics and Statistics 88, 336–347.

Hale, L., Guan, S., 2015. Screen time and sleep among school-aged children and adolescents: a systematic literature review. Sleep medicine reviews 21, 50–58.

Hanushek, E.A., Rivkin, S.G., 2010. Generalizations about using value-added measures of teacher quality. American economic review 100, 267–71.

Ichino, A., Tabellini, G., 2014. Freeing the italian school system. Labour Economics 30, 113–128.

ISTAT, 2021. Gli spostamenti per motivi di studio o lavoro secondo il censimento permanente della popolazione. Censimenti Permanenti: Popolazione e Abitanti URL: https://www.istat.it/it/files/ /2021/05/Censimento_spostamenti_pendolari.pdf.

Jackson, K., Makarin, A., 2018. Can online off-the-shelf lessons improve student outcomes? evidence from a field experiment. American Economic Journal: Economic Policy 10, 226–254.

Kane, T.J., Staiger, D.O., 2008. Are teacher-level value-added estimates biased? an experimental validation of non-experimental estimates, in: National Conference on Value-Added Modeling, Madison, WI, Citeseer. pp. 106–118.

Karam, R., Pane, J.F., Griffin, B.A., Robyn, A., Phillips, A., Daugherty, L., 2017. Examining the implementation of technology-based blended algebra i curriculum at scale. Educational Technology Research and Development 65, 399–425.

Kirschner, P.A., Karpinski, A.C., 2010. Facebook® and academic performance. Computers in human behavior 26, 1237–1245.

Kostyrka-Allchorne, K., Cooper, N.R., Simpson, A., 2017. The relationship between television exposure and children's cognition and behaviour: A systematic review. Developmental review 44, 19–58.

Leuven, E., Lindahl, M., Oosterbeek, H., Webbink, D., 2007. The effect of extra funding for disadvantaged pupils on achievement. The Review of Economics and Statistics 89, 721–736.

Linebarger, D.L., Barr, R., Lapierre, M.A., Piotrowski, J.T., 2014. Associations between parenting, media use, cumulative risk, and children's executive functioning. Journal of Developmental & Behavioral Pediatrics 35, 367–377.

Machin, S., McNally, S., Silva, O., 2007. New technology in schools: Is there a payoff? The Economic Journal 117, 1145–1167.

Malamud, O., Cueto, S., Cristia, J., Beuermann, D.W., 2019. Do children benefit from internet access? experimental evidence from peru. Journal of Development Economics 138, 41–56.

Malamud, O., Pop-Eleches, C., 2011. Home computer use and the development of human capital. The Quarterly journal of economics 126, 987–1027.

McCoy, D., Lyons, S., Morgenroth, E., Palcic, D., Allen, L., 2018. The impact of broadband and other infrastructure on the location of new business establishments. Journal of Regional Science 58, 509–534.

Muralidharan, K., Singh, A., Ganimian, A.J., 2019. Disrupting education? experimental evidence on technology-aided instruction in india. American Economic Review 109, 1426–60.

Mutz, D.C., Roberts, D.F., Vuuren, D.v., 1993. Reconsidering the displacement hypothesis: Television's influence on children's time use. Communication Research 20, 51–75.

Nusche, D., Minea-Pic, A., 2020. Ict resources in school education: What do we know from oecd work.

OECD, 2013. Population mobility among regions .

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

Przybylski, A.K., Wang, J.C., 2016. A large scale test of the gaming-enhancement hypothesis. PeerJ 4, e2710.

Roschelle, J., Feng, M., Murphy, R.F., Mason, C.A., 2016. Online mathematics homework increases student achievement. AERA open 2, 2332858416673968.

Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., Gallagher, L.P., 2010. Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. American Educational Research Journal 47, 833–878.

Rouse, C.E., Krueger, A.B., 2004. Putting computerized instruction to the test: A randomized evaluation of a "scientifically based"' reading program. Economics of Education Review 23, 323–338.

Sanchis-Guarner, R., Montalbán, J., Weinhardt, F., 2021. Home broadband and human capital formation. Available at SSRN 3772087 .

Schmidt, M.E., Anderson, D.R., 2009. The impact of television on cognitive development and educational achievement, in: Children and television. Routledge, pp. 87–106.

Sharif, I., Wills, T.A., Sargent, J.D., 2010. Effect of visual media use on school performance: a prospective study. Journal of Adolescent Health 46, 52–61.

Tondeur, J., Sinnaeve, I., Van Houtte, M., Van Braak, J., 2011. Ict as cultural capital: The relationship between socioeconomic status and the computer-use profile of young people. New media & society 13, 151–168.

Van Deursen, A.J., Van Dijk, J.A., 2019. The first-level digital divide shifts from inequalities in physical access to inequalities in material access. New media & society 21, 354–375.

Vedechkina, M., Borgonovi, F., 2021. A review of evidence on the role of digital technology in shaping attention and cognitive control in children. Frontiers in Psychology 12, 611155.

Ventura, M., Shute, V., Zhao, W., 2013. The relationship between video game use and a performance-based measure of persistence. Computers & Education 60, 52–58.

Vigdor, J.L., Ladd, H.F., Martinez, E., 2014. Scaling the digital divide: Home computer technology and student achievement. Economic Inquiry 52, 1103–1119.

Wolf, D., Irwin, N., 2024. Is it really bridging the gap? fiber internet's impact on housing values and homebuyer demographics. Journal of Regional Science 64, 238–271.

