High-speed Broadband, School Closures and Educational Achievements[1]

Filippo Boeri[2]

Abstract

The pandemic shock has led to a new awareness about the importance of ICT in our society, especially in education. However, there is still limited evidence on the impact of high-speed internet on student performance. In this study, I shed new light on the short-run effects of access to high-speed internet on educational disparities, by analysing a large infrastructural program implemented in Italy between 2015 and 2020. While previous contributions use discontinuous jumps in the available broadband connection speed across space at a given moment in time, this study exploits the actual rollout of an infrastructural policy associated with an increase in 30 Mbps household broadband coverage from 40% to 80% over a 4-year period. The estimation strategy relies on a unique dataset combining panel data on 1,800,000 students belonging to four different cohorts with a rich set of time-varying school- and student-level information and broadband data measured at a very fine spatial scale.

On the one hand, I am able to account for student performance in the previous grade and unobservable school level characteristics. On the other hand, I exploit the exogeneity of the timing of the roll-out of the broadband infrastructure. Since the projects aims to cover 100% of the territory within five years, the timing was mainly dictated by technical considerations and other quasi-random factors which are unlikely to correlate with students' attainments.

Finally, I focus on the mitigating role of high-speed broadband during the Covid-19 pandemic. By exploiting regional differences in the number of days of school closure, I analyse to what extent access to high-speed internet helped reducing learning loss for different student types. Results show an average positive effect of high speed internet broadband on educational achievements. However, this masks substantial heterogeneity: only students with better backgrounds gain from internet speed. School closures significantly amplified this result. Overall, the results suggest that the increasing access to ICT could in the short-run increase the educational gap between low performers with poor family background and other students.

Keywords: ICT, education, economics, internet, broadband, Italy

JEL Codes: I24, H54, D83

---

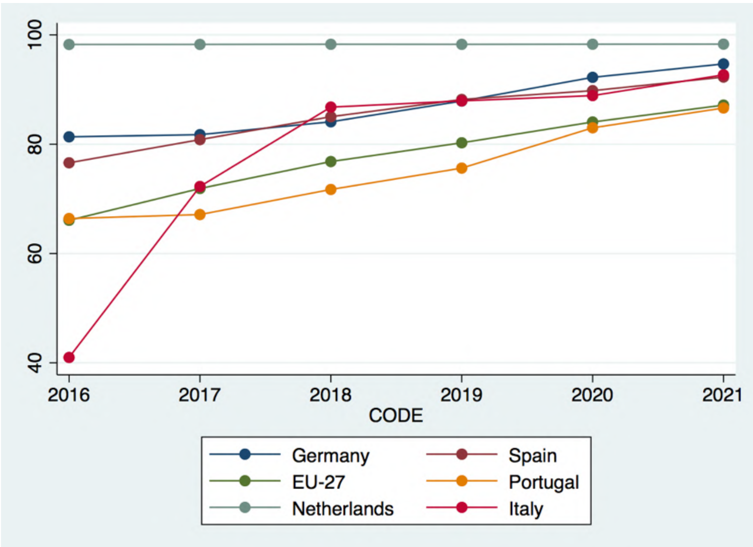
[1] I thank the participants of the American Urban Economics Meeting 2018 at Columbia University and Economic Geography Early Career Seminar 2017 at the LSE for comments and suggestions. I am particularly grateful to Gabriel Ahlfeldt, Mirko Draca, Steve Gibbons, Henry Overman, Anna Piil Damm, Olmo Silva and Daniel Sturm for their insightful comments.

[2] Department of Geography and Environment, London School of Economics; Centre for Economic Performance (CEP)

1 Introduction

Nowadays, home computers have become an essential tool of modern education in the developed countries. According to the last OECD[3] report on ICT and education, access to home computers is now nearly universal in most OECD countries. However, data still show significant disparities in the access to and quality of home computing. The digital divide is often related to various levels of access to high-speed internet connections and for this reason many countries have invested relevant amounts of public funding in order to upgrade information and communication technologies (ICT) with the aim of increasing the available internet connection speeds. In spite of this, the actual impact of ICT on student performance, as on several other social outcomes, is still debated (Machin, Sandra & Silva, 2007; Barrera-Osorio & Linden, 2009; Checchi, Rettore, & Girardi, 2015; Cristia, Ibarrarán, Cueto, & Santiago, 2017; Faber et al., 2015). Recently, Italy, one of the lowest performers in the Pisa tests among OECD countries, started to gradually reduce this historical gap with other European countries in access to next generation access (NGA) broadband services. Fig. 1 illustrates the rapid increase in the share of Italian households with access to 30 Mbps internet, comparing to other European countries.

Figure 1: Evolution of 30mb/s broadband coverage across European Countries



[3] Organisation for Economic Co-operation and Development

Figure 2: Evolution of 30mb/s broadband coverage across municipalities

Notes: the maps show the progressive evolution of 30 mb/ps coverage in Italy between 2015 and 2019. For every year, I access sub-municipality level data recorded at the end of March.

The rapid rollout of broadband infrastructure in Italy took place right before the outbreak of Covid-19. Since the beginning of the pandemic, in a bid to contain the number of cases, most countries imposed severe lockdown measures. Between March 2020 and December 2021, schools were closed for several months in most affected countries. These disruptions have raised concern over possible learning losses, adverse socio-emotional effects and mental health issues among students (OECD 2021). According to the United Nations Development Programme, 2020 was characterised by an unprecedented decrease in human development of about -0.025 (UNDP 2020). The lockdown measures had a particularly severe effect on students from disadvantaged backgrounds, causing a general increase in educational inequalities.

This paper sets out to estimate the causal effect of upgrades to the available internet speed on educational achievements, before and after the covid-19 outbreak. The identification strategy relies on the specific features of a policy implemented by the Italian Government in 2014; the 'National Ultra-Broadband Plan' (NUBP) is a national plan aimed at ensuring 100% coverage at 30 Mbps and 85% coverage at 100 Mbps by 2020. This study exploits two specific characteristics of the plan. In order to cover the whole territory in a relatively short period, whilst minimising public spending, the NUBP was implemented progressively in adjacent territories (see Figure 2). As a result, the timing of the implementation can hardly be related to variables associated with educational outcomes. Furthermore, the analysis is conducted following four cohorts of students over time. This way, I am able to control for possible time-invariant school/municipality characteristics that might be correlated with the ability of local council to "capture" the central legislator and obtain a preferential treatment in the rollout of the infrastructure. In addition, I focus on the broadband supply measure, rather than measuring its actual consumption. This allows me to bypass the endogeneity that may characterise the internet usage measures frequently used in the literature. Prior student's performance and a rich set of time-variant student characteristics allow me to precisely identify the policy effect and to investigate its heterogeneity across different student types.

Figure 3: Weeks of school closure, 2020/21

Results on pre-pandemic period suggest a small but significant positive effect of 30 Mbps broadband on student performance in numeracy subjects, whereas, consistently with the literature, no such effect is found for literacy scores. When heterogeneity in social background is accounted for, this impact remains positive and significant for the best performers in the previous grade, while it becomes non-significant or negative for disadvantaged students. Overall, students' backgrounds seem to play a relevant role in the heterogeneous policy outcome. The pandemic significantly affected student performance regardless of previous performance and family background. In this context, access to high-speed broadband helped to mitigate the learning loss, but at the same time exacerbated inequalities across different social groups.

This study contributes to the literature on social outcomes of the internet in two ways. First, it focusses on the impact of high-speed internet broadband on educational outcomes where previous studies have primarily focused on different outcomes, such as employment and productivity (Akerman et al., 2015), electoral outcomes (Falck et al., 2014; Campante et al., 2017), marriage rates (Bellou, 2015) and housing prices (Alhfeldt et al., 2017). A few studies have focused on the relationship between home computer technology and student achievement. For example, Beltran et al. (2006) use a large panel dataset of US students to explore the causal relationship between computer ownership and various educational attainment levels, with a specific focus on high school graduation. Fairlie and Robinson (2013) conducted a large field study involving almost 8,000 students enrolled in grades 6-10 in 15 different middle and high schools in the United States. Fiorini (2010) use data from the Longitudinal Study of Australian Children (LSAC) to analyse the causal relationship between computer usage and children's cognitive and non-cognitive skills. They are able to test individual skills at two moments in time, instrumenting computer usage with their parents' previous computer ownership. Results generally exhibit computer usage to have a positive effect on cognitive skills, whereas the results are mixed for non-cognitive skills. Studies directly relating internet access to educational outcomes are generally descriptive. To my knowledge, the only exception is Faber et al. (2015). Exploiting randomly placed jumps in the available ICT across neighbouring residences, the authors investigate the causal effect of a sensible increase in available internet speeds on educational outcomes in the United Kingdom. The second relevant contribution of this work is that it exploits the features of a large infrastructure policy, bypassing the endogeneity issues that usually affect similar studies. This study also contributes to a recent literature on the detrimental effects of Covid-19 pandemic on pupils' educational outcomes and inequalities. Other studies focus only consider the role played by student-level characteristics (Contini et al., 2020; Agostinelli et al., 2020) or school-level endowments (Gavosto and Romano 2020), while overlooking the importance of local broadband infrastructures. This stuy, contributes to the literature directly studying the nexus between access to high-internet broadband with the heterogeneous effect of the pandemic on student performance.

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

2 Institutional Background

2.1 The National Ultra-Broadband Plan

In 2014, the Italian Government set up the '*National Ultra-Broadband Plan*' (Piano Nazionale Banda Ultra-Larga - NUBP), a massive program aiming to ensure 100% coverage at 30 Mbps and 85% coverage at 100 Mbps 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 rollout of broadband networks. The national territory was classified into three different areas according to existing or expected future broadband infrastructure deployment:

1. *White areas*: areas where no provider of broadband services is currently operating and where no such provider is expected to enter the market in the coming three years.
2. *Grey areas*: areas where one (infrastructure-based) provider is already active, but another network is unlikely to be developed in the next three years.
3. *Black areas*: areas where there are or there will be in the next three years at least two basic broadband networks of different operators.

The NUBP is based on four main pillars. First, the State guarantees administrative simplification and a reduction in burdens for all of the target regions. Second, private investments are encouraged in black and grey areas through the creation of tax exemption tools for infrastructure operations. Grey areas also benefit from various measures to facilitate the access to financial resources, the establishment of a guarantee fund and access to credit at subsidised rates. Finally, in white areas (often incorrectly defined as "market failure areas[4]") the Public Sector intervenes directly to realise the infrastructures.

