POLITECNICO DI TORINO

Collegio di Ingegneria Gestionale
Corso di Laurea Magistrale in Ingegneria Gestionale

Tesi di Laurea di II livello

The Impact of Broadband Penetration on Students' Performance: the Italian Case

Relatore:

Prof. Carlo Cambini

Correlatore:

Dott. Lorien Sabatino

Candidato:

Davide Dosio

Anno Accademico 2019-2020

Contents

List of figures

List of tables

1. Introduction

The general belief that broadband technologies may improve the educational outcomes of students and the quality of learning led to numerous investments by governments to provide fast Internet connection to schools. In May 2020, the Italian Ministry of Education approved a plan for the expenditure of €400 million, financed by European funds, to bring optical fiber to Italian schools. In addition to school-level investment and government subsidies, families spend a substantial amount of money to bring broadband connection at home. Therefore, in the last ten years, students have experienced continuous growth in connectivity, both at school and at home. How does this increase in connectivity and Internet access affect students' educational achievement? The existing literature investigating the impact of new Information and Communication Technologies (ICT) on students' performance at school find ambiguous or even negative results. On the one hand, high-speed Internet connection may provide real-time access to numerous information, encourage new learning methods that increase students' interest and make available new technologies in education such as virtual classrooms or e-learning. On the other hand, investments in broadband and ICT may substitute expenditures on more efficient educational inputs. Also, teachers may find it hard to integrate ICT in their traditional learning and students may use high-speed Internet to perform distracting activities such as playing video games or watching videos.

This thesis aims to provide further evidence on the impact of ultra-fast broadband on students' performance. The main data source for this study is a student-level dataset collecting language and math test scores for Italian eighth-grade students, over the period 2013-2017. Students' performance is measured on a standardized national test designed by INVALSI, the National Evaluation Institute for the School System. Also, the dataset provides rich information on students' characteristics (e.g. gender, age, birthplace) and family-background (e.g. parental education and employment). The INVALSI dataset is combined with a second panel collecting broadband diffusion data in every Italian SLL (Local Labour System).

The remainder of the thesis is organized as follows. Chapter 2 presents the principal broadband technologies and network architecture. Chapter 3 discusses the empirical evidence from the existing literature regarding the impact of ICTs on students' performance. Chapter 4 describes our dataset and provides summary statistics on the sample of students. Chapter 5 presents the econometric models used in the empirical analysis and our main findings. Lastly, Chapter 6 offers some concluding remarks.

2. Broadband Technologies

In telecommunications, the term *broadband* commonly refers to high-speed data transmission across long distances. In the context of Internet access, the term broadband includes any Internet access that is always on and faster than the traditional dial-up[1] connection. Internet access networks are defined as the last part of the public network which connects the users to the nearest central office of the service provider (this segment is also known as the "last mile"). Access networks with connection speed up to 20 Mbps are considered *broadband* (BB), while access networks with connection speed from 30 Mbps up to 100+ Mbps are called *ultra-fast broadband* (UBB). These networks are usually owned by the national telecom operators. For instance, in Italy, the broadband infrastructure is mainly owned by Telecom Italia. Broadband includes several high-speed technologies that differ in term of the medium which transports the signals: coaxial cable, optical fiber, twisted pair, or radio. In particular, it is possible to distinguish between *fixed networks* that typically refers to all of the wired networks, and wireless or *mobile networks* that exploit radio signals to transmit the information. Following this distinction, the principal broadband technologies are presented in the next sections.

2.1 Fixed access networks

Fixed networks consist of a group of devices that are connected to the Internet by wires, rather than by radio signals. Most of the fixed access networks are based on copper wires and exploit the pre-existing telecommunication infrastructure. The network is typically segmented in two levels, using cabinets in the streets and in the building basements (as shown in Fig. 2.1). The *primary network* is the segment from the central office to the street

[1] The dial-up connection use telephone line to connect PC to the Internet. In particular, dial-up access uses facilities of the public switched telephone network (PSTN) to establish a connection to an Internet Service Provider (ISP). It is characterized by analog data transmission and connection speed up to 56 Kbps. This connection is typical in remote areas where broadband and cable connections are rare.

cabinets and has distances up to 2-3 Km. The segment from the street cabinets to the residential users is the *distribution network* (or secondary network), with typical distances between 100-300 m.

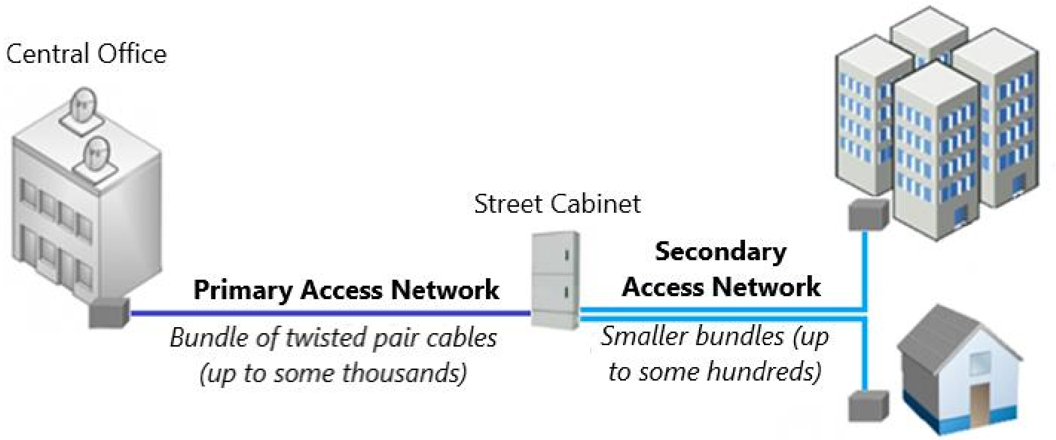


Fig. 2.1: Primary and secondary access network.

The *central office*, also called exchange, is the main switching facility for a telecom operator and handles Internet and telephone service for a determined locality. It connects subscriber home and business lines via a physical pair of wires, and it is the termination of long-distance links. The *street cabinets* are an intermediate small structure, usually located near the habitation of final users, that connects the subscriber line to the exchange point. In 2011, Telecom Italia fixed residential network had near 11.000 central offices, 150.000 street cabinet and 22 million residential users (in terms of apartments reached).

Fig. 2.2: Inside a central office (left-side picture) and a street cabinet (right-side picture).

However, modern access networks are increasingly using optical fiber to improve broadband speeds and service quality (Fig. 2.3). For instance, in Europe, governments and operators are moving to deploy new access solutions partially or totally based on fibers, according to the recent European Commission (EC) directives. In 2010, the EC launched the Digital Agenda for Europe[2] (DAE) allowing every European to access broadband by 2013 and ultra-fast broadband by 2020. The development of the Next Generation Access Networks[3] (NGA) constitutes a crucial step toward achieving this goal.

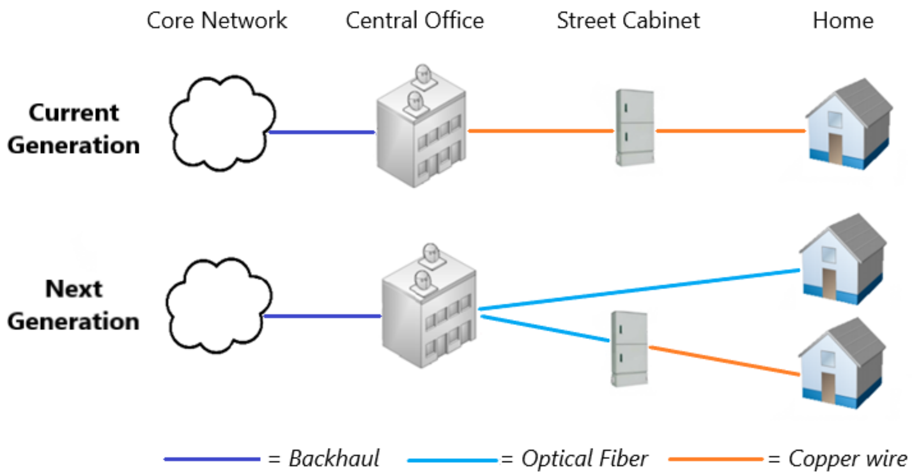


Fig. 2.3: Current generation and Next Generation Access (NGA).

The main broadband technologies based on wired connections are: Digital Subscriber Line (DSL), cable modem, and optical fiber. The choose of the broadband technology depends on several factors such as location (urban or rural area), price, availability, and bundling with other services (e.g. telephone or home entertainment).

2.1.1 Digital Subscriber Line (DSL)

Digital Subscriber Line (DSL) is a family of data communications technologies that enable digital data transmission over copper telephone lines. Figure 2.4 shows a typical DSL connection. On the customer side, several devices are connected to the DSL modem, which is hooked up to a telephone line. The DSL modem turns the digital signals generated by the connected devices into analog electrical signals to be carried over the line. Voice and data traffic are transmitted over a *dedicated* twisted pair that connects the single subscriber to the central office. The transmission of voice and data occurs at different frequencies to reduce interferences. On the telco side, the line is connected to a DSLAM, which concentrates and

---

[2] The Digital Agenda for Europe (DAE) is one of the seven flagship initiatives of the EC in the 10 years strategy "Europe 2020" for smart, sustainable, and inclusive growth. Further information is available at: http://eurlex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC0245R(01)&from=EN.

[3] The Next Generation Access Network (NGA) consists of a not fully copper-based access network, capable of providing broadband services with higher bandwidths than those available with fully copper-based access networks. It provides substantial improvements in broadband speeds and quality of service and it is often used to refer to networks using optical fiber technology.

collects data from a large number of individual DSL connections. Opposed to the DSL modem, the DISLAM turns the analog electric signals into digital signals and route the voice and data traffic to the telephone network and the Internet, respectively.

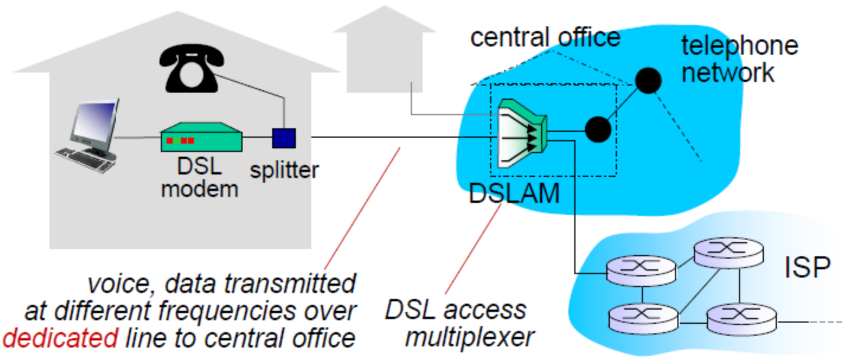


Fig. 2.4: Schematic DSL connection.

There are different types of DSL technologies for Internet access, but the most commonly installed in Europe is the *Asymmetric* DSL (ADSL). Starting around 2000, the ADSL introduced the first "broadband revolution" in access networks by allowing to carry up a few Mbps over the "old" twisted pair, without any cable replacement. ADSL differs from the less common *Symmetric* DSL (SDSL) because data can flow faster in one direction than the other. ADSL is typically faster in the downstream (from the network to the user) than the upstream direction. For this reason, it is used primarily by residential customers downloading content from the Internet. The SDSL is mainly used by business subscribers for services that require significant connection speed, both upstream and downstream (e.g. video conferencing).