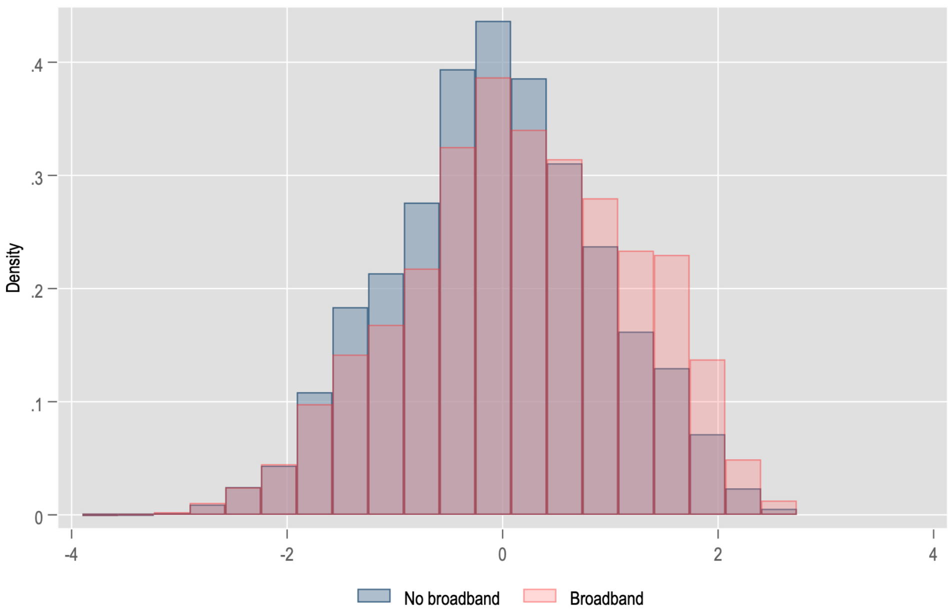
Appendices

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

A. Tables and Figures

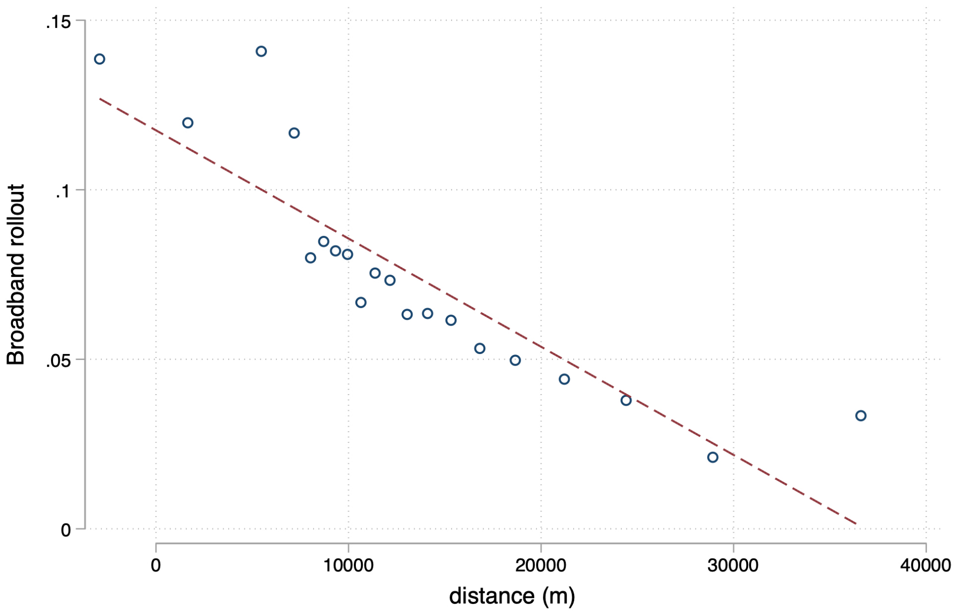
A.1. Figures

Figure A1: Distribution of student ESCS by fibre access, 2015



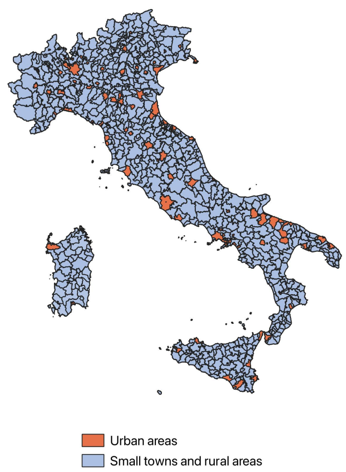
Notes: The figure plots the distribution of student ESCS (Economic, Social and Cultural Status) in 2015 separately for those living in areas already covered by fibre and those in areas without coverage. Students in fibre-covered areas are disproportionately wealthier, reflecting the concentration of early broadband rollout in city centres (see Figure 2). The sample used in the main analysis focuses on students in areas that gained access only as a consequence of the policy-driven expansion of fibre coverage.

Figure A2: Broadband rollout and distance from the infrastructure in t-1



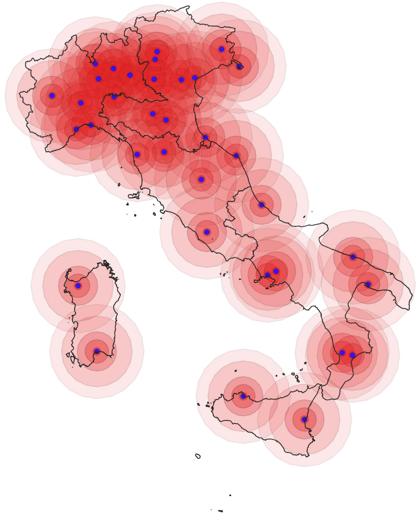
Notes: The figure illustrates the relation between the distance from the closest treated municipality in $t-1$ and the likelihood for a municipality not-yet-treated to get access to the broadband infrastructure in t. Both variables have been residualized with respect to year fixed-effects, in order to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

Figure A3: Homogeneous Travel-to-Work Areas



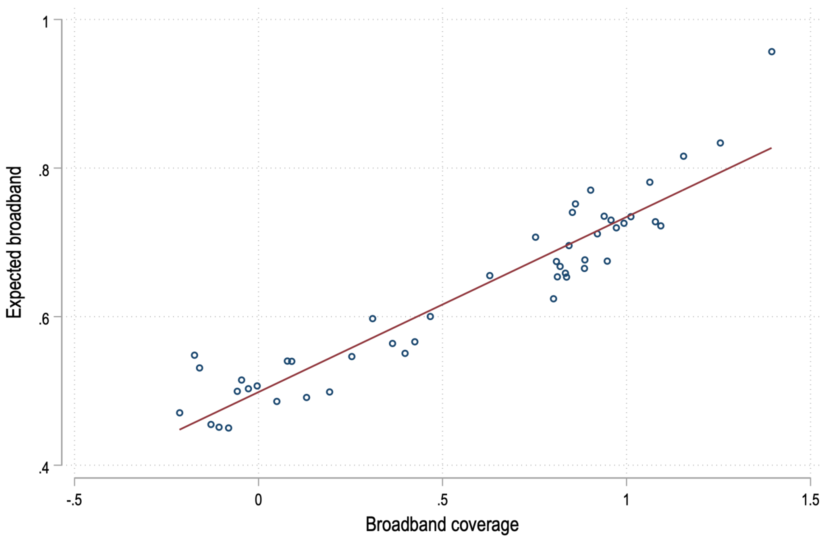
Notes: The figure shows the division of the Italian territory into *homogeneous travel-to-work areas* (HTTWAs). Orange areas denote urban HTTWAs, while blue areas denote small-town and rural HTTWAs. All specifications in this study include HTTWA×year fixed effects, thereby absorbing any cross-HTTWA variation.