[4] These areas are defined According to Article 107(3)(c)TFEU (Treaty of Functioning of the European Union). State aid is allowed in order "*facilitate the development of certain economic activities or of certain economic areas, where such aid does not adversely affect trading conditions to an extent contrary to the common interest*". In the Communication from the Commission on EU State Aid Modernisation (SAM). Brussels, 8.5.2012. COM(2012), it is stated that "*State aid policy should focus on facilitating well-designed aid targeted at market failures and objectives of common European interest. State aid measures can, under certain conditions, correct market failures, thereby improving the efficient functioning of markets and enhancing competitiveness. Further, where markets provide efficient outcomes but these are deemed unsatisfactory from a cohesion policy point of view, State aid may be used to obtain a more desirable, equitable market outcome. In particular, a well-targeted State intervention in the broadband field can contribute to reducing the 'digital divide' between areas or regions where affordable and competitive broadband services are on offer and areas where such services are not.*"

After an initial delay, due to a number of legal disputes, in 2015 the program started to produce positive effects. Between 2015 and 2017 Italy managed to significantly reduce the historical gap with the other large European countries (see Figure 1). The 30 Mbps broadband penetration rate increased between 2014 and 2018, from 20% to almost 80%. The specific characteristics of the broadband infrastructures 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 60%-95%. These significant results were made possible by the availability of EU structural funds (the European Regional Development Fund, European Agricultural Fund for rural development and the Development and Cohesion Fund), which complemented public funds, and the general inadequacy of existing infrastructures, offering an average broadband speed below 2 Mbps.

2.2 The Italian School System

The Italian education system consists of 4 stages: nursery school (children between 3 and 6 years of age), primary education (children between 6 and 11), first grade (lower) secondary school (between 11 and 14 years of age) and second grade (upper) secondary school (from 14 to 19 years of age). Once these stages have been successfully completed, students can access the higher education offered by universities, institutes for Higher Education in Art and Music as well as Higher Technical Institutes. Education is compulsory for ten years, between the ages of 6 and 16. As a result, all students are expected to gain at least a 'Licenza media' (lower secondary school diploma). Over the last decade, the country experienced a reasonable decrease in the number of high school early dropouts. In 2014, only 1.6% (mostly first-generation foreigners) of the population in the 16-19 year-old cohorts did not hold a lower secondary school diploma.

Primary and lower secondary school together form the first cycle of education, lasting 8 years. According to the new ministerial guidelines, the general aim of lower-secondary education is '*the harmonious and comprehensive development of the individual, according to the principles of the Italian Constitution and European cultural tradition, to be achieved through the promotion of knowledge, respect for individual diversity and the active involvement of students and their families*' (Framework for Key Competences for Lifelong

Learning set up by the European Parliament and the Council of the European Union through the Recommendation). The subjects taught in this stage are: Italian, English, a second foreign language, mathematics, science, technology, geography, history, music, art, sports science and Catholic religious education (optional). Schools are expected to provide 30 hours of teaching per week (990 hours per year), allocated according to a common timetable (see A. 1 in the Appendix).

School Councils can offer to some or all classes an 'Extended timetable' (from 36 to 40 hours per week). In this case, the mandatory education goals remain the same, but students are expected to allocate less time to at-home study. At the end of the three-year program, students need to pass a uniform national examination in order to obtain a diploma and to access the following stage. The examination consists of a national written test set by INVALSI (also used by the Institute as a national assessment for grade 8) and four written tests set by a mixed (internal-external) committee. The subjects covered in the tests are Italian language, mathematics, science, ICT and two foreign languages.

In contrast to the following two cycles (upper secondary school and tertiary education), the primary and lower-secondary schools are characterised by a very low, if not entirely absent, degree of autonomy (Ichino & Tabellini, 2014). First, individual schools have almost no autonomy in the design of the education programs. The national Ministry designs the course contents, defines the number of hours to allocate to each subject and authorises a limited number of textbooks for each field. Single institutions are only allowed to use a limited budget to set up laboratories and extra-program activities, such as optional courses. Teachers are allowed to choose among a certain number of authorised textbooks for each subject and can design their classes based on the national program, but need to report each semester to the Ministry. Teaching methods and contents must be consistent with each school's educational offer plan, which in turn must be consistent with the educational goals established at the national level. Second, the homogeneity in the service is also guaranteed by a rigid financial system. Between 97% and 100% of the school budget depends on transfers from the Central Government. Every three years, the Ministry allocates resources based on specific criteria (i.e. number of students, number of disabled students, specific needs, etc.). The uniformity of the system is also guaranteed by the human resource management, which is primarily conducted at the national level. Teachers apply to province-level lists and are assigned a ranking. Vacancies in each school are covered on the basis of teachers' preferences and rankings, with little to no involvement of the school directors. Salaries and career development are defined by national agreements.

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

3 Theoretical framework

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

1. ICT improvement can change the productivity associated with a given amount of time spent studying (MOOC effect).
2. High-speed internet can affect the supply of time spent studying relative to leisure activities (online-gaming effect).

This model simply extends the one proposed by Faber et al. (2015) by taking into account any potential background-biased effect of ICT access on students' productivity. In this framework, students with a better family background are expected to maximise productivity gains, whereas disadvantaged students may be less likely to offset the negative 'gaming effect'. This assumption is built on the rich - although mostly qualitative - literature investigating the factors related to the relationship between student performance and their access to ICT[5]. Consistently, student i's knowledge production function is given by:

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

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

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

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

[5] Several studies suggest that the perception of their own ICT abilities (Aesaert and Van Braak, 2007), their actual competencies (Tounder et al., 2010; Aesaert et al., 2015) as well as their profile of ICT use (Scherer et al., 2010) might be influenced by family background.

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

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

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

The main hypothesis tested is that the interplay between η 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:

$$H_{i2} = \beta(B)lnS + \lambda_i + \mu_s + H_{i1} + \varepsilon \qquad (5)$$

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

4 Data

4.1 Main Sources

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

Data come from four main sources:

a. Schooling outcomes

The Italian National Institute for the Evaluation of the Education System (INVALSI) is a research institute with the status of a legal entity governed by public law. It is responsible for the annual assessment of the competencies of Italian students in both reading and mathematics. Tests are taken at a number of given grades (2, 5, 6, 8 and 10) and at a national level. Historically, the test was optional and suffered from a significant spatial bias in response rate. Since 2012, the test has become compulsory and recently has reached a 97% coverage at secondary school level (2016/2017). Every year, the Institute publishes anonymised microdata on student performance[6]. Since 2008, individual marks have been matched with a rich set of individual information, allowing for control of personal, family and school characteristics. I have information on the test results for the whole population of students in the 2012/2013, 2013/2014, 2014/2015, 2015/2016 2016/17 and 2017/2018 school years at grades 5 and 8 and have retrieved from the dataset individual information for three cohorts of students: 560,000 students born in 2000, who completed grade 5 in 2013 and grade 8 in 2016; 555,000 students born in 2001, who completed grade 5 in 2014 and grade 8 in 2017; 548,000 students born in 2002, who completed grade 5 in 2015 and grade 8 in 2018. Invalsi dataset also provides a rich set of variables concerning student characteristics. In addition to the student-level information, I have access to the level of education (according to the ISCED scale) of both parents, as well as to their occupational status,

[6] As highlighted by a vast literature (see Bertoni, 2013 for a review) a common problem with test-based accountability systems in education is that they generate incentives for teachers, students and school administrators to manipulate test outcomes. For this reason, the dataset reports both raw and cheating-adjusted results, obtained by means of a probabilistic algorithm that takes into account personal information (beyond that available to the public), score fluctuations and suspicious patterns. Moreover, the algorithm takes into account relevant discrepancies with the 5% classes that, for every test, have been assigned an external examiner.

recorded in the Socio-Economic Index of occupational status (ISEI) and an index based on individual-level economic, social and cultural information (ESCS)[7]. I also have access to data on home computer ownership and availability of an internet connection before and after the policy rollout.

b. Ministry of education

In 2011, the Ministry of Education, in accordance with the Community Guideline on public access to information held by public authorities, started to publish data on each State-recognized school, for any grade. Since 2012, all schools provide, on a yearly basis, information concerning the number of students enrolled in each grade by gender and nationality, number of classes, number of teachers, the school basic budget and a self-evaluation document, to be sent to the Ministry at the end of the academic year. Moreover, 80% of schools provide information on staff (age, type of contract, level of training), environment (desktops, ICT technologies, Wi-Fi coverage) and other relevant features. For institutes providing different stages of education, data are gathered for each stage. Moreover, each school provides the full address of each building (*plesso*) belonging to the school.

c. 2011 Italian socio-economic Census

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

[7] The ESCS is a composite score built by the indicators parental education (ISCED), parental occupation (ISEI), and home possessions (HOMEPOS) via principal component analysis. The index was developed by the OECD Pisa (Programme for International Student Assessment) and is often used in international standardised tests. Invalsi builds the HOMEPOS component using information about the availability of a 'quiet place to study', a private room, a desk, encyclopaedias, a personal computer and access to the internet.

micro-data have been linked to a complete digital database in ArcInfo format at a scale of 1:25.000, integrating remote sensing images, IGMI maps and technical maps at regional level with information relating to the municipality. This advanced methodology allowed the Istat to produce detailed geocoded data on the Italian territory, which is divided into 402,000 areas. On average, each section hosts 142 people and for each one, the Istat releases information concerning the number of people living in each division, by gender and age class. Furthermore, the dataset can be matched to information on wage, occupational status and other social features on the basis of a 5-10% population sample living in each division.

d. Infratel area-level data

The Infratel dataset contains information gathered through the monitoring process of the Italian Ultra-Broadband Plan. In 2014, a consultation was carried out in order to collect information on the availability of broadband telecommunication infrastructures and on the private investment plans for the following three years. Infratel divided the entire Italian territory into 94,645 areas and separately assessed each one. Once it had obtained data on internet speed and had assigned each area to one of the four clusters, according to the presence/interest of private actors, Infratel, on behalf of the Ministry of Local Development, started to update the dataset every three months. Data are based on information provided by the private sector at the beginning of the process together with updated data from the ministry relating to the implementation process. As a result, the dataset is able to provide a comprehensive representation only for white areas, where only the public sector operates. Specifically, the dataset provides historical data on:

- Population coverage to at least 100 Mbps.
- Population coverage to at least 30 Mbps.
- Public Administration coverage to at least 100 Mbps.
- Businesses coverage to at least 30 Mbps.
- Businesses coverage to at least 100 Mbps.

Population coverage is measured as the percentage of houses (flats) that have access to 30/100 Mbps internet speed. The dataset can be linked to a spatial dataset providing geocoded boundaries for each section. In this way, the information gathered can be easily matched with other spatial datasets.

e. Final Dataset

In the final dataset, the census areas are matched with the Infratel areas. In this way, I can create a new micro-aggregated spatial unit, corresponding to the Infratel areas, reporting weighted broadband coverage measures, where the weight is the relative share of students living in the area (see section 4.2 for a complete explanation of the full procedure). Information relating to individual students is matched with school data provided by the Ministry of Education and, via the school coordinates, to the data associated with the territory in which each school is located. Thus, for each student, I have the results in the national exams, a rich set of individual and family characteristics, various information about the school attended and the town in which they live and the weighted broadband coverage measured in the proximity to the school. A complete description of the variables used in the analysis is reported in Tables A. 2 and A. 3.

4.2 Broadband Measure and Catchment Areas

Ideally, in order to correctly identify the effect of the policy, I would assign students' homes to treated and control groups. As I do not have information student addresses, I assign students to treatment and control groups by defining geographical catchment areas around each school. Each area c is defined so as to guarantee that student i attending school s lives within the defined borders. In this way, I can assign the whole catchment area to a specific group and subsequently focus the analysis on treated and non-treated pupils living in neighbouring areas. De Simone (2013) developed a method to identify an area within which most resident students would be attending a specific school.