DSL technologies have numerous advantages: (i) DSL Internet is faster than the traditional dial-up connection and provides connection speeds between 2 Mbps up to 50 Mbps, depending on the technology; (ii) unlike dial-up connection, DSL technologies allow the use of broadband Internet and phone at the same time, and more devices can be connected to the same DSL line; (iii) DSL completely reuse the existing copper infrastructure, reducing the investment costs and speeding up its diffusion. However, DSL technologies have some limitations in term of bandwidth and performance. Connection speed depends on the proximity between the subscriber and the provider's central office: the farther the subscriber is from the central office, the more the connection speed reduces. This effect is due to the attenuation of the signal power while running along the copper wires. The available bit rate also depends on interference phenomena from several sources (e.g. twisted pairs in the same bundle or proximity to radio transmitters). As shown in Table 2.1, the maximum range for

DSL technologies without a repeater is 5.5 km and, as the distance from the central office decreases, the data rate increases.

Very high bit-rate DSL2 (VDSL2) is the fastest of all DSL technologies and provides data rates up to 100 Mbps downstream for distances of less than 1.2 km from the central office. VDSL2 reaches high connection speeds using optical fiber between the central office and the street cabinet (this solution is also called *fiber-to-the-cabinet* FTTC).

Type	*Download speed*	*Upload speed*	*Optimal distance from CO*	*Ratified*
HDSL	< 1.5 Mbps	< 1.5 Mbps	< 3.6 Km	1994
SDSL	< 1.5 Mbps	< 1.5 Mbps	< 6.7 Km	1999
ADSL	1.5 to 8 Mbps	16 to 640 Kbps	< 5.5 Km	1999
SHDSL	< 2.3 Mbps	< 2.3 Mbps	< 3.0 Km	2001
ADSL2	8 to 12 Mbps	< 1.0 Mbps	< 3.0 Km	2002
ADSL2+	12 to 24 Mbps	< 1.0 Mbps	< 2.5 Km	2003
VDSL	20 to 50 Mbps	< 20 Mbps	< 1.6 Km	2004
VDSL2	20 to 100 Mbps	< 30 Mbps	< 1.2 Km	2005

Table 2.1: DSL technologies and performance development.

2.1.2 Cable modem

The cable network is the most common access network in the United States, supporting both cable Internet and cable television services. As shown in Fig. 2.5, the cable modem enables cable operators to provide broadband Internet using the same coaxial cable that delivers pictures and sound for television services. In contrast to DSL, where each twisted pair is dedicated to a single subscriber, the coaxial cable is *shared* between many subscribers. The coaxial cable concentrates the data traffic of different subscribers to the cable headend, where the CMTS routes the traffic to the Internet.

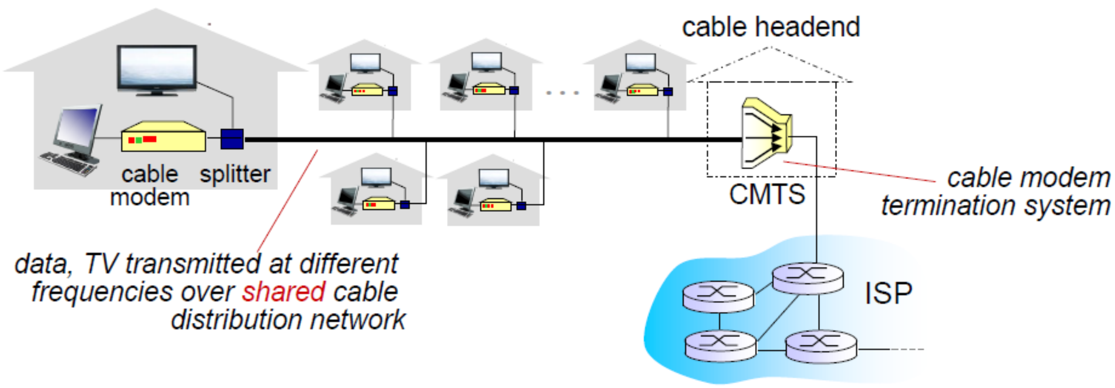


Fig. 2.5: Schematic cable modem connection

The coaxial cable is characterized by four different layers (Fig. 2.6): (i) a centre core conductor based on a copper wire which data and video travel through; (ii) a dielectric plastic insulator surrounding the copper wire; (iii) a metallic layer that shield the cable from electromagnetic interference; (iv) an external plastic layer which protects the internal layer from damage. Unlike the most common copper wires, the coaxial cable can carry high-frequency electrical signals with low losses and interferences, supporting download speed up to 50 Mbps. For these reasons, cable modem provides faster download and upload Internet speed as compared to DSL connections. However, cable speed depends on the traffic load, as the total available network bandwidth is shared among different users: the more subscribers access simultaneously to the Internet, the more the connection speed decreases for the individual user. Furthermore, unlike the dedicated DSL connection, a damage to the coaxial cable leads to a service loss for all connected users.

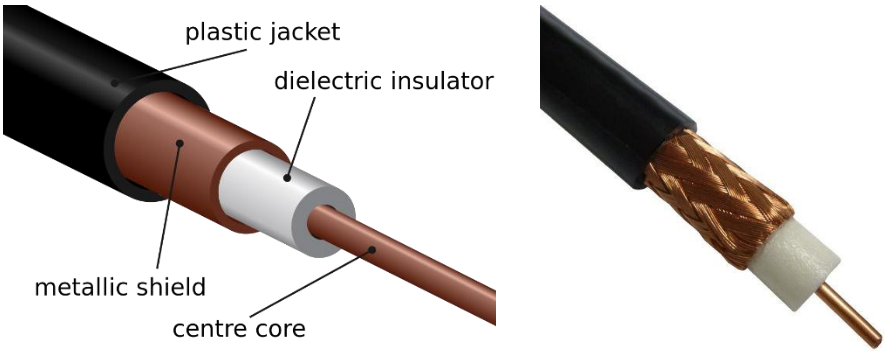


Fig. 2.6: Coaxial cable cutaway (left-side) and a real one (right-side).

2.1.3 Optical access networks

Access networks may also use optical fiber as a communication medium instead of copper wires. Fiber-optic communication is based on the transmission of information by sending pulses of infrared light along a flexible and transparent fiber, made by drawing glass or plastic. Optical fiber is slightly thicker than a human hair (0.25 to 0.5 mm) and a single cable can contain a varying number of optical fibers, from a few up to a couple hundred. The fiber consists of two glass layers (Fig. 2.7): a core surrounded by a cladding material with a lower refraction index. The difference in the refraction index between core and cladding keeps the light inside the core, thanks to the total internal reflection.

Optical fiber has large advantages over the existing copper wire when high bandwidth and long-distance is required. The infrared light propagates through the fiber with lower attenuation than electrical transmission, allowing longer distances between amplifiers or repeaters. Also, due to inherent glass characteristics, optical fiber has the capacity for greater bandwidth, allowing higher data rate transmissions. Besides, fiber is also immune to electromagnetic interference, a problem from which metal wires suffer. For these reasons, fiber-optic communication systems have primarily been installed in long-distance applications (up to dozens Tbps over thousands Km) such as backhaul network or submarine communication cable[4]. In short distance and relatively low bandwidth applications, the electrical transmission is preferred because of its lower cost and the exploitation of pre-existing telecommunication infrastructures.

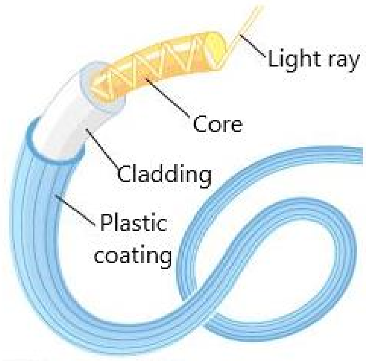


Fig. 2.7: Optical fiber cutaway (left-side) and an Optical Fiber Cable (right-side)

However, in recent years, optical fiber also started to be used in access networks due to the increasing demand for high-speed broadband Internet. As fiber cables can carry much more data than copper wires, the main idea is to use fiber as close as possible to the final user to increase connection performances. In general, depending on how close fiber is to the user, it is possible to identify three optical solutions in access networks (Fig. 2.8): *fiber-to-the-cabinet* (FTTC), *fiber-to-the-building* (FTTB), and *fiber-to-the-home* (FTTH).

- FTTC uses a full optical fiber connection from the central office to the street cabinet. From there, the optical link connects to the existing copper line to deliver broadband service to the home. This solution is regarded as a hybrid solution between fiber and DSL: the exploitation of existing street cabinets and infrastructure makes FTTC a lot cheaper to install than other FTTx solutions. However, using a form of DSL connection between the cabinet and the customer leads to far lower speeds than full fiber connections.

[4] There are dozens of optical fiber cables crossing each ocean to connect continents between them. The first transatlantic optical cable was put into service in 1988 between U.S., France, and Great Britain. By 2002, an intercontinental submarine network of 250.000 km with a capacity of 2.56 Tbps was completed.

- In FTTB solution, the fiber cable directly reaches the boundary of the building, such as the basement or the communication room. From there, an active device connects the optical link to the existing copper network to provide the final connection to each apartment or office. In contrast to FTTC, FTTB connection eliminate the slower copper cable infrastructure between the customer and the provider, reaching bandwidth up to 1 Gbps in download and 400 Mbps in upload.
- Lastly, FTTH refers to the deployment of optical fiber from the central office directly into the home. This solution replaces the existing copper wires within the building and provides connection speeds between 1 Gbps up to 10 Gbps. FTTH solution completely solves any reasonable bandwidth limitation. However, it requires a complete re-cabling of the residential access network and a new optical modem in each apartment. The main problem is the cost: between 500€ and 800€ per user are required to reach a new user in "true" FTTH, most of which related to digging.

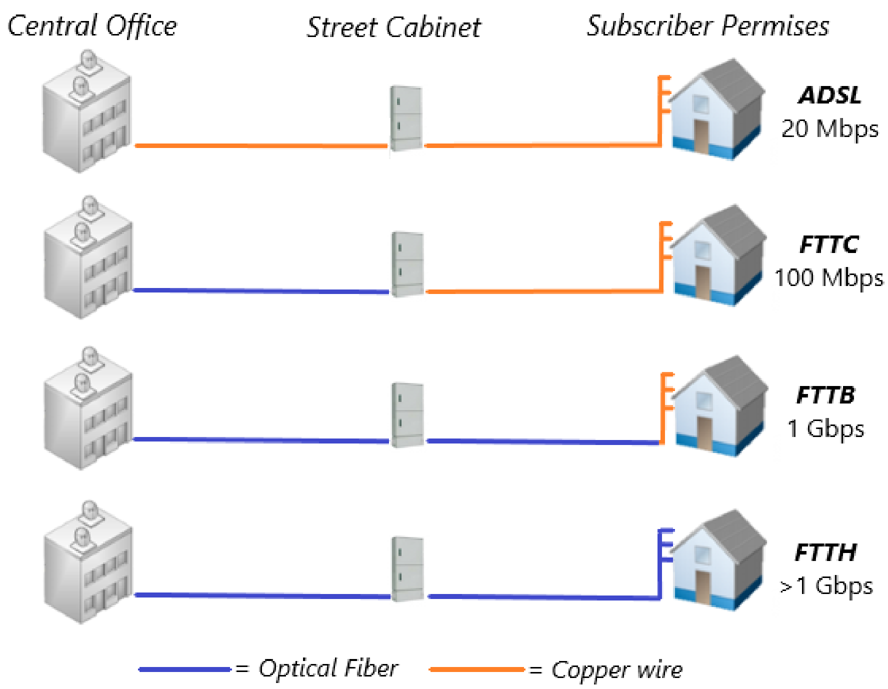


Fig. 2.8: Optical access network solutions (FTTx).