Figure A4: Nodes



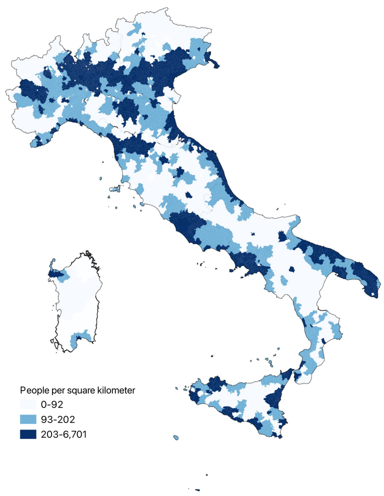
Notes: The figure illustrates the position of the nodes and provides a graphical representation of the proximity of each area to the closest node.

Figure A5: Actual vs Expected Broadband



Notes: The figure plots the relationship between actual and expected broadband availability, net of region×year fixed effects.

Figure A6: HTTWAs by density



Notes: The figure reports the population density by homogeneous travel-to-work area.

A.2. ICT use

Table A1: Broadband access and internet use

	Primary/Secondary School					Lower Secondary School		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A				Daily internet use				
High-speed internet	0.192***	0.128***	0.133***		0.169***	0.110***	0.106***	
	(0.00561)	(0.00739)	(0.00762)		(0.00940)	(0.0117)	(0.0121)	
High-speed internet # Parent with tertiary educ.				0.116***				0.118***
				(0.00895)				(0.0142)
High-speed internet # Parent with secondary educ.				0.153***				0.114***
				(0.00813)				(0.0129)
High-speed internet # Parent with primary educ. or less				0.105***				0.0735***
				(0.0105)				(0.0164)
Observations	26,102	26,098	25,562	25,562	6,250	6,250	6,135	6,135
R-squared	0.142	0.160	0.161	0.161	0.115	0.144	0.143	0.142
Panel B				Online gaming				
High-speed internet	0.192***	0.128***	0.133***		0.169***	0.110***	0.106***	
	(0.00561)	(0.00739)	(0.00762)		(0.00940)	(0.0117)	(0.0121)	
High-speed internet # Parent with tertiary educ.				0.116***				0.118***
				(0.00895)				(0.0142)
High-speed internet # Parent with secondary educ.				0.153***				0.114***
				(0.00813)				(0.0129)
High-speed internet # Parent with primary educ. or less				0.105***				0.0735***
				(0.0105)				(0.0164)
Observations	26,102	26,098	25,562	25,562	6,250	6,250	6,135	6,135
R-squared	0.142	0.160	0.161	0.161	0.115	0.144	0.143	0.142
Panel C				Educational tools				
High-speed internet	0.0703***	0.0677***	0.0574***		0.0563***	0.0542***	0.0430***	
	(0.00706)	(0.00892)	(0.00917)		(0.0109)	(0.0138)	(0.0142)	
High-speed internet # Parent with tertiary educ.				0.101***				0.0858***
				(0.0107)				(0.0169)
High-speed internet # Parent with secondary educ.				0.0528***				0.0342**
				(0.00976)				(0.0151)
High-speed internet # Parent with primary educ. or less				0.00828				0.0170
				(0.0126)				(0.0190)
Observations	26,102	26,098	25,562	25,562	6,250	6,250	6,135	6,135
R-squared	0.143	0.145	0.153	0.154	0.177	0.182	0.192	0.192
Individual characteristics	✓	✓	✓	✓	✓	✓	✓	✓
Domestic ICT	-	✓	✓	✓	✓	-	✓	✓
Family Characteristics	-	-	✓	✓	✓	-	-	✓
Region×Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table reports regression results of the empirical approach discussed in Section 6.3. In *Panel A*, the dependent variable is a dummy equal to one if the student reports daily internet use. In *Panel B* and *Panel C*, the dependent variable indicates, respectively, whether the student engaged in online gaming or used digital educational tools in the previous three months. The key explanatory variable is a dummy equal to one if the household has access to high-speed internet. Covariates include gender, citizenship, commuting time to school, and chronic diseases (individual characteristics); access to a PC, 3G mobile connection, and USB modem (domestic ICT); and parents' occupation and education, household type, and household size (family characteristics). All regressions include region×year fixed effects. Robust standard errors, clustered at the household level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

A.3. Robustness exercises

A.3.1. Alternative Models

Table A2: Quadratic value added model

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband dummy	0.00121	0.0100	0.00158	0.00863	0.00125	0.00823	0.000308	0.00769
	(0.0102)	(0.00844)	(0.0119)	(0.0100)	(0.0119)	(0.0100)	(0.0120)	(0.0101)
Test score (Grade 5)	0.611***	0.601***	0.657***	0.610***	0.656***	0.610***	0.656***	0.610***
	(0.00347)	(0.00254)	(0.00288)	(0.00207)	(0.00288)	(0.00207)	(0.00288)	(0.00207)
(Test score - Grade 5)^2	-0.0448***	-0.0372***	-0.0260***	-0.0214***	-0.0261***	-0.0214***	-0.0261***	-0.0214***
	(0.00183)	(0.00161)	(0.00136)	(0.00131)	(0.00136)	(0.00131)	(0.00136)	(0.00131)
Observations	936,331	910,413	935,598	909,672	935,598	909,672	935,598	909,672
R-squared	0.383	0.382	0.485	0.478	0.485	0.478	0.485	0.478
Year×HTTWA FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 5 School FE	-	-	✓	✓	✓	✓	✓	✓
School variables	-	-	✓	✓	✓	✓	✓	✓
Peer effects	-	-	-	-	✓	✓	✓	✓
Municipality variables	-	-	-	-	-	-	✓	✓