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

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

1. identify the school type (primary, lower secondary or upper secondary) and, consequently, the relevant student population in the census area (population aged 5-9 years, 10-14 years or 15-19 years respectively);
2. compute the distance between school j and the 400 nearest census areas;
3. for each school, neighbouring census areas are sorted by distance (in ascending order);
4. compute the areas' cumulative relevant population;
5. select the closest N areas so that the cumulative relevant population contains a multiple k of the number of students enrolled in school j.
6. For each school, I create two different catchment areas: a small catchment area (k = 1), where at least 80% of the enrolled students live and a large one (k = 3), where all enrolled students (and many non-enrolled) live.

Once the data have been extracted and the catchment areas defined, I can build a proxy for Broadband Access, BA_{ct}^{sh}, obtained as the weighted average of the broadband coverage measured in each catchment area *c*. The weights used are the share of students living in each Infratel area (aggregated from the smaller census areas) to the total catchment area.

$$BA_{ct} = \sum_{i \in c} ICT_{ict} \left(\sum\nolimits_{s \in i} \frac{n_{sict}}{N_{ict}} \right)$$

where ICT_{ict} is the share of buildings with access to internet broadband in the Infratel area *i*, belonging to catchment area c , n_{sict} is the number of students living in the census area s and N_{ict} is the total number of students living in Infratel area *i*, in catchment area *c*.

In the analysis, I drop any large catchment areas where the share of white areas (clusters C and D) is lower than 95%.

5. Empirical Strategy

5.1 Basic model

In the analysis, I first estimate the following model:

$$y_{ist} = \beta BA_{ct} + \gamma X_{ist} + \rho S_{st} + \theta W_{ct} + \mu_s + \nu_u + \lambda_t + \epsilon_{ist} \quad (6)$$

where y_{ist} represent student's i achievement in school s at time t, BA_{ct} is the weighted share of buildings located in catchment area c with access to high-speed internet broadband, X_{ist} is a vector of time-variant individual characteristics, S_{st} is a vector of school-level characteristics and W_{ct} is a vector of average socio-economic characteristics of the school's catchment area. The unobservable school quality is accounted by grade 8 and grade 5 school fixed effects, respectively μ_s, and ν_u. In order to identify the causal effect of high-speed broadband on student achievement, I need to address various sources of bias. Many studies in the literature exploit data on internet usage, which often represents the only available information. The choice of this variable is problematic for a number of reasons. First, survey data, especially when covering technical information, are affected by attrition and measurement error. Moreover, internet usage is likely to be highly correlated with several variables related to the tested outcome, such as social background, profession, individual network and income. In this case, the treatment would not be orthogonal with respect to the unobservable individual characteristics. For all these reasons, I choose to focus on the broadband supply, measured as the number of houses with access to 30 Mbps internet broadband, weighted by the share of students located in each census area.

Broadband access measures are not exempted from endogeneity concerns. First, the empirical strategy must address possible selection bias. This may occur because broadband access can determine a self-selection of different groups across regions with different levels of coverage. Moreover, even excluding an active sorting based on internet speed, there are several reasons to believe that broadband rollout may be far from random. Exchange stations[8] are normally located close to central locations,

[8] Exchange stations are physical infrastructure through which Internet service providers exchange Internet traffic between their networks.

where they can guarantee the best connectivity to high income households and offices. Faber et al. (2015) address this problem by adopting a neighbouring discontinuity design, able to guarantee a high degree of homogeneity between the treated and control group. This strategy is particularly effective in addressing endogeneity concerns, but it inevitably requires the analysis to focus on a small sample of the available data. Moreover, it is possible to perform a neighbouring discontinuity design only when the treatment and outcome variables share the same level of geographical detail. This is not the case for Italian student data, which can only be linked to school catchment areas. Data on distance from the closest local exchange station are exploited by Campante et al. (2017), who study the diffusion of access to high-speed internet using Italian municipal data from 1996 to 2013. The strategy is based on the assumption that the cost of providing ADSL-based broadband services varies depending on its relative position in the pre-existing voice telecommunications infrastructure. Since the pre-existing infrastructure was not randomly distributed, the authors implicitly assume that the correlation between distance and unobserved municipal characteristics has not changed during the period considered, other than through the introduction of high-speed internet. In other words, firms and households may differ in terms of time-invariant unobservables, but are assumed to have, for instance, the same wage/productivity growth. This assumption appears to be relatively strong and would require strong supporting evidence to justify. In fact, the regional economics literature provides rich evidence of rising regional disparities in most developed countries, including Italy (A'Hearn and Venables, 2013).

Instead of relying the existing infrastructure, this paper exploits the specific design of the Italian 'National Ultra-Broadband Plan', which guarantees the timing of the rollout to be exogenous. This assumption is based on three features of the policy. First, the policy targets *white areas*, consisting in small-to-medium towns in the countryside (where the average population is 20,000) and there are no overlaps with the interests of private actors. Second, the introduction of the new fibre broadband does not depend on the pre-existing infrastructure, since no compatible infrastructure exists in *white areas* at the time the policy is implemented. Third, even though some geographical characteristics associated with the lack of coverage of any previous infrastructure may still influence the implementation costs, the policy aims to cover 100% of municipalities within three years. As a consequence, implementation costs can hardly be correlated with the rollout timing. On the other hand, an efficient implementation would imply a progressive geographical coverage. Moreover, the timing and the universal target undermine the risk of a spatial sorting of households based on treatment. These assumptions alone do not allow political bias in the implementation phase to be excluded. Local administrators may lobby to obtain full coverage before neighbouring municipalities, which would result in a selection bias. However, this issue does not appear to be particularly relevant in this context, since the program has been designed by the national Government and Italian political system and does not have a relevant representation at the local authority level. Moreover, Mayors in small towns lack the political power to deliver relevant changes to a national plan, especially when this kind of change would involve higher costs for the whole project. This issue is further addressed by using school fixed effects, that can absorb the ability of local authorities to capture resources from the central government and other time-invariant municipality-level characteristics that could be associated with student performance.

1.5.2 Value added model

Extending the basic framework, student educational attainment can be described through a value-added model (VAM)

$$y_{ist} = \alpha y_{ist-1} + \beta BA_{ct} + \gamma X_{is} + \rho S_{st} + \theta W_s + \mu_s + \nu_u + \lambda_t + \epsilon_{ist} \tag{7}$$

where y_{ist} and y_{ist-1} are, respectively, student i's achievement in school s at time t and *t*–*1*, BA_{ct} is the broadband dummy, X_{is} is the vector of time-invariant individual characteristics, S_{st} is a vector of time-variant school-level characteristics, W_s is a vector of average socio-economic characteristics of the school catchment area, μ_s and ν_u are school fixed effects, λ_t are cohort fixed effects and ϵ_{ist} is the unobserved error term. Controlling for previous performance, I partially take into account potential time-invariant differences between treated and control students. Moreover, even assuming treatment to be uncorrelated with the performance at *t-1*, $(Y_{i0}, Y_{i1} \perp BA_{ct})$, the autoregressive specification allows me to investigate the heterogeneity of the treatment across different social groups. In particular, I can test the impact of the access to high-speed broadband on student performance, as depending on prior achievement, ethnicity, socio-economic background and nursery attendance.

In the economics of education literature, VAMs have been often used to measure the importance of productivity inputs (such as teacher quality or peer effects) on student performance. Recently, a number of concerns have been raised regarding the opportunity of using VAM models to estimate teacher quality (Kane & Staiger 2008; Hanushek & Rivkin, 2010; Rothstein, 2010; Chetty et al., 2014; Condie et al., 2014). Most of these studies have discussed the possible bias resulting from student sorting and the reliability of standardised tests scores as a proxy for student achievement. These issues do not appear to be relevant in the framework of this study. First, as explained in section 1.2.2, student sorting across schools appears to be a negligible phenomenon in Italy. On the other hand, although unobservable characteristics could, in principle, influence the cheating-corrected test scores, they are unlikely to be correlated with the rollout of the broadband infrastructure. To further address these concerns, in section 1.7 I show that moving home is not correlated with neither the treatment nor the outcome.

If student performance exhibits a mean reverting pattern at the tails of the previous performance distribution, the value-added model might fail to describe the learning process. Therefore, I test a quadratic specification:

$$y_{ist} = \alpha y_{ist-1} + \nu y_{ist-1}^2 + \beta BA_{ct} + \gamma X_{is} + \rho S_{st} + \theta W_{st} + \mu_s + \nu_u + \lambda_t + \epsilon_{ist} \tag{8}$$

6. Results

6.1 Baseline models

I start the investigation by studying whether the infrastructural policy had an effect on student achievements. In Table 1, I regress students' performance in grade 8 on the weighted share of household with access to high-speed broadband in the catchment area. All specifications include grade 5 and grade 8 school fixed effects and cohort fixed effects.

Columns (1) and (2) illustrate, respectively, results for math and literacy standardised tests when no covariates are considered in the specification. High-speed internet broadband seems to have a small but significant effect on student performance in math, but no significant effect on literacy scores. Access to high-speed internet raises on average student scores in numeracy subjects by 0.056 standard deviations.

These results are confirmed in columns (3)-(4), when student-level and school-level variables are included. I also include two dummy variables that take value of one when the student had a PC at home and internet connection in grade 5. A 100% change in broadband access is associated with 4.4% standard deviation change in student scores in numeracy subjects, whereas no significant effect is found for literacy scores. Not surprisingly, the internet dummy is associated with a much larger effect on student performance. However, students with access to internet and home computer ownership in grade 5 are likely associated with unobservable individual and family characteristics that could have an effect on students' performance.

In columns (5)-(6) class peer effects (mean class values for a set of student level characteristics) and time-variant school level variables are included. The additional covariates do not change the results.

When I add time-variant municipality-level variables (columns (7) and (8)), the effect of broadband on numeracy test scores is reduced by more than 50% and becomes non-significant.

This result suggest that, once controlling for changes in the socio-economic characteristics of the catchment area, the policy does not have a direct positive effect on average student performance.

In Table 2, I estimate a simple value-added model of cognitive achievement, to investigate the effects of the rollout of the broadband infrastructure on students' learning trajectories.

In columns (1) and (2), the policy is found to have a relevant positive effect on student performance in the numeracy test, but a significant negative effect in the literacy assessment. Treated students are found to obtain, on average, 0.07 standard deviation higher scores than control ones in the normalised test scores. When adding student-level and school-level covariates, the magnitude of the coefficient slightly decrease, bur remains above 0.07 standard deviations. Columns (3)-(6) reveal that, once considered the heterogeneity in class composition and school quality, the policy has a less relevant effect on math scores. The coefficient size decreases and becomes less significant. As in the previous specification, the inclusion of time-variant municipality variables significantly reduces the broadband coefficient, but I still find a positive and significant relation between the policy and student performance in math. On the other hand, broadband is found to have a negative, although not significant, effect on literacy scores.