In general, there are two types of architecture for fiber optic networks (Fig. 2.9): *Point-to-Point* (P2P) and *Passive Optical Networking* (PON).

- The P2P architecture has a dedicated fiber pair from the central office to each residential user or Optical Network Termination (ONT). The core switch is at the central office and each port directly connects to an ONT.

- The PON is a point-to-multipoint infrastructure: the traffic split from one fiber into many using passive optical splitters. A single fiber runs from the Optical Line Terminal (OLT) located at the central office to the passive optical splitter, where the transmission line split off into separate lines. In some situation, there may be two levels of splitting: first in the street cabinet and then in the building basement.

Unlike the dedicated P2P infrastructure, the PON architecture shares a small number of fibers across a set of subscribers. The PON is less expensive and quicker to implement than P2P due to lower fiber investments. However, the PON has a limited level of bandwidth and it is quite difficult to update when the bandwidth requirements change. Usually, "true" FTTH deployment is based on a PON infrastructure, while intermediate FTTC and FTTB solutions are based on a P2P architecture.

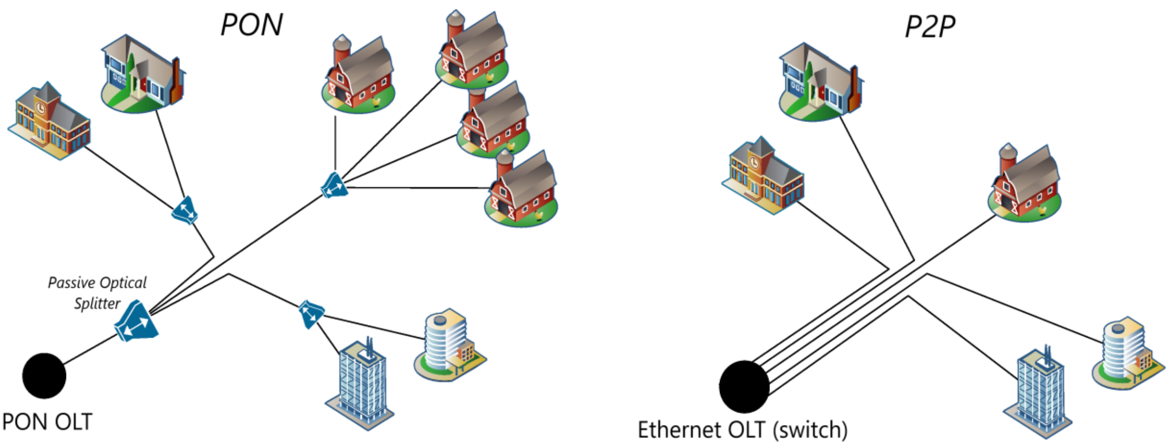


Fig. 2.9: Passive Optical Networking (left-side) and Point-to-Point (right-side) architecture.

2.2 Wireless access networks

Wireless networks use radio frequencies in air to transmit and receive data, instead of physical cables. The main benefits of wireless networks are: (i) increased nodes mobility, allowing them to move in space without getting disconnected from the network; (ii) faster and easier installation, avoiding the costly cabling process; (iii) broadcast transmission with a wider reach, also in places not accessible for wires and cables; (iv) flexibility in network change to meet new configurations. In general, wireless networks can be classified according to their architecture, coverage, and node mobility.

There are two basic architectures for wireless networks: *infrastructure-based* and *Device-to-device* (D2D). The infrastructure-based network requires the use of infrastructure devices to facilitate communication between nodes (e.g. base stations in cellular networks or access points in wireless local networks). All nodes communicate directly with the infrastructure node, which is a gateway towards the wired network. Individual devices cannot communicate directly with each other, but only indirectly through the infrastructure node. Cellular and Wi-Fi are the most important examples of wireless networks with infrastructure. In contrast, the D2D network consists of direct communication between devices, without using infrastructures such as access points or base stations. Devices nearby communicate using direct link rather than sending a radio signal through an infrastructure node, resulting in low latency due to a shorter signal path. For many years D2D networks have been of interest mainly for niche applications, particularly for the emerging Internet of Things (IoT)[5]. However, D2D communication is expected to play a significant role in upcoming cellular networks. Besides, there are also hybrid architectures combining the two basics. As shown in Fig. 2.10, hybrid networks combine rich interconnection among devices with some gateway nodes connected to the Internet (e.g. base stations).

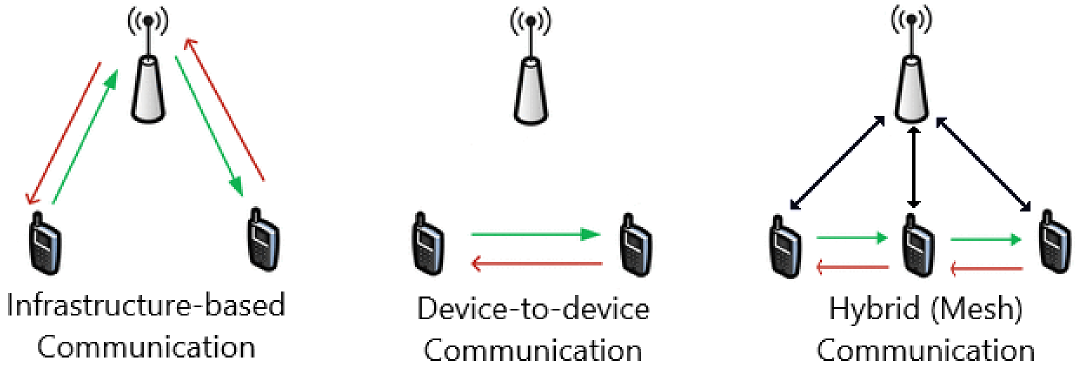


Fig. 2.10: Illustration of Infrastructure-based, Device-to-device and hybrid communication.

Wireless network coverage is defined as the width to which the wireless signal is transmitted, usually in term of geographical area. In term of coverage, there are large differences between wireless solutions.

- The wireless solution with the highest coverage is the *Satellite access network*. Geostationary satellites, positioned 36.000 Km from the earth surface, provide broadband connection with download speeds up to 500 Kbps. Satellite networks are typically used for downstream broadcasting, covering several nations with their signal (range of thousands Km).

[5] Internet of Things (IoT) refers to a system of interrelated, Internet-connected objects and devices that can collect and transfer data over a wireless network. The physical objects are embedded with sensors, software, and other technologies to connect and exchange data with other devices over the Internet.

- Reducing network coverage, we find *cellular* or *mobile networks*. Cellular networks obtain geographic coverage by splitting large areas into smaller “cells”, each served by a base station. In cities, each cell may have a range up to 2 Km, while in rural areas the cell sizes usually range from 1 Km to 20 Km.
- Finally, the *Wireless Local Network* (WLAN), commonly known as Wi-Fi, allows users to move around a coverage area such as home or small office while maintaining the connection. The Wi-Fi router connects to the cable modem and serves as a base station, providing Internet access to the connected devices. Wi-Fi routers may reach up to 50 m indoors and up to 100 m outdoors, depending on the operating frequency.

2.2.1 Cellular networks

A cellular or mobile network is a communication network that use radio frequencies to transmit the information. Cellular networks are infrastructure-based and rely on a set of fixed-location transceivers, known as base stations. The base station provides network coverage for the transmission of voice, data, and other types of content. Each base station covers a delimited area called “cell”, and route communication to and from user’s terminals located in that area. Users can move freely from one cell to another, hooking up to the base station with the strong signal. Ideally, cells should have a regular hexagonal shape. However, cells shape and size depend on many factors such as the type of base station, the morphology of the geographical area (e.g. urban, mountain, and so on), and propagation conditions. Also, base stations have limited transmission capacity, and can only handle a certain number of simultaneously connected users. For this reason, cells tend to be numerous and small in urban areas, where the population density is high. Similarly, cells size is much larger in rural areas, where the population density is lower.

Since the radio frequencies are finite, the main advantage of the geographical division in cells is the reuse of spectrum resources in non-adjacent cells. The radio spectrum refers to the full frequency range from 3 kHz to 300 GHz that may be used for wireless communication. As shown in Fig. 2.11, the same set of frequencies is reused in non-adjacent cells, avoiding interferences and providing guaranteed bandwidth within each cell. The wireless spectrum is an extremely precious resource: the governments use an auction system to sell the rights to transmit signals over specific bands of the spectrum. In 2018, the Italian government completed its sale of wireless spectrum to telecom operators for future 5G services, raising 6.55 billion euros.

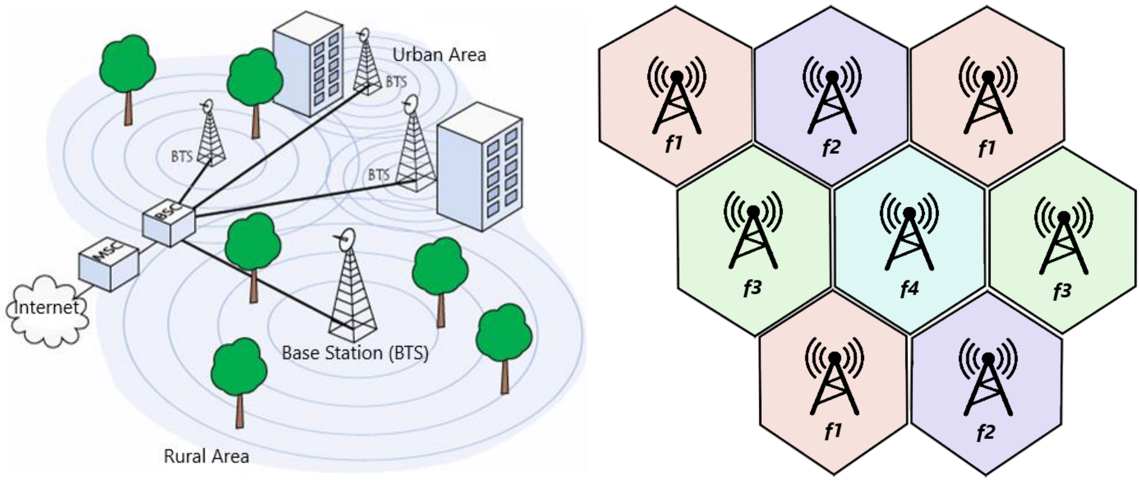


Fig. 2.11: Cellular network topology (right-side) and a conventional network with frequency reusing (left-side).