Notes: This table presents regression results from the quadratic specification of the value-added model in Eq. 3. The dependent variables are students' standardised numeracy and literacy scores in grade 8. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A3: Broadband coverage - value added model

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband coverage	0.0145	0.0341***	-0.0290**	-0.00922	-0.0294*	-0.00951	-0.0264*	-0.0102
	(0.00899)	(0.00706)	(0.0120)	(0.0106)	(0.018)	(0.0106)	(0.0181)	(0.0106)
Test score (Grade 5)	0.576***	0.595***	0.589***	0.576***	0.589***	0.576***	0.589***	0.576***
	(0.00344)	(0.00236)	(0.00306)	(0.00213)	(0.00307)	(0.00214)	(0.00307)	(0.00214)
Observations	1,202,410	1,168,948	1,201,673	1,168,183	1,201,673	1,168,183	1,201,673	1,168,183
R-squared	0.387	0.394	0.453	0.456	0.453	0.456	0.453	0.456
Year×HTTWA FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 5 School FE	-	-	✓	✓	✓	✓	✓	✓
School variables	-	-	✓	✓	✓	✓	✓	✓
Peer effects	-	-	-	-	✓	✓	✓	✓
Municipality variables	-	-	-	-	-	-	✓	✓

Notes: This table presents regression results of the value-added model described in Section 5. The right- and left-hand side variables are the same as in Eq. 3, but I do not apply the stacking procedure outlined in Section 5.1. Instead, I estimate a high-dimensional fixed effects regression, where the variable of interest is the actual weighted average of broadband coverage at the catchment-area level (see Section 4.2). The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A4: Broadband coverage - interactions (1)

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband coverage	-0.0245*	-0.00840			-0.0290**	-0.0134		
	(0.0181)	(0.0106)			(0.0121)	(0.0106)		
Test score (Grade 5)	0.610***	0.588***			0.589***	0.575***	0.593***	0.580***
	(0.00379)	(0.00275)			(0.00307)	(0.00213)	(0.00306)	(0.00213)
Broadband × Test score (Grade 5)	-0.0353***	-0.0212***						
	(0.00426)	(0.00349)						
Test score (Grade 5) - II			0.509***	0.546***				
			(0.00397)	(0.00365)				
Test score (Grade 5) - III			0.946***	0.954***				
			(0.00554)	(0.00409)				
Test score (Grade 5) - IV			1.514***	1.428***				
			(0.00804)	(0.00556)				
Broadband coverage × Test score (Grade 5) - I			0.0101	0.00947				
			(0.0123)	(0.0112)				
Broadband coverage × Test score (Grade 5) - II			0.0170	-0.0114				
			(0.0120)	(0.0107)				
Broadband coverage × Test score (Grade 5) - III			-0.0500***	-0.0220**				
			(0.0123)	(0.0108)				
Broadband coverage × Test score (Grade 5) - IV			-0.0854***	-0.0188				
			(0.0134)	(0.0115)				
Family Background	0.127***	0.133***	0.140***	0.145***	0.110***	0.113***		
	(0.00121)	(0.00120)	(0.00115)	(0.00115)	(0.00167)	(0.00162)		
Broadband coverage × Family Background					0.0287***	0.0341***		
					(0.00233)	(0.00219)		
Family Background - II							0.106***	0.112***
							(0.00361)	(0.00356)
Family Background - III							0.158***	0.160***
							(0.00390)	(0.00367)
Family Background - IV							0.263***	0.270***
							(0.00429)	(0.00421)
Broadband coverage × Family Background - I							-0.0623***	-0.0558***
							(0.0129)	(0.0115)
Broadband coverage × Family Background - II							-0.0413***	-0.0281***
							(0.0124)	(0.0109)
Broadband coverage × Family Background - III							-0.0163	0.00532
							(0.0123)	(0.0109)
Broadband coverage × Family Background - IV							0.00396	0.0253**
							(0.0125)	(0.0109)
Observations	1,201,673	1,168,183	1,201,673	1,168,183	1,201,673	1,168,183	1,201,673	1,168,183
R-squared	0.453	0.456	0.424	0.430	0.453	0.456	0.451	0.454
Year×HTTWA FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓	✓	✓	✓	✓
School variables	✓	✓	✓	✓	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table presents regression results of the value-added model in Eq. 3. The specification is the same as in Table 4, except that I do not apply the stacking procedure described in Section 5.1. Instead, I estimate a high-dimensional fixed effects regression, where the variable of interest is the actual weighted average of broadband coverage at the catchment-area level (see Section 4.2). The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates is reported in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Columns (3)–(4) and (7)–(8) report interactions of the variable of interest with the quartile distribution of, respectively, grade 5 student performance and the ESCS index. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A5: Broadband coverage - interactions (2)

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
(Fam. Background < avg) × (Test score (Grade 5) > avg)	0.934***	0.915***	0.922***	0.892***
	(0.00595)	(0.00455)	(0.00592)	(0.00450)
(Fam. Background > avg) × (Test score (Grade 5) < avg)	0.173***	0.184***	0.160***	0.172***
	(0.00366)	(0.00360)	(0.00365)	(0.00357)
(Fam. Background > avg) × (Test score (Grade 5) > avg)	1.163***	1.138***	1.141***	1.106***
	(0.00554)	(0.00416)	(0.00552)	(0.00409)
Broadband coverage × (Fam. Background < avg) × (Test score (Grade 5) < avg)	-0.0232**	-0.0372***	-0.0224*	-0.0403***
	(0.0117)	(0.0108)	(0.0117)	(0.0107)
Broadband coverage × (Fam. Background < avg) × (Test score (Grade 5) > avg)	-0.103***	-0.0529***	-0.100***	-0.0557***
	(0.0129)	(0.0113)	(0.0129)	(0.0113)
Broadband coverage × (Fam. Background > avg) × (Test score (Grade 5) < avg)	0.0282**	0.0332***	0.0299**	0.0308***
	(0.0118)	(0.0107)	(0.0118)	(0.0107)
Broadband coverage × (Fam. Background > avg) × (Test score (Grade 5) > avg)	-0.0271**	0.00998	-0.0261**	0.00646
	(0.0121)	(0.0107)	(0.0121)	(0.0106)
Observations	1,201,673	1,168,183	1,201,673	1,168,183
R-squared	0.342	0.337	0.346	0.351
Year×HTTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	-	-	✓	✓
Peer effects	-	-	✓	✓
Municipality variables	-	-	✓	✓