These results are in contrast with the 'perfect zero-effect' found by Faber et al. (2017), analysing the effect of high-speed internet on student scores in the UK. In the same way, they contradict Vigdor et al. (2014), that find internet access in North Carolina to be associated with a 2.7% standard deviation decrease in numeracy scores and a 1% decrease for literacy. These results can also be analysed with respect to other policies aimed at improving ICT use in education.

This effect appears to be significantly lower than the one found by Comi et al. (2017), that analyse the ICT-related practices in a small sample of Northern Italian schools. However, the broader effect analysed in this study would justify a lower magnitude.

More generally, the policy is found to have limited effect on student performance when compared to the main policy investigated in the literature. For instance, according to recent literature reviews (Chingos, 2012; Jepsen, 2015), the effect of reducing class size by 10 students on educational performance ranges between 0 to 50% of a standard deviation. However, educational outcomes represent only a second order objective of this kind of infrastructural policy and should be considered together with medium-long term effects on other socio-economic outcomes.

Thus far, I assumed performance in grade 8 to be a linear function of performance in grade 5. In order to account for mean reverting patterns at the lower tale of the test score distribution, in Table 3, I add a quadratic component to the estimation. The new model does not lead to significant changes with respect to the basic VAM, although the policy is found to have a marginally higher effect on numeracy scores (3.8-6.5% standard deviations).

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 socio-economic background. From this point onwards, the analysis focusses exclusively on numeracy test scores, since – consistently with the literature - in previous tables broadband access was found to have a zero effect on literacy performance.

In column (1), I interact the main regressor with the performance in numeracy tests in grade 5. The interaction term is negative and significant, suggesting the policy to be associated with a mean reverting pattern in student performance. In column (2) I further investigate this relationship, interacting the policy dummies with the quartile of the previous test score distribution. Students in the lowest quartile of the performance distribution in grade 5 significantly benefit from the policy, gaining on average 3% of a standard deviation comparing to students without access to high-speed broadband. The effect declines by 50% for students in the second quartile and becomes null in the third quartile. Students in the fourth quartile lose, 10% of a standard deviation comparing to students with no access to high-speed internet. This result is well explained in Figure 5, showing a more consistent mean reversion pattern with respect to grade V for students that benefit from the policy. In column (3) I focus on the second relevant source of heterogeneity, namely the parental background. In this case, I find positive and significant contribution of ESCS index[9] on the policy outcome. In column (4), I find a perfect zero effect of the policy on students in the lowest quartile of the socio-economic distribution, with a positive and significant effect for the other quartiles, increasing from 5.6% to 9.3% of a standard deviation.

[9] Index of individual economic, social and cultural status

Many studies emphasise the importance of parents' education for students' mathematical achievements (among others, Lauder et al., 1999; Zimmer & Toma, 2000; Martins and Veiga, 2010; Vigdor et al., 2014). In countries where women are underrepresented in the labour market, mothers play the main role in supporting children's education. In Italy, where women occupation rate is on average below 49%, BMMJ is expected be a strong predictor of students' performance. Therefore, in column (5), I interact the policy-dummy with the 1-5 BMMJ scale, that describes the occupational level of students' mothers. Results provide strong evidence of an important role of this factor on the effectiveness of the policy. The positive effect of the policy on student performance in math in concentrated among students whose mothers record high occupational levels. Finally, I investigate to what extent the effect of the policy depends on home ownership and home internet connection in grade 5. Not surprisingly, I find a positive and significant effect only for students grew in a family background where ICT devices were already available before the implementation of the policy.

Overall, results highlight an important nexus between the effectiveness of the policy and students' 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 investigate this dynamic, interacting the main regressor with performance in grade 5, within a quadratic VA framework. In columns (1) and (2) I distinguish between students above and below the mean ESCS level. As expected, low performers with above average parental background benefit the most from the policy. Average performance students, on the other hand, only benefit from the policy if they can rely on a favourable family background. These results are consistent with the patterns illustrated in Figure 6. Overall, high-speed broadband is fund to have two distinctive effects on the performance distribution. On the one hand, the policy reduces the dispersion in student performance. On the other hand, this effect is less pronounced for disadvantage students, causing a further rise in the performance gap between students with different socioeconomic backgrounds.

In table 6 I extend the analysis to the academic year 2020/21, and I analyse how the number of weeks of school closure affected student performance in the affected cohort. The specification includes the full set of individual and school-level time-variant covariates, together with time and school fixed effects (grade 5 and 8). In column (1), I find every week of school closure to be associated with a 0.7% standard deviation decline in student performance. Column (2) suggest that lockdown is less severe for high performers in grade (5). Columns (3) and (4) show that family background (ESCS) and access to access to high-speed internet broadband are important mitigating factors. Finally, the positive sign of the double interaction in table (5) suggest that the mitigating role of broadband could benefit mostly advantaged students, potentially increasing inequality in student achievements.

8 Robustness Checks

The analysis focus on "white areas", where no provider of broadband services was willing to invest at the beginning of the period. In these areas, the broadband rollout during the period is driven only by the public intervention. The strategy allows me to isolate the exogenous variation of internet speed, avoiding endogeneity issues. However, in order to support the external validity of the results, in Table A. 4 I replicate the main analyses using the whole sample. In columns (1) and (2) I estimate the base model using the full set of controls, interacting the variable of interest with student performance in class 5, as in columns (7) and (8) of Table 4. A 10% increase in high speed broadband coverage in the area is found to increase average student performance in numeracy subjects by 50% of a standard deviation, against the 30% increase recorded in the restricted sample. However the effect of high-speed internet is less dependent on prior performance. Conversely, broadband rollout has an average negative effect on literacy scores, particularly relevant for students with poor performance in class 5. In column (3) and (4) I interact the variable of interest with student's mother educational level, as in column (5) of Table 4. Table 4: VAM - InteractionsA gain, results are consistent, but the effect of broadband rollout outside white areas is less dependent on family background. Similar results are found in columns (5) and (6), where I interact the variable of interest with the ESCS score recorded in grade 5. Overall, high-speed broadband is found to have a positive effect on low-performers over the whole territories, but its contribution to the reduction of student performance dispersion is particularly relevant when we focus on white areas. The difference between regions might be due to some confounding factors that might foster average performance in non-white areas. Alternatively, we difference could be explained with the higher dispersion that characterised poorer white areas at the beginning of the period.

So far, I have documented the effect of high-speed broadband on students' performance. Unfortunately, the available data do not allow me to identify the underlying mechanisms. In particular it is not possible to distinguish between the direct effect of high-speed internet supply from home hardware upgrade that the policy might have fostered. However, it is still possible to exploit information on pc ownership in grade 5 to shed light on the effect of the internet upgrade for students that already had a personal computer and access to the internet prior to the policy implementation. In Table A. 5 I replicate the main specifications for this subsample of the student population. In columns (1)-(4), I estimate the value added model using numeracy and literacy scores as an outcome variable. The results can be compared with columns (7) and (8) of Table 2 and Table 3. The effect on both numeracy and literacy scores is lower in magnitude with respect to the main results and not significant.

In columns (5) – (6) I interact the variable of interest with student performance in grade 5. In this case, results are very close to the ones reported in column (1) of Table 4. Low-performers are found to benefit the most from the policy, regardless of the availability of a home computer before the broadband rollout. Finally, in columns (7)-(8) I interact the main regressor with the ESCS score reported in grade 5. Once again, the results are consistent with the ones reported in Table 4. The higher benefit gained by advantaged students does not seem to be explained by the availability of devices to take advantage of the internet upgrade.

9. Conclusions

In this study, I tested the impact of access to high-speed internet broadband on educational achievements. To this aim, I exploited a large infrastructural program implemented by the Italian government. Available data allowed me to investigate the heterogeneous effect of the treatment with respect to students' performance in previous grades and their family background. Overall, broadband access is found to have a positive and significant effect on educational achievements. Consistently with the literature, ICTs play a relevant role in numeracy subjects but do not seem to be relevant for literacy. The effect largely depends on the parental socio-economic background and prior performance. Low performers in grade 5, when benefiting from a rich cultural background, might take advantage of the new learning devices available thanks to the new infrastructure in order to reduce the achievement gap with their peers. On the other hand, the productivity gain is less pronounced for pupils with a poor background. This might be due to the fact that, without parental supervision, the online-gaming effect might overcome any possible positive effect on learning productivity. Alternatively, it might simply reflect a financial constraint, with poorer families not able to provide their children with the hardware necessary to benefit from the new infrastructure.

A similar result is found when analysing the effect of school closure during the Covid-19 pandemic. The detrimental effect of the lockdown on student performance is partially compensated by access to high-speed internet. However, the mitigating effect significantly depends on student background.

Although more work is required to better understand the underlying mechanisms, the preliminary empirical evidence may have relevant policy implications. ICT upgrading programs can be beneficial for students but need to be accompanied by training programs and other policies aimed at allowing disadvantaged students to access the benefits. More generally, infrastructural policies, as well as other place-based policies, are likely to affect students' overall performance and inequalities in the educational attainment.

References

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

Aesaert, K., & Van Braak, J. (2015). Gender and socioeconomic related differences in performance-based ICT competences. *Computers & Education*, 84, 8-25.

Aesaert, K., van Braak, J., Van Nijlen, D., & Vanderlinde, R. (2015). Primary school pupils' ICT competences: Extensive model and scale development. *Computers & Education*, 81, 326-344.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad (2015). The skill complementarity of broadband internet. *The Quarterly Journal of Economics* 130(4), 1781-1824.

Ahlfeldt, Gabriel, Pantelis Koutroumpis, and Tommaso Valletti (2017). Speed 2.0: Evaluating access to universal digital highways. *Journal of the European Economic Association* 15(3), 586-625.

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

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. National Poverty Center Working Paper Series 06-01. *National Poverty Center, University of Michigan*.

Bauernschuster, Stefan, Oliver Falck, and Ludger Woessmann (2014). Surfing alone The Internet and social capital: Evidence from an unforeseeable technological mistake. *Journal of Public Economics* 117 : 73-89.

Bellou, A. (2015). The impact of Internet diffusion on marriage rates: evidence from the broadband market. *Journal of Population Economics*, 28(2), 265-297

Bertoni, M., Brunello, G., & Rocco, L. (2013). When the cat is near, the mice won't play: The effect of external examiners in Italian schools. *Journal of Public Economics*, 104, 65-77

Brunello, G., Checchi, D. (2005). School quality and family background in Italy. *Economics of Education Review*, 24(5), 563577.

Bhuller, Manudeep, et al. (2013). Broadband internet: An information superhighway to sex crime. *Review of Economic Studies* 80(4) 1237-1266.

Campante, F., Durante, R., & Sobbrio, F. (2017). Politics 2.0: The multifaceted effect of broadband internet on political participation. *Journal of the European Economic Association*, *16*(4), 1094-1136

Checchi, D., & Flabbi, L. (2013). Intergenerational Mobility and Schooling Decisions in Germany and Italy: The Impact of Secondary School Tracks. *Rivista di Politica Economica*, (3), 7-57

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*(3), 295-320.

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(9), 2593-2632

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

De Simone, Gianfranco, and Andrea Gavosto (2013). Patterns of value-added creation in the transition from primary to lower secondary education in Italy. *XXVIII national conference of labour economics*, Rome.