Cellular network has evolved through various generations since the early 1970s. The first generation of mobile networks, known as 1G, was launched in 1980 providing voice service. However, 1G technology suffered from several drawbacks: coverage was poor and characterized by large cell clusters, service quality was low, calls were not encrypted, and there was no compatibility between different systems. The 1G technology was replaced in the early 1990s by the second generation (2G). The main difference between the two cellular systems is that 1G used analog radio signal, while 2G was the first digitally based generation. The 2G solved some limitation of its predecessor introducing smaller cells and cell clusters, digitally encrypted calls, and higher sound quality. Also, 2G allowed for the first time services such as text messaging (SMS) and multimedia messaging (MMS). The first 2G upgrade (called 2.5G) implemented a packet-switching domain, allowing for data service in addition to the voice service. The third generation (3G), launched in 2001, required large investments and was characterized by higher data transfer capabilities: 3G was four-time faster than 2G. This led to the rise of new services such as mobile TV, Internet access, video conferencing, and video streaming. The fourth generation (4G) is the current standard around the globe and succeeded the 3G in 2009. The 4G technology builds upon what 3G offers but it is characterized by faster speed and lower latency. 4G cellular networks performance allow services such as high-definition video, real-time streaming, gaming, and IP telephony. 4G technology is rapidly becoming a mature market. In Italy, the 4G coverage (% homes) was around 99% in 2019. Also, the 4G availability (% time) raised from 79% in 2019 to 89,6%

in 2020[6]. The maturity of 4G technology pushed telecom operators to focus on the development of the new generation standard for broadband cellular networks. 5G is currently under development and the technology is likely to appear by the early 2020s. 5G promises significant improvements over 4G: faster data rates, higher connection density, and much lower latency. Also, some of the plans for 5G include device-to-device communication. The increased speed is achieved by using higher-frequency radio waves than current cellular networks. However, high-frequency waves have a shorter range and so require smaller cells. The superior connectivity offered by 5G has the great potential to transform industries and business models, enable new applications, and improve quality of life. 5G is the foundation for realizing the full potential of IoT, connecting and managing a growing number of devices. Besides, 5G will boost innovation across many industries, from banking to healthcare, allowing innovations such as remote surgery, autonomous driving, smart homes, and so on.

Generation	***Period***	***Frequency***	***Data rate***	***Characteristics***
1G	1980s	150-900 MHz	< 2.4 Kbps	First wireless commun.
2G	1990s	0.9-1.8 GHz	14.4 - 384 Kbps	Digital signals
3G	2000s	1.6-2.0 GHz	144 Kbps - 2 Mbps	Digital broadband
4G	2010s	2.0-8.0 GHz	100 Mbps - 1 Gbps	High speed
5G	(2020s)	3.0-300 GHz	> 1 Gbps	-

Table 2.2: Evolution of mobile technologies and their performance.

[6] 4G coverage data are taken from the Digital Economy and Society Index (DESI) report in 2019. 4G availability data are taken from OpenSignal.com measures in May 2019 and May 2020.

3. Literature Review

3.1 Impact of broadband technologies on economic development

Broadband Internet fulfill many characteristics of a *General Purpose Technology* (GPT). This notion is used extensively to discuss the role of technical change in economic growth, such technologies may have considerable impact upon the economy. According to Bresnahan and Trajtenberg[1] (1995), a technology is considered "general purpose" if follows three key characteristics: (i) pervasiveness, (ii) inherent potential for technical improvement, and (iii) innovational complementarities. GPT are those technologies characterized by the potential of pervasive use in a wide range of different sectors and, as the technology evolves and spread throughout the economy, fostering generalized productivity gains. Furthermore, most GPTs play the role of enabling technologies, opening up new opportunities for innovation. The deployment of broadband infrastructure yields both direct and indirect positive effects on economic activities. Direct investments in broadband infrastructure increase employment and economic activities necessary for creating and expanding the network, while indirect benefits derive from productivity gains associated to the adoption of the technology. High-speed Internet promote economic growth and productivity by accelerating the distribution of large batches of data and ideas, facilitating decentralized information processing, and lowering the storage cost of data. Moreover, cheaper information dissemination facilitates the adoption of new technologies and accelerate the development of new products, processes, and business model. This results in more efficient business processes and innovation acceleration. On the demand side, broadband internet increase market transparency and intensify competition.

The expected benefits of this technology led to numerous investments for the development of the infrastructure in the last decade. In 2010, the European Commission (EC) launched

[1] Bresnahan, T., & Trajtenberg, M. (1995). General purpose technologies: 'Engines of growth'? Journal of Econometrics, 65, 83–108.

the Digital Agenda for Europe[2] (DAE) which specifies the targets to be achieved by 2020 in terms of next generation access networks: (i) universal availability of at least 30Mbps connection speed, (ii) 50 percent or more of European households with at least 100 Mbps connection speed, and (iii) high-speed broadband in industrial areas. The EC estimated an investment between 38 and 58 billion euros needed to achieve 30Mbps coverage for all by 2020, and between 181 and 268 billion euros to provide 100 Mbps services to half of the European households. The importance of this technology and the size of the investments has attracted growing attention from both policy makers and academic researchers. There are many empirical studies focusing on the impact of broadband on different economic outcomes. In particular, we can distinguish the studies in two categories: (i) macro-studies assess the causal impact of broadband deployment at the country level on macroeconomic measures, such as GDP or employment; (ii) micro-studies focus on the causal impact of broadband on within-country economic outcomes, such as local employment or firm-level productivity. The following sections present some empirical studies that explore these areas.

3.1.1 Effects on GDP

Czernich et al. (2011) evaluates the impact of broadband infrastructure on economic growth using an annual panel of 25 OECD[3] countries in 1996-2007. The variation in broadband penetration in each country is exploited to measure its impact on per capita GDP growth. However, broadband penetration may be endogenous to the economic growth process: (i) individuals in high-income countries have more financial resources to adopt broadband services, resulting in more rapid broadband penetration; (ii) economically more developed countries can allocate more state funds for the development of broadband infrastructure; (iii) broadband diffusion could be related to the diffusion of other technologies like mobile phones and computers, making it difficult to isolate its effect. To address these problems, the authors develop an instrumental-variable (IV) approach approximating the actual values of the broadband penetration rate with the predicted values of a diffusion model. The diffusion model is based on the fact that broadband standards (e.g. ADSL, VDSL) rely on

[2] The Digital Agenda for Europe (DAE) is one of the seven flagship initiatives of the EC in the 10 years strategy "Europe 2020" for smart, sustainable, and inclusive growth. Further information is available at: http://eurlex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC0245R(01)&from=EN.

[3] The Organization for Economic Co-operation and Development (OECD) was founded in 1961 with the main aim of stimulating economic progress and trade. Currently, the organization has 34 countries as members on 4 continents in the world.

the pre-existing telephony network: the extent of the infrastructure determine the "ceiling" of the broadband penetration in each country. The existing networks only affect the supply side, excluding the demand-side effects previously mentioned. The authors find a significant positive effect of broadband introduction and penetration on economic growth. The IV results show that, after a country has introduced broadband, GDP per capita is 2.7 to 3.9 percent higher on average than before its introduction. In terms of diffusion, an increase in broadband penetration rate by 10 percent-point leads to a 0.9-1.5 percent-point annual growth in per-capita GDP.

Rohman and Bohlin (2012) focus on the impact of broadband speed on economic growth in 34 OECD countries. The macroeconomic data on GDP over the period 2008-2010 were collected from OECD databases, while the speed data were collected from Ookla, a company that provides Internet diagnostic services. Broadband speed is strongly related with broadband development. In order to avoid the endogeneity problem of speed data, the authors adopt a two-stage least-squares. In the first stage, the broadband speed is estimated as a function of some independent variables: broadband penetration rate, broadband subscription price, urban density, proportion of the urban population and telecom revenue. The IV is employed in the second stage to estimate the causal effect on GDP per capita. The control variables in the econometric model are urban density and growth, urban population, and labor force. Fig. 3.1 shows a positive relationship between broadband speed and GDP per capita.

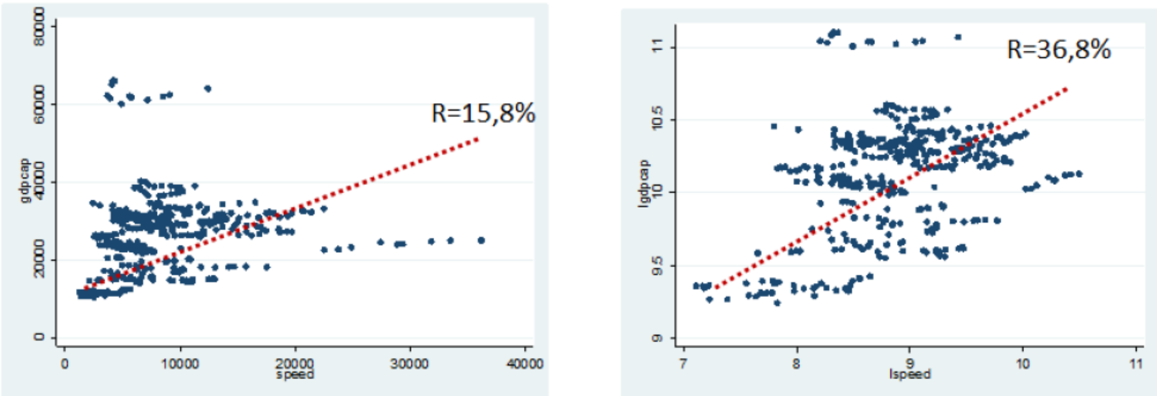


Fig. 3.1: Relationship between broadband speed and GDP per capita level (left side) and growth (right side)

The results of the analysis show that the hypothetical impact of broadband speed on economic growth is statistically significant: doubling the speed leads to an increase of 0.3 percent-point in GDP growth. However, the size of the impact largely depends on the existing economic growth in each country and the size of the coefficient of speed.

The studies presented here suggest that both the adoption and development of broadband technology have an important role for the economy, explaining large investments in recent years.

3.1.2 Effects on employment and labor market

Forman et al. (2012) study the effect of firm-level Advanced Internet Investment[4] on local wages growth. The database contains rich firm-level data over the period 1995-2000, including number of employees, personal computers per employee and use of Internet applications, for establishment located in different counties in the US. In order to solve the endogeneity problem of AII on county-level wages, the authors use measures of local telecommunications infrastructure costs, local industry, and the programming capabilities of related locations as instrument for local Internet investment. The preliminary results show a statistically significant positive correlation between AII and local wage growth. However, this relationship is primarily found in only about 6 percent of US counties (163 out of 2580). Those counties are characterized by high income, large population, high skills, and large number of IT-intensive firms. In top counties, advanced Internet led to an increase in wage growth of about 23 percent-point, while in other counties the wage growth increased by only 1 percent (Fig. 3.2).

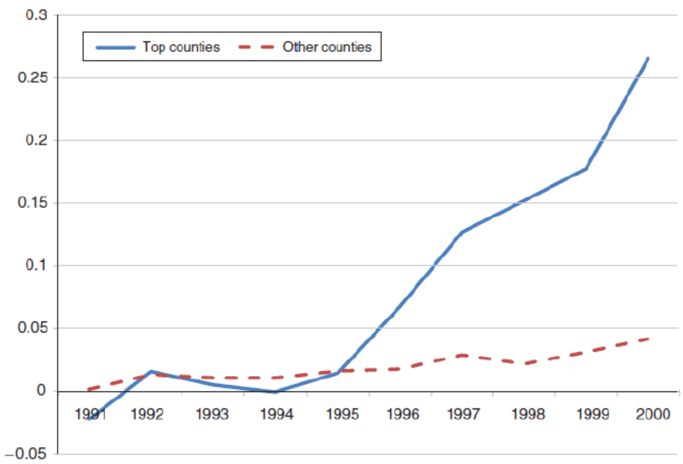


Fig. 3.2: Year-by-year marginal effect of AII on county-level wage growth

In short, while Internet investment is widespread, the payoffs are not: only high-income and high-educated counties are able to receive the benefits of these investments, increasing

[4] Advanced Internet Investment (AII) refers to substantial investments in e-commerce or e-business applications. These investments enabled productivity advances and required skilled labor to implement and operate.

regional wage and income inequality. Empirical evidence from this study run against policies subsidizing Internet infrastructure in low density location: infrastructure growth has little impact without appropriate supply of skilled labor.