Notes: This table presents regression results of the value-added model in Eq. 3. The specification is the same as in Table 5, except that I do not apply the stacking procedure described in Section 5.1. Instead, I estimate a high-dimensional fixed effects regression, where the variable of interest is the actual weighted average of broadband coverage at the catchment-area level (see Section 4.2). The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates is reported in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A6: Broadband coverage - 30mbps vs 100mbps

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband coverage (30mbps)	-0.0266**	-0.0101		
	(0.0122)	(0.0107)		
Broadband coverage (100mbps)	0.00199	-0.000324		
	(0.0125)	(0.0107)		
Fam. Background	0.113***	0.118***		
	(0.00144)	(0.00144)		
Test score (Grade 5)	0.589***	0.576***		
	(0.00307)	(0.00214)		
(Fam. Background < avg) × (Test score (Grade 5) > avg)			0.923***	0.892***
			(0.00591)	(0.00450)
(Fam. Background > avg) × (Test score (Grade 5) < avg)			0.160***	0.172***
			(0.00365)	(0.00357)
(Fam. Background > avg) × (Test score (Grade 5) > avg)			1.141***	1.106***
			(0.00552)	(0.00409)
Broadband coverage (30mbps) × (Fam. Background < avg) × (Test score (Grade 5) < avg)			-0.0157	-0.0352***
			(0.0119)	(0.0109)
Broadband coverage (30mbps) × (Fam. Background < avg) × (Test score (Grade 5) > avg)			-0.113***	-0.0628***
			(0.0132)	(0.0115)
Broadband coverage (30mbps) × (Fam. Background > avg) × (Test score (Grade 5) < avg)			0.0310***	0.0349***
			(0.0120)	(0.0109)
Broadband coverage (30mbps) × (Fam. Background > avg) × (Test score (Grade 5) > avg)			-0.0266**	0.00358
			(0.0123)	(0.0107)
Broadband coverage (100mbps) × (Fam. Background < avg) × (Test score (Grade 5) < avg)			-0.0266**	-0.0228*
			(0.0130)	(0.0124)
Broadband coverage (100mbps) × (Fam. Background < avg) × (Test score (Grade 5) > avg)			0.0576***	0.0295**
			(0.0157)	(0.0131)
Broadband coverage (100mbps) × (Fam. Background > avg) × (Test score (Grade 5) < avg)			-0.00166	-0.0176
			(0.0135)	(0.0126)
Broadband coverage (100mbps) × (Fam. Background > avg) × (Test score (Grade 5) > avg)			0.00489	0.0105
			(0.0137)	(0.0112)
Observations	1,201,673	1,168,183	1,201,673	1,168,183
R-squared	0.453	0.456	0.346	0.351
Year×HTTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓

Notes: This table reports estimates from the same model as in Table A3 (columns (7)–(8)) and Table A5 (columns (3)–(4)), augmented with a term capturing the share of buildings with access to 100 Mbit/s. The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

A.3.2. Alternative Samples

Table A7: White areas only

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	0.0191	0.0203*		
	(0.0135)	(0.0122)		
Test score (Grade 5)	0.647***	0.609***		
	(0.00352)	(0.00252)		
Fam. Background	0.112***	0.119***		
	(0.00156)	(0.00156)		
(Fam. Background < avg) × (Test score (Grade 5) > avg)			0.959***	0.906***
			(0.00576)	(0.00446)
(Fam. Background > avg) × (Test score (Grade 5) < avg)			0.172***	0.188***
			(0.00372)	(0.00366)
(Fam. Background > avg) × (Test score (Grade 5) > avg)			1.184***	1.126***
			(0.00522)	(0.00410)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) < avg)			0.0300**	0.00614
			(0.0135)	(0.0123)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) > avg)			-0.0764***	-0.0564***
			(0.0154)	(0.0137)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) < avg)			0.0531***	0.0632***
			(0.0137)	(0.0124)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) > avg)			0.0155	0.0213*
			(0.0141)	(0.0124)
Observations	676,760	658,979	676,760	658,979
R-squared	0.491	0.483	0.376	0.371
Year×TTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓

Notes: This table reports regression results from the value-added model in Eq. 3, estimated on the subsample of *white areas* (see Section 3.1). The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A8: Stayers only

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	0.000212	0.0245**		
	(0.0136)	(0.0112)		
Test score (Grade 5)	0.667***	0.620***		
	(0.00344)	(0.00240)		
Fam. Background	0.110***	0.115***		
	(0.00154)	(0.00154)		
(Fam. Background < avg) × (Test score (Grade 5) > avg)			0.981***	0.918***
			(0.00569)	(0.00455)
(Fam. Background > avg) × (Test score (Grade 5) < avg)			0.177***	0.191***
			(0.00383)	(0.00375)
(Fam. Background > avg) × (Test score (Grade 5) > avg)			1.205***	1.135***
			(0.00524)	(0.00413)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) < avg)			-2.92e-05	0.00785
			(0.0129)	(0.0113)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) > avg)			-0.0676***	-0.0375***
			(0.0151)	(0.0127)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) < avg)			0.0250*	0.0649***
			(0.0129)	(0.0112)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) > avg)			0.00735	0.0247**
			(0.0134)	(0.0112)
Observations	661,734	644,606	661,734	644,606
R-squared	0.493	0.483	0.371	0.366
Year×HTTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓

Notes: This table reports regression results from the value-added model in Eq. 3, estimated excluding students who attended grade 8 in a different municipality from the one where they attended grade 5. The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A9: Breakdown by school quality