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.

Faber, B., Sanchis-Guarner, R., & Weinhardt, F. (2015). ICT and education: Evidence from student home addresses (No. w21306). *National Bureau of Economic Research*.

Falck, O., Gold, R., & Heblich, S. (2014). E-lections: Voting Behavior and the Internet. *American Economic Review*, *104*(7), 2238-65.

Fairlie, R. W., Beltran, D. O., & Das, K. K. (2010). Home computers and educational outcomes: Evidence from the NLSY97 and CPS. *Economic inquiry*, *48*(3), 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*(3), 211-40.

Fiorini, M. (2010). The effect of home computer use on children’s cognitive and non-cognitive skills. *Economics of Education review* 29(1), 55-72.

Gibbons, S., Machin, S., & Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, 75, 15-28

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100(2), 267-71

Ichino, A., & Tabellini, G. (2014). Freeing the Italian school system. *Labour Economics*, 30, 113-128

Infratel (2017), National Broadband Plan and National ultrawide band Plan, Rome.

INVALSI, 2012a. Sistema Nazionale di Valutazione A.S. 2012/2013 Rilevazione degli apprendimenti.

INVALSI, 2015a. Sistema Nazionale di Valutazione A.S. 2015/2016 Rilevazione degli apprendimenti.

ISTAT (2011). Italian general census. Italian National Institute of Statistics, Rome.

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 Modelling*, Madison, WI (Vol. 47, pp. 106-118)

OECD (2013) PISA 2012 results in focus: what 15-year-olds know and what they can do with what they know. Oecd - Paris

Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education finance and policy*, 4(4), 537-571

Tondeur, J., Van Braak, J., & Valcke, M. (2007). Towards a typology of computer use in primary education. *Journal of Computer Assisted Learning*, 23(3), 197-206

Varian, Hal R. (2010). Computer mediated transactions. *American Economic Review* 100, no. 2: 1-10.

Verbeek, Marno, and Francis Vella (2005). Estimating dynamic models from repeated cross-sections. *Journal of econometrics* 127(1), 83-102.

Vigdor, Jacob L., Helen F. Ladd, and Erika Martinez (2014). Scaling the digital divide: Home computer technology and student achievement. *Economic Inquiry* 52(3), 1103-1119.

Tables

Table 1: Basic Model

	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband	0.0322***	0.0748***	0.0442***	0.0227	0.0450***	0.0231	0.0202	0.0257
	(0.0102)	(0.00845)	-0.0142	-0.0164	-0.0143	-0.0164	-0.0143	-0.0164
Male			0.115***	-0.121***	0.117***	-0.121***	0.117***	-0.121***
			(0.00221)	(0.00210)	(0.00218)	(0.00211)	(0.00218)	(0.00211)
I Gen. migrant			-0.137***	-0.317***	-0.125***	-0.314***	-0.126***	-0.314***
			(0.00778)	(0.00864)	(0.00781)	(0.00871)	(0.00781)	(0.00872)
Nursery			0.00613**	-0.00997***	0.00615**	-0.00995***	0.00542**	-0.00974***
			(0.00261)	(0.00264)	(0.00261)	(0.00264)	(0.00260)	(0.00264)
Early enrolled			0.00248	-0.0822***	0.00254	-0.0822***	0.00223	-0.0822***
			(0.00370)	(0.00387)	(0.00369)	(0.00387)	(0.00369)	(0.00387)
BFMJ			0.0501***	0.0467***	0.0499***	0.0467***	0.0500***	0.0466***
			(0.00126)	(0.00132)	(0.00126)	(0.00132)	(0.00125)	(0.00132)
BMMJ			0.0366***	0.0362***	0.0366***	0.0362***	0.0365***	0.0362***
			(0.00106)	(0.00110)	(0.00106)	(0.00110)	(0.00105)	(0.00110)
FISCED			0.132***	0.121***	0.131***	0.121***	0.132***	0.121***
			(0.00128)	(0.00130)	(0.00128)	(0.00130)	(0.00128)	(0.00130)
MISCED			0.120***	0.0979***	0.120***	0.0979***	0.120***	0.0979***
			(0.00131)	(0.00130)	(0.00131)	(0.00130)	(0.00131)	(0.00130)
Internet dummy			0.139***	0.163***	0.139***	0.163***	0.138***	0.164***
			(0.00305)	(0.00327)	(0.00305)	(0.00327)	(0.00304)	(0.00327)
Computer Dummy			0.0737***	0.0733***	0.0738***	0.0733***	0.0731***	0.0735***
			(0.00237)	(0.00248)	(0.00237)	(0.00248)	(0.00236)	(0.00248)
Constant	0.00229	-0.0632***	-1.184***	-1.111***	-1.182***	-1.112***	-1.184***	-1.122***
	(0.00588)	(0.00577)	(0.0127)	(0.0117)	(0.0127)	(0.0117)	(0.0153)	(0.0129)
Observations	1,435,017	1,126,031	853,716	822,688	853,716	822,688	853,716	822,688
R-squared	0.1	0.09	0.192	0.190	0.192	0.190	0.194	0.190
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipaloty variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents the results of the linear regression model reported in equation (6). The dependent variables are students' numeracy and literacy scores in grade 8. Sample includes all students located in white areas. The ESCS index is a proxy for student individual economic, social and cultural status, based on known and unknown family characteristics. MISCED/FISCED and BMMJ/BFMJ measure, respectively, the educational level and occupational status of students' parents. All regressions include grade V school, grade VII school and time 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 2: Value-added model

	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband	0.0728***	-0.0305**	0.0615***	-0.0241	0.0621***	-0.0240	0.0338**	-0.0129
	(0.0158)	(0.0153)	(0.0162)	(0.0166)	(0.0162)	(0.0166)	(0.0163)	(0.0165)
Test score - Grade 5	0.588***	0.621***	0.565***	0.599***	0.565***	0.599***	0.567***	0.601***
	(0.00315)	(0.00201)	(0.00326)	(0.00213)	(0.00326)	(0.00213)	(0.00321)	(0.00209)
Male			0.0222***	0.0140***	0.0233***	0.0138***	0.0228***	0.0141***
			(0.00183)	(0.00177)	(0.00177)	(0.00177)	(0.00177)	(0.00176)
I Gen. migrant			-0.0402***	-0.169***	-0.0307***	-0.170***	-0.0310***	-0.169***
			(0.00629)	(0.00676)	(0.00624)	(0.00680)	(0.00623)	(0.00679)
Nursery			0.000324	-0.00233	0.000318	-0.00232	-0.000645	-0.00186
			(0.00220)	(0.00222)	(0.00220)	(0.00222)	(0.00218)	(0.00221)
Early enrolled			0.0335***	-0.0578***	0.0336***	-0.0578***	0.0333***	-0.0576***
			(0.00325)	(0.00330)	(0.00325)	(0.00330)	(0.00323)	(0.00329)
BFMJ			0.0267***	0.0157***	0.0266***	0.0157***	0.0266***	0.0156***
			(0.00110)	(0.00109)	(0.00110)	(0.00109)	(0.00108)	(0.00109)
BMMJ			0.0170***	0.0144***	0.0169***	0.0144***	0.0167***	0.0144***
			(0.000852)	(0.000856)	(0.000852)	(0.000856)	(0.000847)	(0.000855)
FISCED			0.0734***	0.0377***	0.0734***	0.0377***	0.0733***	0.0375***
			(0.00109)	(0.00105)	(0.00109)	(0.00105)	(0.00109)	(0.00105)
MISCED			0.0722***	0.0254***	0.0721***	0.0254***	0.0720***	0.0253***
			(0.00111)	(0.00105)	(0.00111)	(0.00105)	(0.00110)	(0.00105)
Internet dummy			0.0560***	0.0674***	0.0559***	0.0674***	0.0547***	0.0677***
			(0.00264)	(0.00269)	(0.00264)	(0.00269)	(0.00262)	(0.00269)
Computer Dummy			0.0377***	0.0234***	0.0377***	0.0234***	0.0367***	0.0237***
			(0.00200)	(0.00201)	(0.00200)	(0.00201)	(0.00199)	(0.00201)
Constant	0.106***	-0.108***	-0.589***	-0.482***	-0.587***	-0.483***	-0.580***	-0.492***
	(0.0107)	(0.00811)	(0.0141)	(0.0113)	(0.0141)	(0.0114)	(0.0176)	(0.0130)
Observations	1,140,089	1,091,535	853,716	822,320	853,716	822,320	853,716	822,320
R-squared	0.415	0.441	0.452	0.456	0.452	0.456	0.456	0.457
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents the results of the value added model reported in equation (7). The dependent variables are students' numeracy and literacy scores in grade 8. Sample includes all students located in white areas. The ESCS index is a proxy for student individual economic, social and cultural status, based on known and unknown family characteristics. MISCED/FISCED and BMMJ/BFMJ measure, respectively, the educational level and occupational status of students' parents. All regressions include grade V school, grade VII school and time fixed effects. Robust standard errors are clustered at the school level and reported in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 3: Quadratic value-added model

	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband	0.0768***	-0.0299*	0.0656***	-0.0238	0.0661***	-0.0237	0.0381**	-0.0126
	(0.0157)	(0.0153)	(0.0160)	(0.0166)	(0.0160)	(0.0166)	(0.0161)	(0.0165)
Test score - Grade 5	0.935***	0.879***	0.889***	0.855***	0.889***	0.855***	0.883***	0.855***
	(0.0145)	(0.0129)	(0.0139)	(0.0148)	(0.0139)	(0.0148)	(0.0137)	(0.0149)
(Test score - Grade 5)^2	-0.350***	-0.253***	-0.325***	-0.248***	-0.325***	-0.248***	-0.317***	-0.247***
	(0.0156)	(0.0121)	(0.0150)	(0.0138)	(0.0150)	(0.0138)	(0.0148)	(0.0138)
Male	0.935***	0.879***	0.0223***	0.0141***	0.0234***	0.0139***	0.0229***	0.0142***
			(0.00181)	(0.00177)	(0.00176)	(0.00176)	(0.00175)	(0.00176)
I Gen. migrant			-0.0368***	-0.166***	-0.0274***	-0.167***	-0.0278***	-0.166***
			(0.00629)	(0.00675)	(0.00623)	(0.00679)	(0.00622)	(0.00679)
Nursery			0.000240	-0.00248	0.000235	-0.00247	-0.000716	-0.00201
			(0.00219)	(0.00222)	(0.00219)	(0.00222)	(0.00217)	(0.00221)
Early enrolled			0.0333***	-0.0587***	0.0334***	-0.0588***	0.0331***	-0.0585***
			(0.00323)	(0.00330)	(0.00323)	(0.00330)	(0.00322)	(0.00330)
BFMJ			0.0259***	0.0151***	0.0259***	0.0152***	0.0259***	0.0151***
			(0.00109)	(0.00109)	(0.00109)	(0.00109)	(0.00108)	(0.00109)
BMMJ			0.0163***	0.0141***	0.0162***	0.0141***	0.0161***	0.0141***
			(0.000847)	(0.000853)	(0.000847)	(0.000853)	(0.000843)	(0.000853)
FISCED			0.0731***	0.0379***	0.0730***	0.0379***	0.0729***	0.0377***
			(0.00109)	(0.00105)	(0.00109)	(0.00105)	(0.00108)	(0.00105)
MISCED			0.0723***	0.0259***	0.0722***	0.0259***	0.0721***	0.0257***
			(0.00110)	(0.00105)	(0.00110)	(0.00105)	(0.00110)	(0.00105)
Internet dummy			0.0534***	0.0654***	0.0533***	0.0655***	0.0522***	0.0658***
			(0.00263)	(0.00270)	(0.00263)	(0.00270)	(0.00261)	(0.00270)
Computer Dummy			0.0372***	0.0233***	0.0372***	0.0233***	0.0362***	0.0236***
			(0.00200)	(0.00201)	(0.00200)	(0.00201)	(0.00198)	(0.00201)
Constant	0.107***	-0.111***	-0.580***	-0.482***	-0.579***	-0.483***	-0.571***	-0.492***
	(0.0106)	(0.00810)	(0.0140)	(0.0113)	(0.0139)	(0.0114)	(0.0174)	(0.0130)
Observations	1,140,089	1,091,535	853,716	822,320	853,716	822,320	853,716	822,320
R-squared	0.107***	0.442	0.454	0.457	0.454	0.457	0.458	0.458
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES	YES	YES
Municipality variables	NO	NO	NO	NO	NO	NO	YES	YES