Akerman et. al (2015) examine how broadband Internet adoption affects labor productivity and wages of different types of workers, exploiting the broadband diffusion in Norway over the period 2001-2007. The dataset contains detailed firm-level information on productivity, municipality-level data on wages, and information about employee length of education. The authors define an employee as “skilled” if he or she has a university degree, while individuals with lower titles are considered “unskilled”. The broadband diffusion in Norway is the result of a public program called National Broadband Policy[5]. The authors consider the roll-out of broadband availability as exogenous as it seems not to depend on the interaction between supply and demand, the evidence presented suggest non-correlation with key observable correlates of labor outcomes and productivity. The results show that the availability and adoption of broadband internet by firms increase wages and employment of skilled workers, while decrease wages of unskilled individuals (Fig. 3.3). In terms of labor productivity, a 10 percent-point increase in broadband availability raises output by 0.4 percent for given inputs.

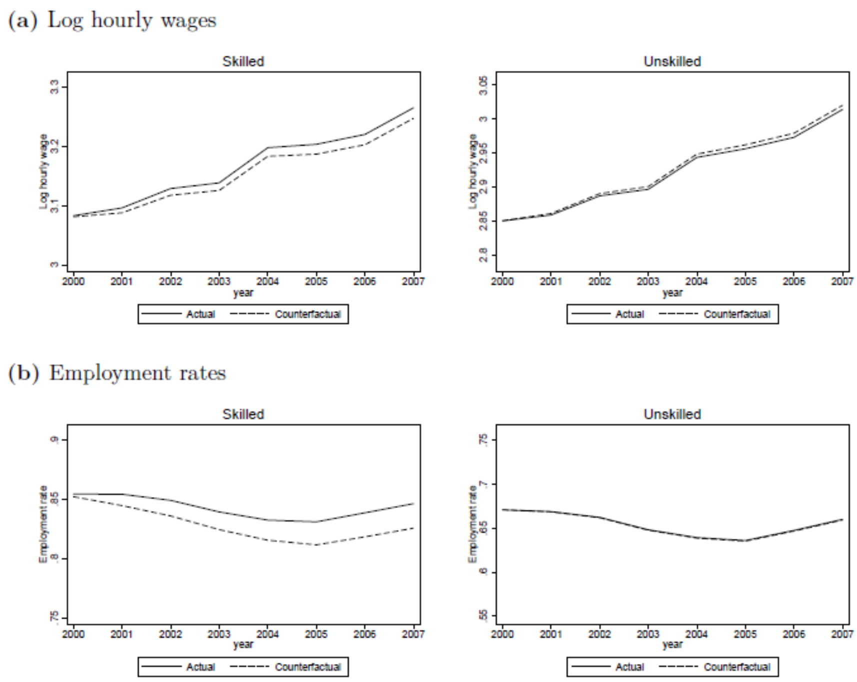


Fig. 3.3: Comparison between actual and predicted counterfactual time trends for wages and employment

[5] The National Broadband Policy was introduced by the Norwegian Parliament in the late 1990s. The aim of the public program was to ensure supply of broadband internet to every area of the country at uniform price, and to ensure that public sector quickly adopted broadband internet.

This finding suggests skill complementarity of broadband adoption by firms: broadband internet complements skilled workers in executing problem-solving and information-intensive tasks, while substitutes unskilled workers in performing routine tasks (more amenable to automatization). This led to what is called *skill bias* of ICT[6]: high skilled workers appropriate all benefits from internet technologies. Taken together, these findings suggest the need for public intervention to redistribute benefits.

Bai (2016) examines the relationship between different broadband speed levels and US county-level employment rate for the period 2011-2014. The research distinguishes three broadband categories: ultra-fast broadband with download speed higher than 1Gb/s, fast broadband characterized by download speed between 100MB/s and 1Gb/s, and normal broadband with download speed lower than 100Mb/s. The results of the first-differenced model show a positive correlation between the availability of all three types of broadband and the changes in the county employment rate. In particular, the positive impact of ultra-fast broadband is lower than that of fast broadband, while the effect of fast broadband is greater than the one of normal broadband, suggesting a diminishing return in the employment rate as speed increases.

In conclusion, broadband technologies not only have a direct positive effect on employment and labor market due to the development and construction of the infrastructure, but also have an indirect positive effect due to the adoption of the technology. However, while the impact tends to be positive on macroeconomic variables, the effects on local economies are not clearly defined. Broadband diffusion seems to increase inequality between urban and rural areas, and between high- and low-skilled workers.

3.2 Impact of ICT technologies on student performance

In recent years, schools have experienced a large infusion of new information and communication technologies. ICT investment in education is mainly divided into three areas: (i) investment in computer hardware, (ii) investment in instructional computer software, and (iii) investment in Internet connections. Figure 3.4 shows how the OECD

[6] Information and Communication Technologies (ICT) refers to technologies that provide access to information through telecommunications. This includes the Internet, wireless networks, cell phones and other communication mediums.

average computers-student ratio raised from 0.13 in 2000 to 0.77 in 2015[7]. The increase in the computers-student ratio is evidence of substantial investment in school ICT resources.

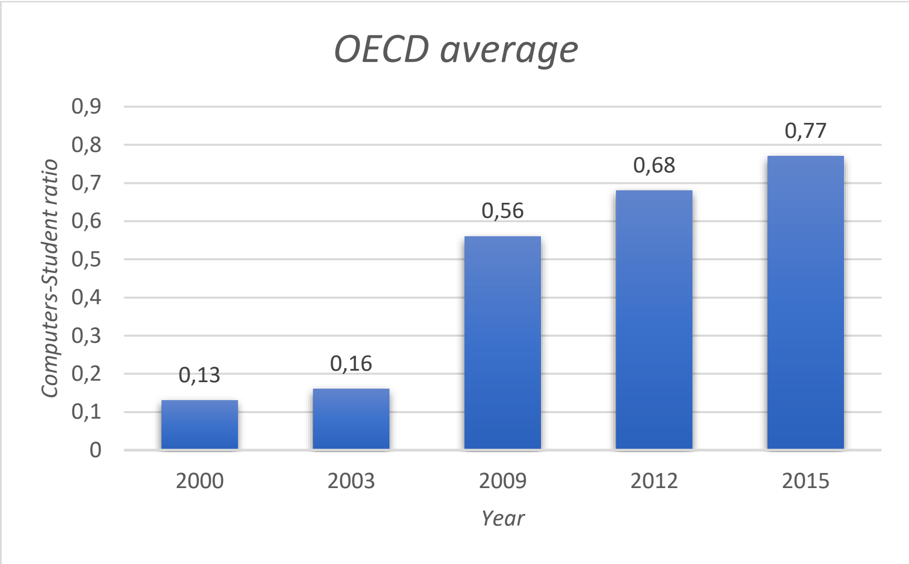


Fig. 3.4:OECD average computers-student ratio in schools from 2000 to 2015. (Source: PISA reports)

Figure 3.5 shows both the increase in the number of computers and the proportion of instructional computers connected with the Internet in U.S. public schools: in 1995 only 8% of computers had Internet access, while in 2008 almost every computer had internet access (98%)[8], suggesting that it has become the norm in schools.

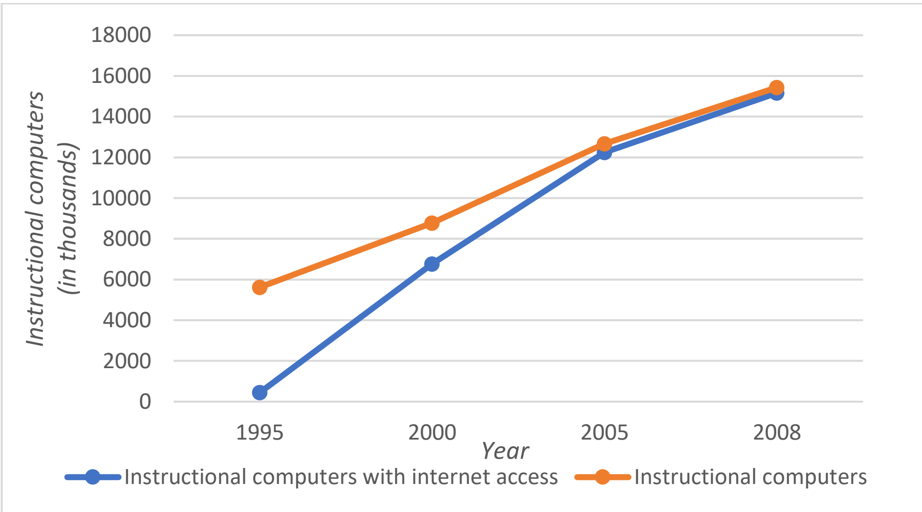


Fig. 3.5: Number and Internet access of instructional computers in U.S. public schools 1995-2008. (Source: U.S. National Center for Educational Statistics)

[7] Program for International Student Assessment (PISA), 2000, 2003, 2006, 2009, 2012, and 2015.
[8] U.S. Department of Education, National Center for Education Statistics, Fast Response Survey System (FRSS), *Internet Access in U.S. Public Schools and Classrooms: 1994-2005* and *Educational Technology in U.S. Public Schools: Fall 2008*.

In addition to school-level investment and government subsidies in ICT, also families spend a substantial amount of money on computers, software, and Internet connection: current level of access to home computers and Internet connections among students are very high (Fig. 3.6).