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband Coverage	-0.00816	-0.00964			-0.0256	-0.00482		
	(0.0204)	(0.0165)			(0.0158)	(0.0145)		
Fam. Background	0.126***	0.134***			0.126***	0.132***		
	(0.00196)	(0.00196)			(0.00157)	(0.00154)		
(Test score (Grade 5) > avg	0.587***	0.574***			0.598***	0.582***		
	(0.00505)	(0.00365)			(0.00392)	(0.00258)		
(Fam. Background < avg) # (Test score (Grade 5) > avg)			0.920***	0.884***			0.931***	0.902***
			(0.00897)	(0.00702)			(0.00793)	(0.00589)
(Fam. Background > avg) # (Test score (Grade 5) < avg)			0.160***	0.171***			0.161***	0.175***
			(0.00563)	(0.00564)			(0.00484)	(0.00470)
(Fam. Background > avg) # (Test score (Grade 5) > avg)			1.150***	1.105***			1.142***	1.111***
			(0.00823)	(0.00626)			(0.00742)	(0.00546)
Broadband # (Fam. Background < avg) # (Test score (Grade 5) < avg)			0.0145	-0.0329*			-0.0359**	-0.0437***
			(0.0192)	(0.0169)			(0.0158)	(0.0146)
Broadband # (Fam. Background < avg) # (Test score (Grade 5) > avg)			-0.0815***	-0.0452***			-0.104***	-0.0631***
			(0.0207)	(0.0172)			(0.0172)	(0.0154)
Broadband # (Fam. Background > avg) # (Test score (Grade 5) < avg)			0.0594***	0.0405**			0.0202*	0.0238*
			(0.0196)	(0.0172)			(0.0127)	(0.0144)
Broadband # (Fam. Background > avg) # (Test score (Grade 5) > avg)			-0.0158	0.0104			-0.0252	0.00296
			(0.0198)	(0.0166)			(0.0162)	(0.0145)
Observations	481,166	467,369	481,166	467,369	695,000	675,773	695,000	675,773
R-squared	0.461	0.462	0.356	0.358	0.461	0.464	0.351	0.358
Year#TTWA FE	✓	✓	✓	✓	✓	✓	✓	✓
V grade School FE	✓	✓	✓	✓	✓	✓	✓	✓
VIII Grade School FE	✓	✓	✓	✓	✓	✓	✓	✓
School variables	✓	✓	✓	✓	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓	✓	✓	✓	✓
School quality	Low	Low	Low	Low	High	High	High	High

Notes: This table reports regression results from the value-added model in Eq. 3, estimated separately for schools with a teacher quality index above or below the median (see Section 4.4 for details on variable construction). The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A10: Broadband - including students with no internet in t-3

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	0.00253	0.00500		
	(0.0117)	(0.0102)		
Fam. Background	0.113***	0.118***		
	(0.00130)	(0.00126)		
Test score (Grade 5)	0.641***	0.603***		
	(0.00312)	(0.00214)		
(Fam. Background < avg) × (Test score (Grade 5) > avg)			0.946***	0.905***
			(0.00504)	(0.00372)
(Fam. Background > avg) × (Test score (Grade 5) < avg)			0.171***	0.191***
			(0.00303)	(0.00299)
(Fam. Background > avg) × (Test score (Grade 5) > avg)			1.178***	1.126***
			(0.00455)	(0.00338)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) < avg)			0.000518	-0.00804
			(0.0113)	(0.0101)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) > avg)			-0.0670***	-0.0639***
			(0.0131)	(0.0113)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) < avg)			0.0315***	0.0474***
			(0.0113)	(0.00993)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) > avg)			0.0149	0.0120
			(0.0116)	(0.0101)
Observations	1,091,294	1,074,827	1,091,294	1,074,827
R-squared	0.483	0.479	0.371	0.369
Year×HTTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓

Notes: This table reports regression results from the value-added model in Eq. 3, estimated on a sample that includes students without any internet connection three years before the baseline period. The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates included in the model is provided in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A11: Broadband - including observations with missing values

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	0.00467	0.00818		
	(0.0117)	(0.00998)		
Fam. Background	0.113***	0.118***		
	(0.00130)	(0.00126)		
Test score (Grade 5)	0.673***	0.644***		
	(0.00297)	(0.00205)		
(Fam. Background $<$ avg) $\times$ (Test score (Grade 5) $>$ avg)			0.961***	0.926***
			(0.00503)	(0.00386)
(Fam. Background $>$ avg) $\times$ (Test score (Grade 5) $<$ avg)			0.183***	0.200***
			(0.00310)	(0.00312)
(Fam. Background $>$ avg) $\times$ (Test score (Grade 5) $>$ avg)			1.201***	1.157***
			(0.00455)	(0.00348)
Broadband $\times$ (Fam. Background $<$ avg) $\times$ (Test score (Grade 5) $<$ avg)			-0.00381	-0.0110
			(0.0113)	(0.00991)
Broadband $\times$ (Fam. Background $<$ avg) $\times$ (Test score (Grade 5) $>$ avg)			-0.0644***	-0.0644***
			(0.0131)	(0.0110)
Broadband $\times$ (Fam. Background $>$ avg) $\times$ (Test score (Grade 5) $<$ avg)			0.0273**	0.0470***
			(0.0112)	(0.00983)
Broadband $\times$ (Fam. Background $>$ avg) $\times$ (Test score (Grade 5) $>$ avg)			0.00932	0.00505
			(0.0115)	(0.00976)
Observations	993,586	972,606	992,478	971,237
R-squared	0.474	0.457	0.367	0.351
Year×HTTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓

Notes: This table reports regression results from the value-added model in Eq. 3, estimated on a sample including observations with missing values. The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A12: Broadband - excluding pandemic years

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	-0.000455	0.00779		
	(0.0121)	(0.0102)		
Fam. Background	0.113***	0.118***		
	(0.00144)	(0.00144)		
Test score (Grade 5)	0.650***	0.607***		
	(0.00345)	(0.00238)		
(Fam. Background $<$ avg) $\times$ (Test score (Grade 5) $>$ avg)			0.953***	0.897***
			(0.00555)	(0.00424)
(Fam. Background $>$ avg) $\times$ (Test score (Grade 5) $<$ avg)			0.166***	0.184***
			(0.00356)	(0.00348)
(Fam. Background $>$ avg) $\times$ (Test score (Grade 5) $>$ avg)			1.180***	1.115***
			(0.00512)	(0.00382)
Broadband $\times$ (Fam. Background $<$ avg) $\times$ (Test score (Grade 5) $<$ avg)			-0.00242	-0.0119
			(0.0119)	(0.0104)
Broadband $\times$ (Fam. Background $<$ avg) $\times$ (Test score (Grade 5) $>$ avg)			-0.0690***	-0.0582***
			(0.0136)	(0.0115)
Broadband $\times$ (Fam. Background $>$ avg) $\times$ (Test score (Grade 5) $<$ avg)			0.0282**	0.0454***
			(0.0118)	(0.0102)
Broadband $\times$ (Fam. Background $>$ avg) $\times$ (Test score (Grade 5) $>$ avg)			0.00823	0.0110
			(0.0121)	(0.0102)
Observations	839,885	816,272	839,885	816,272
R-squared	0.480	0.475	0.369	0.366
Year×HTTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓

Notes: This table reports regression results from the value-added model in Eq. 3, estimated excluding the pandemic years. The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions, based on observed family characteristics; it is constructed using the OECD ESCS index. A complete list of covariates included in the model is provided in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

A.3.3. Exogeneity of the rollout

Table A13: Pseudo fibre

VARIABLES	(1) Expected broadband	(2) Expected broadband	(3) Expected broadband
Actual broadband	0.236***	0.108***	0.00786
	(0.00914)	(0.00676)	(0.00498)
Observations	31,151	31,147	31,026
R-squared	0.287	0.654	0.871
Region-Year FE	✓	-	-
Province-Year FE	-	✓	-
HTTWA-Year FE	-	-	✓

Notes: The table reports the correlation between the actual share of buildings with broadband access and the pseudo-fibre proxy constructed using the method described in Section B. The analysis is conducted at the catchment-area level. Robust standard errors, clustered at the catchment-area level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A14: Broadband coverage - actual connection vs pseudo fibre

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.
Broadband	-0.0261**	-0.0102		
	(0.0121)	(0.0106)		
Test score (Grade 5)	0.589***	0.576***		
	(0.00307)	(0.00214)		
Fam. Background	0.127***	0.133***		
	(0.00121)	(0.00120)		
Expected broadband	0.0182	0.000793		
	(0.0122)	(0.0101)		
(Fam. Background < avg) × (Test score (Grade 5) > avg)			0.887***	0.876***
			(0.00667)	(0.00512)
(Fam. Background > avg) × (Test score (Grade 5) < avg)			0.159***	0.172***
			(0.00408)	(0.00400)
(Fam. Background > avg) × (Test score (Grade 5) > avg)			1.112***	1.099***
			(0.00632)	(0.00456)
Expected broadband × (Fam. Background < avg) × (Test score (Grade 5) < avg)			-0.0276**	-0.0134
			(0.0117)	(0.0102)
Expected broadband × (Fam. Background < avg) × (Test score (Grade 5) > avg)			0.0736***	0.0318***
			(0.0127)	(0.0108)
Expected broadband × (Fam. Background > avg) × (Test score (Grade 5) < avg)			-0.0276**	-0.0140
			(0.0119)	(0.0102)
Expected broadband × (Fam. Background > avg) × (Test score (Grade 5) > avg)			0.0540***	0.00687
			(0.0122)	(0.0101)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) < avg)			-0.00385	-0.0338***
			(0.0118)	(0.0108)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) > avg)			-0.124***	-0.0681***
			(0.0129)	(0.0113)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) < avg)			0.0482***	0.0374***
			(0.0120)	(0.0108)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) > avg)			-0.0417***	0.00443
			(0.0123)	(0.0107)
Observations	1,201,673	1,168,183	1,201,673	1,168,183
R-squared	0.453	0.456	0.346	0.351
Year×HTTWA FE	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓
School variables	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓

Notes: This table presents estimates from the value-added model in Eq. 3. The specification is identical to that in Table A5, but includes the predicted broadband term described in Section B. The dependent variables are students' standardised numeracy and literacy scores in grade 8. Family background captures students' economic, social, and cultural conditions based on observed family characteristics, and is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year fixed effects. Robust standard errors, clustered at the school level, are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Table A15: Breakdown by density

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.	(9)	(10) Lit.	(11) Num.	(12) Lit.
Broadband Coverage	-0.0340	0.0312			-0.00878	0.00663			-0.0362**	-0.0276*		
	(0.0290)	(0.0277)			(0.0216)	(0.0187)			(0.0167)	(0.0145)		
Fam. Background	0.127***	0.132***			0.129***	0.135***			0.125***	0.132***		
	(0.00298)	(0.00286)			(0.00212)	(0.00203)			(0.00166)	(0.00167)		
(Test score (Grade 5) > avg	0.577***	0.568***			0.585***	0.574***			0.594***	0.578***		
	(0.00687)	(0.00502)			(0.00515)	(0.00355)			(0.00433)	(0.00301)		
(Fam. Background < avg) × (Test score (Grade 5) > avg)			0.893***	0.877***			0.898***	0.890***			0.948***	0.898***
			(0.0131)	(0.0102)			(0.0103)	(0.00794)			(0.00858)	(0.00646)
(Fam. Background > avg) × (Test score (Grade 5) < avg)			0.172***	0.178***			0.160***	0.172***			0.155***	0.170***
			(0.00815)	(0.00842)			(0.00658)	(0.00635)			(0.00524)	(0.00505)
(Fam. Background > avg) × (Test score (Grade 5) > avg)			1.110***	1.095***			1.126***	1.114***			1.162***	1.104***
			(0.0126)	(0.00986)			(0.00986)	(0.00721)			(0.00787)	(0.00582)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) < avg)			-0.0376	-0.0117			-0.0212	-0.0420**			-0.0180	-0.0461***
			(0.0291)	(0.0284)			(0.0210)	(0.0189)			(0.0161)	(0.0146)
Broadband × (Fam. Background < avg) × (Test score (Grade 5) > avg)			-0.122***	-0.0160			-0.0752***	-0.0398**			-0.119***	-0.0730***
			(0.0322)	(0.0291)			(0.0226)	(0.0197)			(0.0177)	(0.0154)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) < avg)			0.00223	0.0466*			0.0399*	0.0414**			0.0357**	0.0235
			(0.0294)	(0.0280)			(0.0212)	(0.0189)			(0.0162)	(0.0146)
Broadband × (Fam. Background > avg) × (Test score (Grade 5) > avg)			-0.0319	0.0432			-0.00227	0.0147			-0.0445***	-0.00463
			(0.0304)	(0.0285)			(0.0218)	(0.0185)			(0.0166)	(0.0146)
Observations	150,866	146,693	150,866	146,693	321,619	313,028	321,619	313,028	724,839	704,228	724,839	704,228
R-squared	0.458	0.458	0.352	0.357	0.447	0.455	0.338	0.350	0.456	0.457	0.349	0.352
Year×HTTWA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Grade 8 School FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Grade 5 School FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
School variables	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Peer effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Municipality variables	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
People per km2	0-92	0-92	0-92	0-92	93-202	93-202	93-202	93-202	203-6,701	203-6,701	203-6,701	203-6,701