Notes: This table presents the results of the quadratic value-added model reported in equation (8). The dependent variables are students' numeracy and literacy scores in grade 8. Sample includes all students located in white areas. The ESCS index is a proxy for student individual economic, social and cultural status, based on known and unknown family characteristics. MISCED/FISCED and BMMJ/BFMJ measure, respectively, the educational level and occupational status of students' parents. All regressions include grade V school, grade VII school and time 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 4: VAM - Interactions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Num.	Num.	Num.	Num.	Num.	Num.	Num.
Broadband	0.00779	0.0308*	0.00534	0.0604**	-0.0499**	-0.00936	-0.0125
	(0.0160)	(0.0173)	(0.0160)	(0.0257)	(0.0197)	(0.0161)	(0.0169)
Test score - Grade 5	0.602***		0.565***	0.579***	0.572***	0.567***	0.567***
	(0.00423)		(0.00322)	(0.00317)	(0.00324)	(0.00321)	(0.00321)
Broadband # Test score (Grade 5)	-0.0574***						
	(0.00557)						
2.WLE_MAT_200_V_4		0.140***					
		(0.00602)					
3.WLE_MAT_200_V_4		0.305***					
		(0.00883)					
4.WLE_MAT_200_V_4		0.484***					
		(0.0130)					
2.WLE_MAT_200_V_4#c.broadband		0.0126**					
		(0.00621)					
3.WLE_MAT_200_V_4#c.broadband		-0.0105					
		(0.00769)					
4.WLE_MAT_200_V_4#c.broadband		-0.107***					
		(0.0120)					
ESCS- Grade 5			0.0310***				
			(0.00236)				
ESCS- Grade 5#c.broadband			0.0244***				
			(0.00291)				
2.ESCS - II				0.0912***			
				(0.00665)			
3. ESCS - III				0.104***			
				(0.00708)			
4.ESCS - IV				0.115***			
				(0.00784)			
2.ESCS - II # broadband				0.0526***			
				(0.00778)			
3. ESCS - III # broadband				0.0759***			
				(0.00841)			
4.ESCS - IV # broadband				0.0927***			
				(0.00961)			
BMMJ- Grade 5					0.0124***		
					(0.00231)		
BMMJ- Grade 5#c.broadband					0.0148***		
					(0.00261)		
1.computer_dummy						0.0232***	
						(0.00348)	
1.computer_dummy#c.broadband						0.0209***	
						(0.00447)	
1.internet_dummy							0.0417***
							(0.00439)
1.internet_dummy#c.broadband							0.0215***
							(0.00575)
Constant	-0.580***	-0.798***	-0.423***	-0.228***	-0.540***	-0.570***	-0.569***
	(0.0176)	(0.0189)	(0.0184)	(0.0677)	(0.0199)	(0.0176)	(0.0179)
Observations	853,716	853,716	852,192	618,512	772,495	853,716	853,716
R-squared	0.457	0.460	0.457	0.394	0.462	0.456	0.456
V grade School FE	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
School variables	YES	YES	YES	YES	YES	YES	YES
Peer effects	YES	YES	YES	YES	YES	YES	YES
Municipality variables	YES	YES	YES	YES	YES	YES	YES

Notes: This table presents regression results of the value-added model reported in equation (7). The dependent variable corresponds to students' numeracy scores in grade 8. In columns (1) and (2) the broadband access dummy is interacted with previous performance. In column (3), (4) and (5) the main explanatory variable is interacted, respectively, with family socio-economic scores and the occupational status of the mother. In column (6) and (7) the variable of interest is interacted with the availability of a home computer and internet access in grade V. All regressions include grade V school, grade VII school and time 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 5: Interactions – quadratic model

	ESCS<0 Num.	ESCS>0 Num.	ESCS<0 Lit.	ESCS>0 Lit.
Broadband	-0.0212	-0.0143	0.0110	0.0183
	(0.0373)	(0.0216)	(0.0434)	(0.0170)
Test score - Grade 5	0.718***	0.794***	0.596***	0.959***
	(0.0628)	(0.0267)	(0.0763)	(0.0187)
Broadband # Test score - Grade 5	0.310	0.474***	0.0428	-0.725***
	(0.206)	(0.153)	(0.288)	(0.140)
(Test score - Grade 5)^2	-0.0480	-0.182***	0.0308	-0.377***
	(0.0694)	(0.0279)	(0.0749)	(0.0182)
Broadband # (Test score - Grade 5)^2	-0.408*	-0.522***	-0.0624	0.743***
	(0.218)	(0.161)	(0.297)	(0.142)
Constant	-0.284***	-0.228***	-0.0502	-0.238***
	(0.0330)	(0.0239)	(0.0377)	(0.0159)
Observations	229,158	624,101	218,418	603,403
R-squared	0.489	0.461	0.494	0.462
V grade School FE	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
School variables	YES	YES	YES	YES
Peer effects	YES	YES	YES	YES
Municipality variables	YES	YES	YES	YES

Notes: This table presents regression results of the quadratic value-added model reported in equation (8). The dependent variable corresponds to students' numeracy scores in grade 8. The broadband dummy is interacted with students' performance recorded in grade 5. Results in columns (2) and (3) concern, respectively, student below and above the average family socio-economic score. The sample covers all students located in white areas. All regressions include grade V school, grade VII school and time 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.

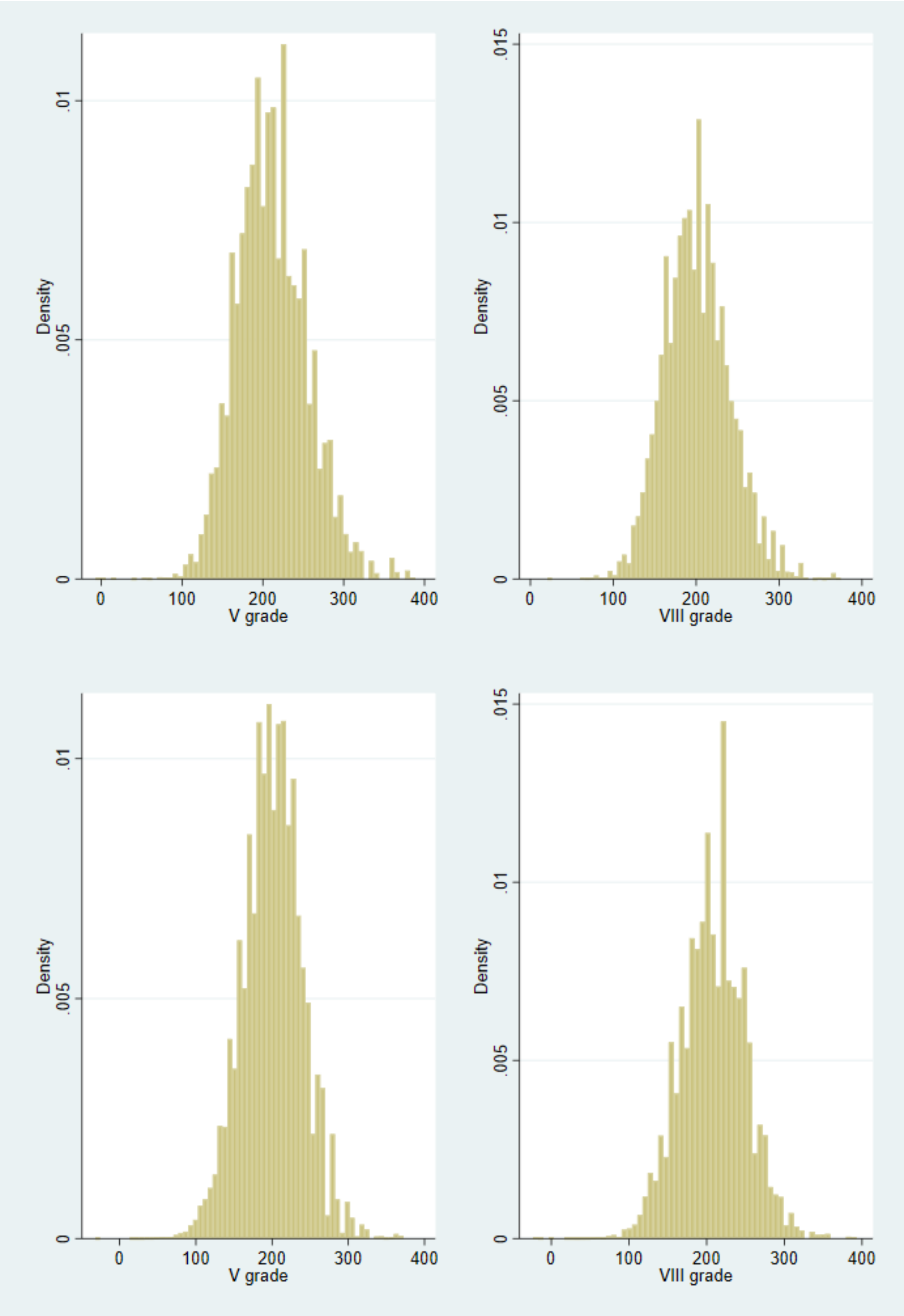
Table 6: School closures and broadband

VARIABLES	(1) Num.	(2) Num.	(3) Num.	(4) Num.	(5) Num.
School Closures	-0.00772***	-0.00930***	-0.00800***	-0.00966***	-0.00914***
	(0.000696)	(0.000902)	(0.000705)	(0.00146)	(0.00138)
Test score - Grade 5	0.601***	0.593***	0.566***	0.598***	0.578***
	(0.00299)	(0.00299)	(0.00289)	(0.00286)	(0.00288)
School Closures # Test score - Grade 5		0.00295***			
		(0.000353)			
ESCS (Grade 5)			0.254***		0.104***
			(0.00123)		(0.00227)
School Closures # ESCS (Grade 5)			0.00189***		0.00528***
			(0.000241)		(0.000677)
broadband30				0.00430	0.00824
				(0.0160)	(0.0159)
School Closures # broadband				0.00267***	0.00167***
				(0.00142)	(0.00195)
Broadband # ESCS (Grade 5)					0.0436***
					(0.00303)
School Closures # Broadband # ESCS (Grade 5)					0.00191***
					(0.000768)
Constant	0.0957***	0.115***	0.102***	0.111***	0.119***
	(0.00130)	(0.00164)	(0.00133)	(0.0112)	(0.0110)
Observations	1,529,966	1,529,966	1,511,025	1,529,966	1,294,728
R-squared	0.115	0.420	0.169	0.420	0.437
V grade School FE	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Indvidual covariates	YES	YES	YES	YES	YES
School variables	YES	YES	YES	YES	YES
Peer effects	YES	YES	YES	YES	YES
Municipality variables	YES	YES	YES	YES	YES

Notes: This table presents regression results of the value-added model reported in equation (7). The dependent variable corresponds to students' numeracy scores in grade 8. In column (2) the school closure variable is interacted with previous performance. In column (3) and (4) the main explanatory variable is interacted, respectively, with family socio-economic scores and access to 30mps internet connection. In column (5), the three covariates are included in a two-way interaction. All regressions include grade V school, grade VII school and time 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.