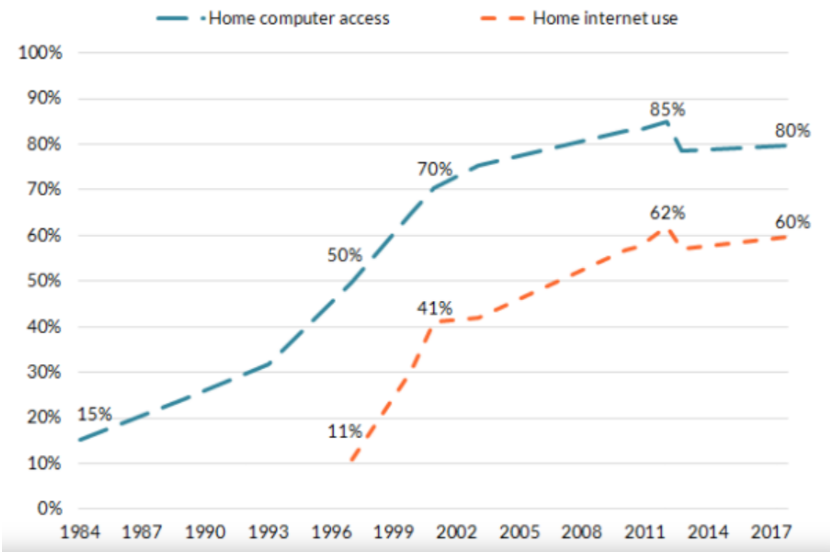


Fig. 3.6: Percentage of children ages 3 to 17 who have access to computer and use Internet at home in U.S. 1984-2015. (Source: U.S.Census Bureau)

The growth in ICT investments to improve the quality of education has attracted the attention of both policy makers and economists. Among policy makers, it is widely believed that ICT investment plays a major role in raising educational standards, while economists tend to have the opposite view. Hypotheses are suggesting that computers may both further and hinder student learning. Therefore, the expected net effect may depend on factors supporting the positive or negative effects. Computers may help students in their learning process by making education available at home and less dependent on teachers' quality. As well, the Internet help students exploiting enormous information possibilities for school purpose. However, computers and the Internet can be distracting, reducing the time spent on doing homework or learning. Furthermore, investment in ICT technology necessarily offsets investment in traditional inputs that may be more or less efficient, and time allocated to using technology may displace traditional classrooms activities. These tradeoffs imply that theoretical expectations of the effects of ICT are ambiguous, so a better understanding of how ICT technology affect educational outcomes is critical. Next sections present a general overview of the existing literature regarding the effects of computers, the Internet, and software such as Computer-aided Instruction (CAI) on educational outcomes.

3.2.1 Impact of Computers

Using a computer may affect economic outcomes in at least two ways: (i) computers skills may have direct effects on productivity and wages; (ii) computers can be used for learning other skills, such as math or science[9], that have a positive impact on labor-market outcomes. For these reasons, governments made large investments in school computerization to improve educational outcomes. The studies presented here try to find evidence on the effectiveness of government intervention and computers on students' performance. In particular, the literature focuses on two primary contexts in which computers may be used for instruction: (i) classroom use in schools, and (ii) home use by students.

Angrist and Lavy (2002) exploit the large-scale computerization effort funded by the Israeli State lottery in 1994 to estimate the effect of the new technology on pupils' test scores. The funds were part of a program called Tomorrow-98[10] to computerize the Israeli education system. The authors use scores from a test conducted by the National Institute for Testing and Evaluation (NITE) and carried out by 4th- and 8th-grade students in June 1996. Among the 200 sampled schools that participated in the test, only 122 applied for the program. The results of the 2SLS estimates show that, despite the program increased the use of computers in schools, there is no evidence that this translated in higher test scores. Results for fourth-graders show lower math scores in the group that was awarded computers: an increase in intensity of CAI by one unit reduces math test scores by 7-9 points. Fourth-grade verbal scores and eighth-grade math scores are also lower in the program group, though these differences are not significant.

Fuchs and Woessmann (2004) estimate the relationship between the availability and use of computers and students' educational achievement. The authors use student-level micro-data of the PISA[11] dataset, an international student achievement test conducted in 2000 by the OECD. The PISA database offers information about availability and use of computers both at home and at school, it also offers information on students' family-background and schools characteristics. The first bivariate analysis shows a positive correlation between student

[9] Borghans and Weel (2004) find that math and writing performance yield positive returns on the labor market. Similarly, other studies suggest that math achievement is related to productivity (e.g. Bishop 1992).

[10] In 1994, the Israeli State lottery used lottery profits to sponsor various social programs, including the Tomorrow-98 program. Between 1994 and 1996, the first three years of the program, 35.000 computers were installed in 905 schools.

[11] The Programme for International Student Assessment (PISA) tested 15-year-old student performance in reading, math, and science in 32 developed and emerging countries in 2000. Further information on the study at: www.pisa.oecd.org

achievement and availability of computers both at home and a school. However, students with better-educated parents or parent with high-paying jobs tend to perform better in terms of educational achievement and tend to have more computers at home. Also, the availability of school computers is strongly correlated with the availability of other school resources that play a role in student performance. So, there is an *identification problem*: computers effect can easily be confounded by effects of other factors. The omission of these variables from the empirical equation will bias the estimated effect of computers. To overcome the problem, the authors directly control for family-background and school characteristics in a multivariate analysis. The multivariate analysis shows a negative effect for the availability of home computers and insignificant effect for the availability of school computers. The authors also study the effect of the actual use of computers as a communicational and educational device. They find a positive relationship between students' performance and the use of computers at home for emailing, webpage access and educational software. Thus, the mere availability of computers at home seems to distract students from learning, while using computers in constructive ways have a positive effect. At schools, the relationship between use of computers and students achievement shows an inverted U-shape: students who never use computers at school show lower performance than students who sometimes use computers, but students who use them several time a week perform even lower. A possible explanation, also found in Angrist and Lavy (2002), is that computer use might decrease student learning. The pattern suggests the existence of an optimal level of computers use at school.

A recent trend in educational policy is to ensure that every student has his or her laptop or tablet computer. However, there is concern that these programs may not yield sufficiently learning outcomes to justify the large investment. In this sense, Grimes and Warschauer (2008) examine the impact of the implementation of a one-to-one laptop program in three diverse schools in California. The authors compare performance on the California Standard Test (CST[12]) between the three schools in the laptop program and their peers in the same district without laptops. The 2004 CST test, before the laptop program, provided a baseline for score changes in the first two years of the program (2005 and 2006). The results (Fig. 3.7) show that language scores of laptop student declined relative to non-laptop students in the first year and recovered in the second year, resulting in no significant differences over

[12] The California Standard Test (CST) measures the performance of students undergoing primary and secondary education in California, from second to eighth grades. CST show how well students are doing in relation to the state standards in both math and language.

the two years. Math scores of laptop students improved relative to their non-laptop peers both years. However, this effect cannot completely be attributed to the laptop program because laptops were used less in mathematics than in language. The initial decline in language scores may have been due to the complexity of introducing a fundamental change in learning, rather than a disadvantage of using laptops.

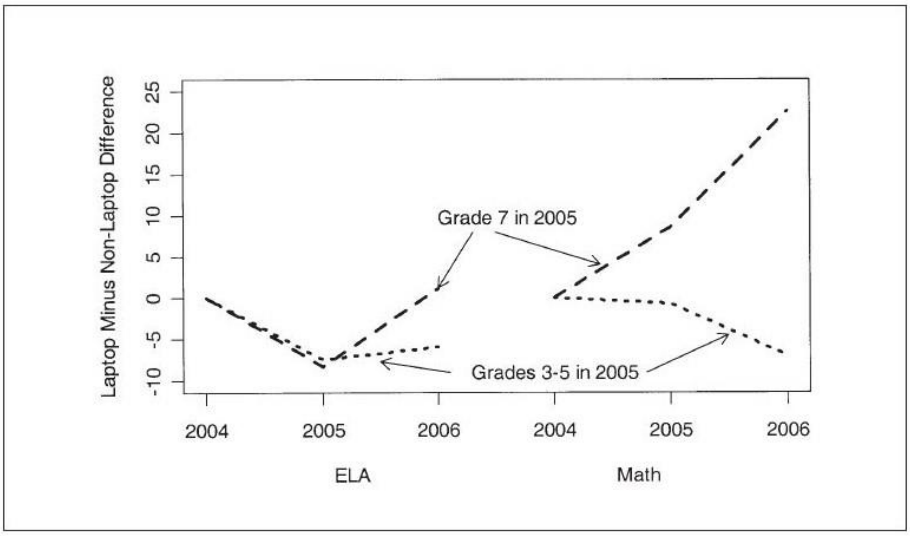


Fig. 3.7: Differences between scores of laptop and non-laptop students.

Barrera-Osorio and Linden (2009) exploit a large-scale computerization program in Colombia to evaluate its effect on student performance. The “Computers for Education” program was created in 2002 to integrate computers, donated by the private sector, into the teaching in public schools. The authors conduct a two-year randomized evaluation of the program using a sample of 97 schools and 5201 children from third to eleventh grade. In this process, 48 schools received the program (treatment group) and 49 schools were assigned to a control group. They assess the causal effect of the program by directly comparing the average responses of the national Colombian exam (Saber) in the treatment and control groups. The results show that the program successfully increased the number of computers in the schools and also increased students’ use of computers. However, the program had little impact on students’ math and Spanish test scores. The main reason for these results may be the implementation of the program: surveys of both teachers and students suggest that teachers failed to incorporate the new technology into their classroom teaching.

Fairlie and Robinson (2013) provide evidence of the educational impact of home computers. The authors conduct a large field randomized experiment with 1123 students from sixth- to tenth-grade in 15 Californian schools. Half of the student were randomly selected to receive a home computer, while the other half served as a control group. The schools provided detailed administrative data on students, including grades and scores of a standardized test from the STAR[13] program, both pre- and post-treatment. The experiment finds that computer ownership increased the total hours of computers use: children in the treatment group spend additional hours on schoolwork, but also on games, social networking, and other forms of entertainment, compared with the control group. However, the results show no evidence of an effect on educational outcomes, including both grades and standardized test scores.

In conclusion, the presented evidence focusing on the impact of computers and programs for the computerization of school yield mixed results at best on educational outcomes. Expanding the supply of computers in schools alone may not lead to actual benefits from the use of new technology. A careful implementation plan is required, and the implementation of the technology may take a long time to be effective. Home computers as well seem to be ineffective in raising student performance: computer ownership alone is unlikely to have an impact on short-term outcomes.

3.2.2 Impact of Computer-aided Instruction (CAI)

Computer-aided instruction is defined as the use of specific software programs on computers to teach concept and skills in classrooms. The studies presented here aim to evaluate whether and how computer-aided instruction is more effective than traditional teaching.

Rouse and Krueger (2004) assess the impact of an instructional computer program, known as Fast ForWord, on student performance. Fast ForWord is a group of computer software programs used in U.S. public schools and designed to improve language and reading skills. The study was conducted in an U.S. urban school district, with a total of over 20.000 students. The treatment group is composed of students randomly selected to participate in the FFW program, while the rest of the students comprised the control group. At the end of the program, the authors evaluated the effect of FFW on four tests designed to reflect both language and reading skills, having both pre- and post-test scores for each outcome. Using the Diff-in-Diff for the four test outcomes, the authors estimate large and statistically test score gains for the treatment group. However, the student in the control group posted nearly identical gains during the same period, resulting in a small and statistically insignificant

[13] The California Standardized Testing and Reporting (STAR) program is a group of four standardized tests that cover mathematics, science, and language arts. All students in grades 2-11 participate in the STAR program.

effect of FFW on student's language outcomes. Using IV models, the authors estimate statistically significant test score gains in only one of the four tests. In conclusion, the results suggest that the use of the FFW program does not translate in actual language acquisition or reading skills.

Banerjee et al. (2007) test the effect of a computer-assisted learning program in India from 2002 to 2004. The program offered two hours of computer instruction per week for fourth-grade children, during which they play games involving math problems. The program was allocated using random assignment across a set of schools: the final sample for the study consisted in 55 schools that received the instruction program (treatment group) and 56 schools as the comparison group. All the student took a pre-test at the beginning of the school year and a post-test at the end of the same. The results of the analysis show that, in the post-test, the math scores are significantly greater in the treatment schools than in comparison schools: math scores increased by 0.35 standard deviations the first year, and 0.47 the second year. Contrary to Angrist and Levy (2002) and Rouse and Krueger (2004), this study suggests that the insignificant effects of CAI in developed countries may not hold in developing countries. In developed countries CAI replaces time spent in well-equipped classrooms with high quality instructors, while in developing countries computers may replace teachers with less motivation and training.