Notes: This table reports regression results from the value-added model in Eq. 3, estimated separately for three groups defined by HTTWA-level population density. The dependent variables are students' standardised numeracy and literacy scores in grade 8. ***Family background*** captures students' economic, social, and cultural conditions, based on observed family characteristics, and is constructed using the OECD ESCS index. A complete list of covariates is provided in Table 1. All regressions include student-level covariates, grade 5 and grade 8 school fixed effects, and HTTWA×year 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 A16: Political influence

	(1)	(2)	(3)	(4)	(5)	(6)
			Broadband access			
Same political party	-0.162***	-0.113***	-0.115***	0.0345	-0.0239	0.00967
	(0.0263)	(0.0286)	(0.0283)	(0.0297)	(0.0317)	(0.0378)
Observations	7,379	7,379	7,367	7,365	7,297	5,897
R-squared	0.005	0.033	0.142	0.367	0.486	0.610
			Excess broadband			
Same political party	-0.0292	0.0442	0.000320	0.00397	-0.0625*	-0.0490
	(0.0308)	(0.0334)	(0.0336)	(0.0361)	(0.0377)	(0.0421)
	7,379	7,379	7,367	7,365	7,297	5,897
	0.000	0.025	0.111	0.307	0.459	0.613
Municipality FE	-	✓	✓	✓	✓	✓
Year FE	-	-	✓	-	-	-
Year×region FE	-	-	-	✓	-	-
Year×province FE	-	-	-	-	✓	-
Year×HTTWA FE	-	-	-	-	-	✓

Notes: This regression tests whether local councils were able to influence the rollout of the broadband infrastructure. Robust standard errors are clustered at the municipality level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

B. Expected broadband

In this section, I discuss the construction of the *expected broadband* proxy. The approach is based on the assumption that the sequential deployment of the infrastructure depends on the proximity to the closest node and on other factors that can be considered as-good-as-random or that are absorbed by the additional terms included in the main specification. The likelihood for a census area to be connected to the network is presented as a function of the logarithmic distance between the census area v and the nearest node n, $distance_{nv}$, and the node's capacity to efficiently deploy fibre coverage across space, $decay_{n,t}$:

$$E(BA_{vn,t}) = \frac{1}{1 + \exp(\phi_t distance_{nv} - \psi_t decay_{n,t})}. \tag{B1}$$

The probability of being connected to the broadband in year t decreases with distance and increases with the decay parameter, which proxies node n's capability. The parameters ϕ_t and $decay_{n,t}$ are estimated through logistic regression in each year t, using the actual availability of fibre based on $distance_{nv}$ and $decay_{n,t}$.

Once the expected broadband proxy has been computed at the census tract-level, I construct a catchment area-level proxy using the same methodology presented in Section 4.2. The proxy $E(BA_{ct})$ is constructed as the weighted average of the expected broadband coverage measured in the catchment area of school c. The weights used are the share of students living in each c area v over the total catchment area.

$$E(BA_{ct}) = \sum_{v} E(BA_{cvt}) \left(\frac{n_{vc,t-l}}{N_{c,t-l}} \right) \tag{B2}$$

C. Toy Model

This section presents a basic toy model to link the theoretical hypotheses developed in Section 2 to the empirical analysis. I study the effect of changes in access to high-speed internet on learning outcomes using a simple production function. Following Sanchis-Guarner et al. (2021) I distinguish two main mechanisms:

1. ICT improvement can change the productivity associated with a given amount of time spent studying: ed-tech resources (+) vs the negative consequences of excessive screen time (-).

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

This model simply extends the one proposed by Sanchis-Guarner et al. (2021) 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 maximise productivity gains, whereas disadvantaged students may be less likely to offset the negative 'online-gaming effect' and more exposed to excessive screen time and other harmful practices. 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:

$$H_{i2} = A_{i2} L_{i2}^{\alpha} H_{i1} \tag{C3}$$

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_c^A and μ_c^L). Access to high-speed internet (S) affects both the student productivity shifter and the time spent studying:

$$A_{i2} = S^{\delta(B)}\lambda_i\mu_c\varepsilon_i \tag{C4}$$

$$L_{i2} = S^{\eta}\tilde{\lambda}_i\tilde{\mu}_c\tilde{\varepsilon}_i \tag{C5}$$

Following a basic labour supply equation, η captures the price effect, since S affects the relative attractiveness of studying compared to leisure activities, online or offline. On the other hand, δ, which depends on students' background (B), captures the effect on individual productivity. Depending on the level of parental supervision, students' productivity can be enhanced by Edu-tech resources or negatively affected by the harmful consequences of a misuse of digital technologies.

Substituting (2) and (3) into (1) and taking logs, I obtain the following equation:

$$\ln H_{i2} = [\alpha\eta + \delta(B)]\ln S + \lambda_i^* + \mu_c^* + \ln H_{i1} + \varepsilon_i^* \tag{C6}$$

Where $\mu_c^* = \ln\mu_c + \alpha\ln\tilde{\mu}_c$, $\lambda_i^* = \ln\lambda_i + \alpha\ln\tilde{\lambda}_i$, and $\varepsilon_i^* = \ln\varepsilon_i + \alpha\ln\tilde{\varepsilon}_i$. The hypothesis tested is that the interplay between $\alpha\eta$ and δ 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:

$$\ln H_{i2} = \beta(B)\ln S + \lambda_i^* + \mu_c^* + \ln H_{i1} + \varepsilon_i^* \tag{C7}$$

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