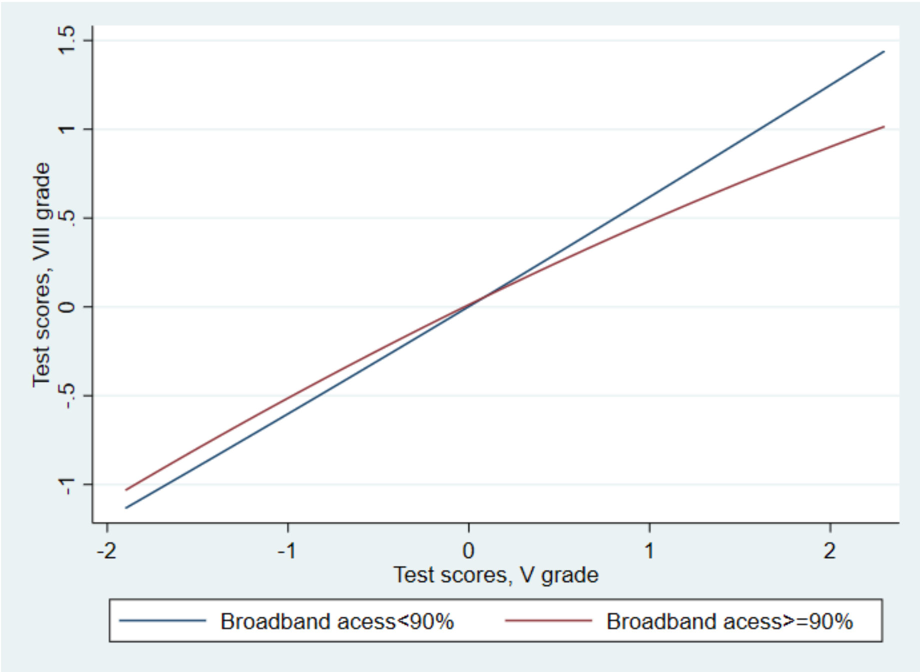
Figures

Figure 4: Test score distributions



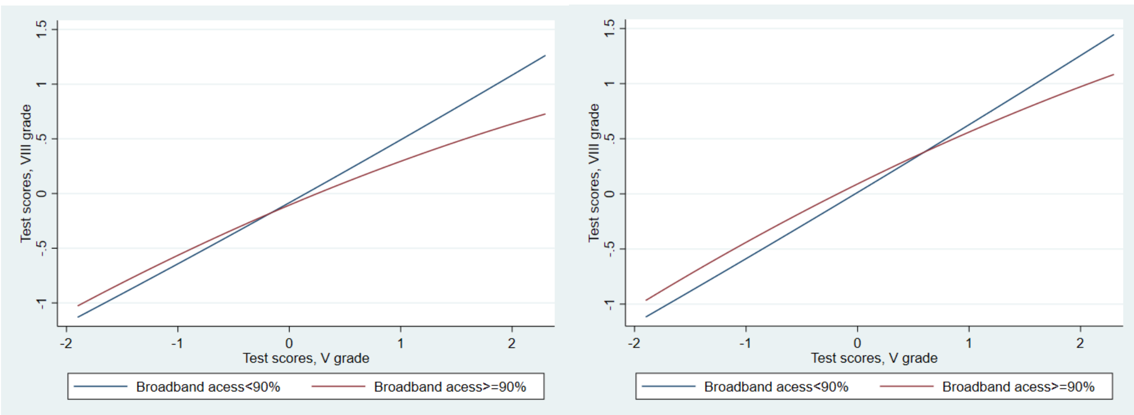
Notes: The histograms report the literacy and numeracy test score distribution for the same cohort of students in grade 5 and grade 8.

Figure 5: Grade 8 numeracy scores by performance in grade 5 and treatment



Notes. The graph reports the quadratic prediction for 8-grade students' test scores based on previous performance in grade 5.

Figure 6: Grade 8 test scores by previous performance, treatment and family background



Notes. The graph reports the quadratic prediction for students' test scores based on previous performance and family background (ESCS).

Appendix

A. 1: Subjects

Subjects	Hours per week	Hours per year
Italian Language, History, Geography	9	297
In-depth studies in literary subjects	1	33
Mathematics and Science	6	198
Technology	2	66
English	3	99
Second foreign language	2	66
Art and design	2	66
Sport science	2	66
Music	2	66
Catholic religious education	1	33
Total	30	990

Notes. The table reports hours studied by subject in Italian lower-secondary school.

A. 2 Final Dataset (VIII grade)

Variable	N	Mean	S.D.	Min	Max
WLE_MAT_200	776,610	199.30	39.06	20.6	374.2
WLE_ITA_200	626,258	209.90	41.07	-23.6	392.9
ESCS	385,483	-0.10	0.96	-3.2	2.1
male	776,685	0.51	0.50	0	1
full-time	776,685	0.21	0.41	0	1
National	776,685	0.95	0.21	0	1
EU-born foreigner	776,685	0.97	0.17	0	1
Non-EU born foreigner	776,685	0.03	0.17	0	1
Young	776,685	0.09	0.28	0	1
Old	776,685	0.07	0.26	0	1
Attended nursery	776,685	0.19	0.39	0	1
Attended pre-school	776,685	0.77	0.42	0	1
Foreign father	776,685	0.10	0.30	0	1
Non-Euborn father	776,685	0.07	0.26	0	1
Foreign mother	776,685	0.12	0.33	0	0
Non-Euborn mother	776,685	0.08	0.28	0	1
National parents	776,685	0.91	0.29	0	1
Mother's occupation: unemployed	776,685	0.04	0.20	0	0
Mother's occupation: sah	776,685	0.37	0.48	0	1
Mother's occupation: Skilled	776,685	0.07	0.26	0	1
Mother's occupation: Merchant	776,685	0.07	0.25	0	1
Mother's occupation: white collar	776,685	0.19	0.39	0	1
Mother's occupation: Unskilled	776,685	0.13	0.33	0	1
Father's occupation: unemployed	776,685	0.05	0.21	0	1
Father's occupation: sah	776,685	0.00	0.05	0	1
Father's occupation: Skilled	776,685	0.16	0.36	0	1
Father's occupation: Merchant	776,685	0.19	0.39	0	1
Father's occupation: white collar	776,685	0.15	0.35	0	1
Father's occupation: Unskilled	776,685	0.29	0.45	0	1
Mother's education: Primary	776,685	0.03	0.17	0	1
Mother's education: Lower secondary	776,685	0.31	0.46	0	1
Mother's education: Upper secondary	776,685	0.41	0.49	0	1
Mother's education: non-tertiary	776,685	0.02	0.13	0	1
Mother's education: Tertiary	776,685	0.10	0.30	0	1
Father's education: Primary	776,685	0.03	0.18	0	1
Father's education: Lower secondary	776,685	0.36	0.48	0	1
Father's education: Upper secondary	776,685	0.36	0.48	0	1
Father's education: non-tertiary	776,685	0.01	0.10	0	1
Father's education: Tertiary	776,685	0.08	0.27	0	1
I generation foreinger	776,685	0.03	0.18	0	1
II generation foreinger	776,685	0.05	0.22	0	1
BFMJ	640,550	3.66	0.98	1	5
BMMJ	668,914	2.95	1.12	1	5
HISEI	687,611	3.77	0.91	1	5
MISCED	656,296	2.69	0.93	1	5
FISCED	668,109	2.83	0.97	1	5
HISCED	682,725	3.02	0.99	1	5
Class share of male students	776,685	0.51	0.11	0	1
Class share of I generation foreign students	776,685	0.03	0.05	0	1
Class share of II generation foreign students	776,685	0.05	0.07	0	1
Log wage per capita	778,248	9.29	0.30	8.4	10.5
Earners share	778,248	62.85	8.55	38.6	118.5
Foreigners share	778,248	6.31	4.13	0.2	31.4
High earners share	778,248	2.84	1.76	0.0	30.9
High income share	778,248	13.63	6.76	0.0	67.4

Notes. The table reports descriptive statistics for the main variables of interest.

A. 3: Final Dataset (V grade)

Variable	N	Mean	S.D.	Min	Max
WLE_MAT_200_V	628,861	213.90	43.41	-32.0	372.4
WLE_ITA_200_V	775,752	199.10	39.19	-7.1	388.5
ESCS	619,170	-0.06	0.92	-3.3	2.6
male	628,861	0.50	0.50	0	1
full-time	628,861	0.95	0.22	0	1
National	628,861	0.97	0.18	0	1
EU-born foreigner	628,861	0.98	0.14	0	1
Non-EU born foreigner	628,861	0.02	0.14	0	1
Young	628,861	0.09	0.29	0	1
Old	628,861	0.02	0.13	0	1
Attended nursery	628,861	0.19	0.39	0	1
Attended pre-school	628,861	0.79	0.41	0	1
Foreign father	628,861	0.08	0.28	0	1
Non-Euborn father	628,861	0.06	0.23	0	1
Foreign mother	628,861	0.11	0.31	0	1
Non-Euborn mother	628,861	0.07	0.26	0	1
National parents	628,861	0.92	0.28	0	1
Mother's occupation: unemployed	628,861	0.04	0.21	0	1
Mother's occupation: sah	628,861	0.36	0.48	0	1
Mother's occupation: Skilled	628,861	0.07	0.26	0	1
Mother's occupation: Merchant	628,861	0.07	0.25	0	1
Mother's occupation: white collar	628,861	0.20	0.40	0	1
Mother's occupation: Unskilled	628,861	0.12	0.32	0	1
Father's occupation: unemployed	628,861	0.05	0.21	0	1
Father's occupation: sah	628,861	0.00	0.06	0	1
Father's occupation: Skilled	628,861	0.16	0.36	0	1
Father's occupation: Merchant	628,861	0.19	0.39	0	1
Father's occupation: white collar	628,861	0.15	0.36	0	1
Father's occupation: Unskilled	628,861	0.28	0.45	0	1
Mother's education: Primary	628,861	0.02	0.15	0	1
Mother's education: Lower secondary	628,861	0.29	0.45	0	1
Mother's education: Upper secondary	628,861	0.42	0.49	0	1
Mother's education: non-tertiary	628,861	0.02	0.13	0	1
Mother's education: Tertiary	628,861	0.10	0.30	0	1
Father's education: Primary	628,861	0.03	0.17	0	1
Father's education: Lower secondary	628,861	0.35	0.48	0	1
Father's education: Upper secondary	628,861	0.37	0.48	0	1
Father's education: non-tertiary	628,861	0.01	0.10	0	1
Father's education: Tertiary	628,861	0.08	0.27	0	1
I generation foreinger	628,861	0.02	0.15	0	1
II generation foreinger	628,861	0.05	0.21	0	1
BFMJ	517,049	3.67	0.98	1	5
BMMJ	537,441	2.96	1.13	1	5
HISEI	551,508	3.78	0.90	1	5
MISCED	528,955	2.71	0.93	1	5
FISCED	536,177	2.87	0.97	1	5
HISCED	545,991	3.06	0.99	1	5
Class share of male students	628,861	0.51	0.12	0	1
Class share of I generation foreign students	628,861	0.03	0.05	0	1
Class share of II generation foreign students	628,861	0.05	0.08	0	1
computer_dummy	619,046	0.73	0.45	0	1
internet_dummy	619,046	0.85	0.35	0	1

Notes. The table reports descriptive statistics for the main variables of interest.