Barrow et al. (2009) evaluate the impact of a well-defined instructional computer program, known as I Can Learn, designed to improve pre-algebra and algebra skills. The study was conducted in three large US urban school districts during the 2003-2004 school year and included 17 schools, 152 classes and 3541 students. The authors implemented a within-school random assignment at the classroom level: they randomly selected classroom of students to be taught using CAI (treatment group) or using traditional teaching (control group). The test was designed by the Northwest Evaluation Association (NWEA) and was customized to target specific algebra skills outlined in each districts' course objectives. The students took the NWEA pre-test at the beginning of the school year and the post-test at the end of the program. Results show a positive effect of CAI on students' achievement in math: students who are randomly assigned to classes using the instructional program score at least 0.17 of a standard deviation higher than students assigned to traditional classrooms. They also find evidence that the effects are larger in large size classrooms. Traditional classrooms have high student heterogeneity, the lessons are covered at the same pace, and teachers have less time for individualized instruction with each student. CAI instruction is completely individualized, and students can move at their own pace, resulting in higher effectiveness for some students.

In conclusion, the evidence from economic studies is mixed and suggest that the characteristics of the intervention are important. The effects of CAI differ in the extent to which it is a substitute or a supplement to traditional instruction. Also, the effect depends on the grade of development of a country, the quality of traditional teaching, and the teaching subject. CAI seems to be more effective in math teaching and in developing countries, where computers may replace to some extent low-quality teaching. However, these studies focus on well-defined instruction programs and not all CAI software may be equally effective. The presented results suggest that CAI deserves additional rigorous evaluation and policy attention, since it may be much easier for schools to implement than other interventions.

3.2.3 Impact of Broadband Internet

Broadband Internet, as well as other ICT technologies, is perceived as a potential tool to improve quality in education and student's performance. High-speed Internet connection may provide real-time access to numerous information, foster new learning methods, increase students' interest, and enable the adoption of new technologies in education such as virtual classrooms or e-learning. For these reasons, many governments announced national programs to make broadband Internet available in schools. However, high-speed Internet access may also have negative effects: teachers may find it hard to adapt traditional learning with ICTs and students may use broadband to perform distracting activities such as playing video games or chatting. The studies presented in this section try to evaluate whether broadband Internet improve or worst student performance.

Goolsbee and Guryan (2006) exploit the U.S. E-Rate[14] program to evaluate the effect of government subsidy on Internet investment by schools, and whether an increase in Internet access improves student performance. The study covers the period from 1996 to 2000 and uses data from Californian schools. First, the authors focus on the effectiveness of the E-Rate program and find that the subsidy significantly increases Internet investment: a ten-percent-point increase in the E-Rate subsidy leads to an increase in the growth rate of Internet

[14] The E-Rate program began in 1998 and provides up to $2.25 billion per year of subsidies to school Investment in Internet and communication technology. The aim of the program is to get classrooms connected to the Internet, particularly at disadvantaged school to reduce the digital divide.

access of 0.78 classrooms per year. Second, they present estimates of the effect of the subsidy on student test scores. Student achievement was measured with the Stanford Achievement Test[15], given every year to each public school student in California, starting from 1997-1998 school year. The results show no evidence that the increased rate of Internet connectivity in schools has any measurable effect on student achievement, in any subject tested (math, reading and science).

Belo et al. (2014) evaluate the impact on students' performance of the actual usage of broadband Internet in schools. The authors collect broadband usage and school performance data from more than 900 schools in Portugal. Student achievement is measured by the school's average score at the national exam for 9th- and 12th-grade students, while broadband usage is evaluated as the sum of upload and download traffic exchanged between the school and the Internet. Fig. 3.8 shows an increase in Internet usage in schools since the introduction of ADSL in 2005. It also shows that average exam scores increased from 2005, suggesting a positive impact of Internet on students' achievement.

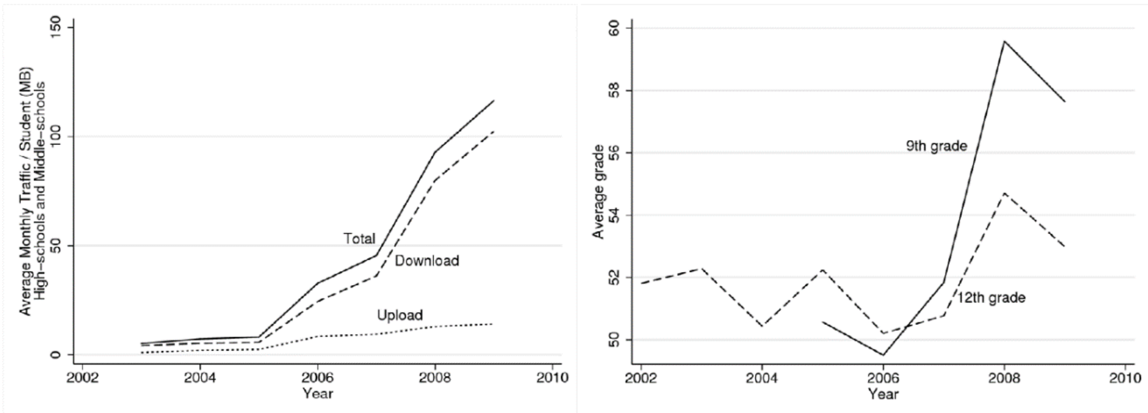


Fig.3.8: Internet traffic evolution in schools (left side) and average exam scores evolution (right side), 2002-2009.

The simple OLS estimation shows a positive and statistically significant relationship between Internet usage and test scores. However, potential unobserved factors might have influenced both Internet usage and test scores during the period of analysis (e.g. new school resources might have raised both test scores and Internet use), leading to a biased estimate. To overcome the identification problem, the authors use Internet connection quality as an instrument for broadband Internet use, assuming that schools with better connection are more likely to use the Internet. They use the distance between the school and the ISP's central

[15] The Stanford Achievement Test is a set of standardized tests used by the school district in the U.S. The test is available in 13 levels, one per grade, and measures students' achievement in various subjects, such as reading, math, and science.

office as a proxy for connection quality, exploiting the fact that ADSL lose performance as distance increases. The 2SLS estimates show a strong negative effect of broadband use on 9th-grade students' test scores: the test score declines by about 8.9% for schools using a high level of broadband. This negative effect seems to affect boys rather than girls: boys tend to perform more distracting activities on the Internet, such as playing online games or watching videos. From the study also emerges that schools with the worst performance before the introduction of broadband in 2005 suffer the most.

Silva et al. (2014) assess the impact of broadband availability in Brazilian schools on students' academic performance. The authors exploit the "Broadband at Schools" program, announced by the Brazilian government in 2008, aiming to bring broadband access to all urban public schools. Students' test scores are collected from the INEP[16] database that contains the average language and math standardized scores of the 9th-grade students in 2007, 2009, and 2011, for all public schools. The standardized test scores are put together to evaluate the performance at the school level. They also use data from the National Telecommunications Agency (ANATEL) to identify schools that participate in the program and the broadband adoption date in each school. Considering the schools' participation in the program as exogenous, the authors find a small but positive relationship between broadband availability and schools' performance: an increase in the number of months the school participates in the program leads to an improvement in the school average standardized scores, both in language and math.

Hazlett et al. (2016) evaluate the impact of E-Rate subsidies on students' achievement in North Carolina high schools, from 2000 to 2013. Differently from Goolsbee and Guryan (2006), the authors first focus on the effect of increasing E-rate spending per student, then they focus on how the number of students per Internet-connected computer affects test scores. Students achievement is measured using school-level data from SAT[17] scores. As shown in Fig. 3.9, the authors find a small negative correlation between E-Rate funding and test scores: a one percent increase in E-Rate spending per student decreases the average math score for a school by .0278%. This effect persists also looking at the number of students per

[16] The National Institute for Educational Studies and Research "Anìsio Teixeira" (INEP) is responsible for assessing education in Brazil. Also, it is accountable for educational statistics that help developing educational policies by the government.

[17] The Scholastic Assessment Test (SAT) is a standardized test, often used for college admissions in the United States. The test evaluates students in four subjects: reading, writing, language, and math.

connected computer: increasing the ratio between students and Internet-connected computers by one percentage point worsts SAT math scores by 6.82 percent.

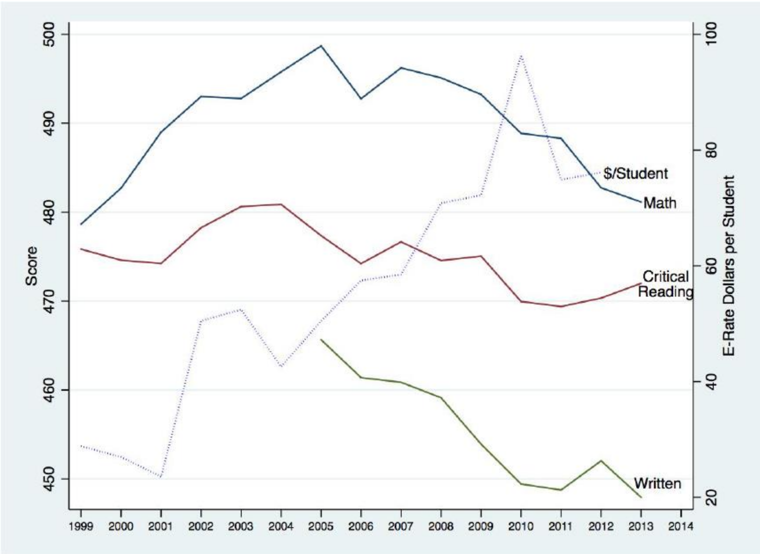


Fig. 3.9: North Carolina average SAT scores and E-Rate funding per student 1999-2014.

Grimes and Townsend (2017) estimate the effect of ultra-fast broadband on school-level performance in New Zealand. The New Zealand government announced the "Ultra-Fast Broadband Initiative" in 2009 to develop a fiber broadband network with school connection prioritized. By 2016, fiber broadband had been available to almost all state schools. The authors exploit differences in the exogenous timing of ultra-fast broadband adoption in schools to identify its causal impact on student achievement. School performance was tested using New Zealand's National Standards[18] to find the proportion of student at or above the national standard in math, reading and writing. The results find a small but positive relationship between the availability of fiber broadband and school performance: the availability of ultra-fast broadband increases National Standards passing rates by about one percentage point per year, in each subject. In contrast to Belo et al. (2014), they find larger benefits in low-socioeconomic schools and no evidence of gender benefit differences. They also find no differences in effect sizes among ethnic groups or between urban and rural schools.

Tajuddin and Rohman (2018) examine the relationship between the broadband penetration rate in Malaysia and state-level educational outcomes in primary and secondary school. The authors use broadband penetration rates data for all states supplied by the Malaysian

[18] In 2010, the New Zealand government introduced a policy of National Standards in reading, writing, and mathematics. Each student is measured against standards of expected competency for their age and graded as above or below the standard for each subject.

Communication and Multimedia Commission (MCMC), while students test scores in national examinations are collected from the Ministry of Education. The correlation analysis shows a very low correlation between broadband penetration and primary school test scores, but a greater correlation with secondary school scores (Fig. 3.10). The negative linear relationship in Fig. 3.10 (right side) is because lower grading values are linked with better performance. This evidence may find an explanation in the fact that subjects in primary school are more simplified and do not require ICT based learning. However, these results are preliminary and must be taken carefully: the existence of a simple correlation does not identify the causal relationship between the two variables.