A. 4: All students

VARIABLES	(1) Num.	(2) Lit.	(3) Num.	(4) Lit.	(5) Num.	(6) Lit.
Broadband	0.0511***	-0.0191*	-0.00287	0.00183	0.0589***	-0.0150
	(0.0116)	(0.0106)	(0.0150)	(0.0137)	(0.0114)	(0.00983)
Test score - Grade 5	0.622***	0.579***	0.585***	0.603***	0.560***	0.586***
	(0.00384)	(0.00279)	(0.00346)	(0.00213)	(0.00337)	(0.00211)
Broadband # Test score (Grade 5)	-0.0995***	0.0157***				
	(0.00546)	(0.00375)				
BMMJ (Grade 5)			0.0673***	0.0411***		
			(0.00172)	(0.00160)		
Broadband # BMMJ (Grade 5)			0.0171***	-0.00545**		
			(0.00238)	(0.00227)		
ESCS (Grade 5)					0.109***	0.0740***
					(0.00179)	(0.00163)
Broadband # ESCS (Grade 5)					0.0362***	-0.00594**
					(0.00257)	(0.00238)
Male	0.0392***	0.0169***	0.0382***	0.0220***	0.0427***	0.0621***
	(0.00171)	(0.00166)	(0.00171)	(0.00161)	(0.00154)	(0.00223)
Full-time	0.0777***	0.00448	0.0884***	0.00997**	0.0774***	0.00857***
	(0.00577)	(0.00487)	(0.00586)	(0.00469)	(0.00552)	(0.00171)
I Gen. migrant	-0.0394***	-0.169***	-0.0424***	-0.166***	-0.0274***	0.00143
	(0.00599)	(0.00641)	(0.00567)	(0.00617)	(0.00486)	(0.00132)
Nursery	0.00169	-0.000874	0.0324***	0.0155***	0.0121***	0.000794
	(0.00213)	(0.00214)	(0.00219)	(0.00216)	(0.00210)	(0.00143)
Pre-primary school	0.0303***	0.0161**	0.0389***	0.00447	0.0397***	0.00324*
	(0.00703)	(0.00700)	(0.00708)	(0.00667)	(0.00653)	(0.00183)
Early enrolled	0.0348***	-0.0605***	0.0609***	-0.0475***	0.0495***	0.391***
	(0.00308)	(0.00321)	(0.00301)	(0.00311)	(0.00268)	(0.0888)
old	-0.191***	-0.180***	-0.195***	-0.178***	-0.183***	-0.236***
	(0.00744)	(0.00822)	(0.00705)	(0.00788)	(0.00597)	(0.0358)
BFMJ	0.0273***	0.0144***				
	(0.00100)	(0.00102)				
BMMJ	0.0174***	0.0134***				
	(0.000804)	(0.000786)				
FISCED	0.0774***	0.0396***				
	(0.00108)	(0.000981)				
MISCED	0.0734***	0.0278***				
	(0.00109)	(0.000999)				
Internet dummy	0.0578***	0.0693***	0.0649***	0.0728***	0.0454***	0.0621***
	(0.00252)	(0.00256)	(0.00246)	(0.00246)	(0.00219)	(0.00223)
Computer Dummy	0.0339***	0.0273***	0.0332***	0.0270***	-0.00209	0.00857***
	(0.00193)	(0.00188)	(0.00189)	(0.00185)	(0.00171)	(0.00171)
Class share of males	-0.00198	0.00135	-0.00338**	0.00139	-0.00242	0.00143
	(0.00166)	(0.00143)	(0.00162)	(0.00139)	(0.00155)	(0.00132)
Class share of foreign students	-0.00548***	0.00115	-0.00754***	0.000462	-0.00749***	0.000794
	(0.00168)	(0.00151)	(0.00167)	(0.00149)	(0.00159)	(0.00143)
Class size	0.0178***	0.00185	0.0211***	0.00414**	0.0213***	0.00324*
	(0.00238)	(0.00199)	(0.00236)	(0.00191)	(0.00223)	(0.00183)
Wage per capita	-0.715***	0.431***	-0.615***	0.420***	-0.580***	0.391***
	(0.122)	(0.0891)	(0.123)	(0.0908)	(0.121)	(0.0888)
Earners share on population	0.493***	-0.227***	0.481***	-0.241***	0.497***	-0.236***
	(0.0607)	(0.0354)	(0.0583)	(0.0369)	(0.0632)	(0.0358)
Foreign share on population	-0.328***	0.206***	-0.306***	0.197***	-0.312***	0.209***
	(0.0408)	(0.0339)	(0.0393)	(0.0337)	(0.0409)	(0.0332)
High earners share on population	0.310***	-0.0955*	0.376***	-0.123**	0.393***	-0.135**
	(0.0976)	(0.0517)	(0.0890)	(0.0509)	(0.102)	(0.0546)
High incomes share on total income	0.246***	-0.130***	0.206***	-0.137***	0.187***	-0.117***
	(0.0720)	(0.0440)	(0.0689)	(0.0432)	(0.0715)	(0.0450)
						0.0167***
Constant	-0.642***	-0.449***	-0.328***	-0.309***	-0.0416***	-0.126***
	(0.0118)	(0.0101)	(0.0128)	(0.0107)	(0.00939)	(0.00784)
Observations	1,004,432	973,426	1,063,400	1,030,157	1,305,019	1,261,554
R-squared	0.438	0.447	0.422	0.445	0.420	0.440
V grade School FE	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
School variables	NO	NO	YES	YES	YES	YES
Peer effects	NO	NO	NO	NO	YES	YES
Municipaloty variables	NO	NO	NO	NO	NO	NO

A. 5: Students with PC and internet access in grade 5

VARIABLES	(1) Num.	(2) Num.	(3) Lit.	(4) Lit.	(5) Num.	(6) Lit.	(7) Num.	(8) Lit.
Broadband	0.0253	0.0298*	-0.0105	-0.00976	0.0310*	-0.00987	0.0266	-0.00886
	(0.0170)	(0.0168)	(0.0168)	(0.0168)	(0.0167)	(0.0168)	(0.0170)	(0.0168)
Test score - Grade 5	0.533***	1.015***	0.563***	0.867***	0.614***	0.565***	0.531***	0.561***
	(0.00533)	(0.0188)	(0.00348)	(0.0188)	(0.00525)	(0.00412)	(0.00533)	(0.00348)
(Test score - Grade 5)^2		-0.482***		-0.294***				
		(0.0201)		(0.0177)				
Male	0.0280***	0.0277***	0.0120***	0.0128***	0.0256***	0.0121***	0.0283***	0.0114***
	(0.00297)	(0.00295)	(0.00300)	(0.00299)	(0.00295)	(0.00300)	(0.00298)	(0.00300)
I Gen. migrant	-0.0653***	-0.0599***	-0.158***	-0.156***	-0.0585***	-0.158***	-0.0562***	-0.147***
	(0.0110)	(0.0109)	(0.0115)	(0.0115)	(0.0110)	(0.0116)	(0.0110)	(0.0116)
Nursery	-0.000857	-0.00122	-0.00476	-0.00472	-0.000946	-0.00476	-0.00242	-0.00612
	(0.00390)	(0.00387)	(0.00389)	(0.00387)	(0.00386)	(0.00389)	(0.00390)	(0.00388)
Early enrolled	0.0223***	0.0231***	-0.0594***	-0.0605***	0.0227***	-0.0594***	0.0218***	-0.0587***
	(0.00537)	(0.00537)	(0.00544)	(0.00545)	(0.00536)	(0.00544)	(0.00539)	(0.00545)
BFMJ	0.0290***	0.0281***	0.0171***	0.0166***	0.0289***	0.0171***	0.0141***	0.00481**
	(0.00174)	(0.00173)	(0.00174)	(0.00174)	(0.00173)	(0.00174)	(0.00199)	(0.00194)
BMMJ	0.0160***	0.0153***	0.0126***	0.0124***	0.0156***	0.0126***	0.0108***	0.00812***
	(0.00149)	(0.00148)	(0.00143)	(0.00142)	(0.00148)	(0.00143)	(0.00151)	(0.00144)
FISCED	0.0873***	0.0862***	0.0416***	0.0419***	0.0872***	0.0416***	0.0713***	0.0288***
	(0.00193)	(0.00190)	(0.00179)	(0.00178)	(0.00192)	(0.00178)	(0.00212)	(0.00197)
MISCED	0.0777***	0.0776***	0.0315***	0.0319***	0.0778***	0.0316***	0.0665***	0.0224***
	(0.00197)	(0.00195)	(0.00180)	(0.00180)	(0.00195)	(0.00180)	(0.00202)	(0.00190)
Broadband # Test score (Grade 5)					-0.159***	-0.00246		
					(0.00784)	(0.00600)		
ESCS- Grade 5							0.0353***	0.0454***
							(0.00325)	(0.00311)
ESCS- Grade 5#c.broadband							0.0237***	-0.0129***
							(0.00426)	(0.00411)
Constant	-0.654***	-0.644***	-0.190***	-0.194***	-0.641***	-0.190***	-0.507***	-0.0664
	(0.0618)	(0.0606)	(0.0485)	(0.0483)	(0.0612)	(0.0484)	(0.0622)	(0.0488)
Observations	332,932	332,932	324,139	324,139	332,932	324,139	332,538	323,758
R-squared	0.412	0.417	0.440	0.441	0.417	0.440	0.413	0.440
V grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
VIII Grade School FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
School variables	YES	YES	YES	YES	YES	YES	YES	YES
Peer effects	YES	YES	YES	YES	YES	YES	YES	YES