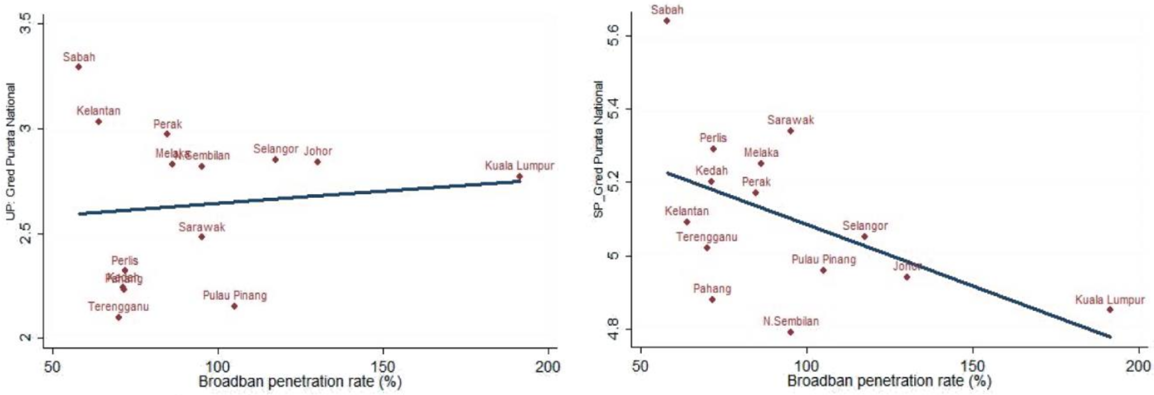


Fig. 3.10: Correlation between broadband penetration rate and primary (left side)/ secondary (right side) school test scores.

In conclusion, empirical evidence yields ambiguous results: relatively small positive or negative effects of broadband on educational outcomes. The government expenditures to introduce broadband access in schools do not necessarily contribute to increasing students' performance. Broadband development alone seems to be ineffective, suggesting the needs to join it with complementary interventions to support its use in beneficial ways. Anyway, broadband may still be beneficial for students in alternative ways not related to scores on standardized tests. The analyzed studies suggest a positive effect of broadband in developing countries rather than developed countries (similar to CAI case), and that teachers and schools need some time to get used with this technology to fully benefit from it.

Table 3.1: Overview: studies regarding the impact of technology on student performance.

	Test score data	*Other data*	*Main results*	*Methodology*
Computers				
Angrist and Levy (2002)	NITE test in June 1996. 200 sampled Jewish schools (only 122 applied for the program).	Tomorrow-98 implementation schedules and computer infrastructure in schools.	Lower Math scores for 4th graders in the group that was awarded computers. (-)	Diff-in-Diff, IV approach
Fuchs and Woessmann (2004)	Student-level micro-data from PISA database.	/	Availability: negative effect for home computers, insignificant for school computers. (-) Use: positive for home computers, inverted U-shape relationship for school computers. (+)	Cross-section, bivariate and multivariate analysis
Grimes and Warschauer (2008)	California Standard Test (CST) from 2004 to 2006.	Surveys and interviews at students, teachers, and principals of the schools for qualitative analysis.	Laptop students failed to keep up with non-laptop student in the first year (-) but made gains in the second year. (+)	T-tests
Barrera-Osorio and Linden (2009)	Saber Test. 97 sampled Colombian schools (only 48 received the program).	Questionnaires administered to students, teachers, and schools. Student-, teacher- and school-level information.	The program seems to have little effect on students' math and Spanish test scores. (0)	RCT, Diff-in-Diff
Fairlie and Robinson (2013)	Administrative school data on educational outcomes (e.g. grades). Standardized test from the STAR program.	Survey with additional information on students' and households' characteristics.	No evidence of an effect of home computer ownership on educational outcomes. (0)	RCT
CAI				
Rouse and Krueger (2004)	Pre- and post-test scores. 4 standardized language tests submitted to several schools in a U.S. district.	Student characteristics, student participation at the program (e.g. days of training).	Instructional computer program improves some aspects of students' language skills, but it does not translate in actual language acquisition or reading skills. (0)	RCT, IV approach
Banerjee et al. (2007)	Pre- and post-test scores. 111 sampled schools (only 55 received the instruction program).	/	The program increased math scores by 0.36 standard deviations the first year, and 0.54 the second year. (+)	RCT, Diff-in-Diff
Barrow et al. (2009)	Pre- and post-test scores. Customized NWEA test. 154 sampled schools.	Administrative data on student (e.g. sex and race/ethnicity) and student attendance/engagement.	Students using the instructional program score at least 0.17 standard deviations higher that students assigned to traditional classrooms. (+)	RCT, IV approach

Table 3.1: Overview: studies regarding the impact of technology on student performance.

	Test score data	*Other data*	*Main results*	*Methodology*
Broadband				
Goolsbee and Guryan (2006)	Stanford Achievement Test (SAT) in Californian schools.	Data on the technology owned by schools. Application data from the E-Rate program. Demographic data for each school.	The increase in Internet access had no measurable impact on student achievement. (0)	IV approach
Belo et al. (2014)	Anonymous student-level data from 9^{th}- and 12^{th}-grade national exam in the period 2005-2009.	School download and upload traffic data. Regional data (e.g. population density, average earnings).	The more broadband use is detrimental for 9^{th}-grade students' test score. This adverse effect is reinforced for boys and for school with lower performance before the broadband introduction. (-)	IV approach
Silva et al. (2014)	INEP dataset contains average Portuguese and math standardized scores of 9^{th}-grade students in years: 2007, 2009, 2011.	Three questionnaires collecting information regarding schools, teachers, and students. ANATEL data on schools participating in the program and broadband date of adoption.	The participation of the schools in the program slightly improves both mathematics and Portuguese test scores. (+)	System GMM
Hazlett et al. (2016)	School-level data on SAT scores from 2000 to 2013, for all public North Carolina high schools.	Data on E-Rate expenditures from USAC for each school or school district.	No gain in student test scores associated with E-Rate subsidy levels. Small negative effect on SAT scores. (-)	IV approach
Grimes and Townsend (2017)	New Zealand's National Standards test scores from 2011 to 2015.	Data on fiber broadband access in schools. Demographic data for each school.	The availability of ultra-fast broadband in schools increases the test's passing rates by one percentage point per year. (+)	Diff-in-Diff
Tajuddin and Rohman (2018)	Test scores for national examinations in primary and secondary school.	Broadband penetration rates per 100 inhabitants' data for all states.	Positive correlation between broadband penetration rate and secondary school educational outcomes (+).	Correlation analysis

4. Dataset Description

According to the existing literature, the impact of the broadband on students' performance is evaluated using two main data sources: standardized test scores and broadband penetration rates. In this study, I use as main data source a rich student-level dataset collecting Italian and Math scores for students enrolled in the grade eight, over the period 2013-2017. Students' performance is assessed with a standardized test designed by INVALSI, the National Evaluation Institute for the School System. The INVALSI national test is administered annually to students from second-grade up to high-school, and assesses learning achievement in Italian, Math, and English. Also, the test contains a questionnaire collecting family-background characteristics (e.g. parental education and work status), and student's characteristics (e.g. gender, age, birthplace). The INVALSI dataset is combined with a second panel collecting broadband diffusion data in every Italian SLL (Local Labour System). In particular, the dataset contains the coverage percentages of several broadband technologies such as ADSL, fiber-based UBB, and 4G mobile networks. The dataset also includes SLL-level structural data (e.g. population density, GDP per capita).

In numerical terms, the dataset contains 2.529.261 observations from 6.167 different schools in all the 20 Italian administrative regions. Table 4.1 summarizes how the observations are distributed over the period under consideration. In conclusion, this chapter aims to describe and summarize, through descriptive tables or charts, the data used in the estimation models.

Table 4.1: Distribution of observations over the period 2013-2017.

Years	Observations	%
2013	509.318	20,14%
2014	525.468	20,78%
2015	520.172	20,57%
2016	524.428	20,73%
2017	449.875	17,79%
Tot	2.529.261	100,00%

4.1 Descriptive statistics for demographic variables

The following descriptive statistics help to better characterize the sample of students considered in this study. This section aims to describe data on students' family-background and students' characteristics, collected through the INVALSI questionnaire. This section focuses on describing the following demographic variables: (i) students' gender; (ii) kindergarten frequency; (iii) students' and parents' birthplace; (iv) parents' educational background, distinguishing between high- and low-educated; (v) parents' job situation, distinguishing between employed high- and low-skilled and unemployed.

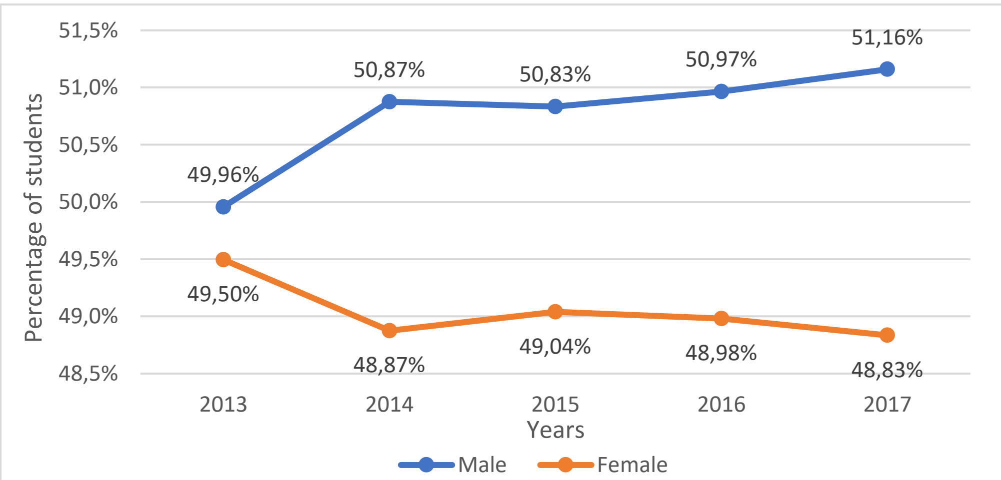
4.1.1 Students' gender

Table 4.2 presents frequency data on students' gender over the period 2013-2017. In general, both genders are equally represented over the sample, with a small prevalence of male students in each year. Furthermore, as shown in Fig. 4.1, the proportion of male student increased by about one percentage point, from 49,96% in 2013 to 51,16% in 2017. As opposed, the percentage of female student slightly decreased from 49,50% to 48,83%.

Table 4.2: Students gender frequency table over the period 2013-2017.

Years	Male	%	Female	%	Missing	%
2013	254.441	49,96%	252.089	49,50%	2.788	0,55%
2014	267.330	50,87%	256.819	48,87%	1.319	0,25%
2015	264.424	50,83%	255.092	49,04%	656	0,13%
2016	267.277	50,97%	256.869	48,98%	282	0,05%
2017	230.153	51,16%	219.693	48,83%	29	0,01%
Overall	1.283.625	50,75%	1.240.562	49,05%	5.074	0,20%

Figure 4.1: Percentage of students by gender from 2013 to 2017.



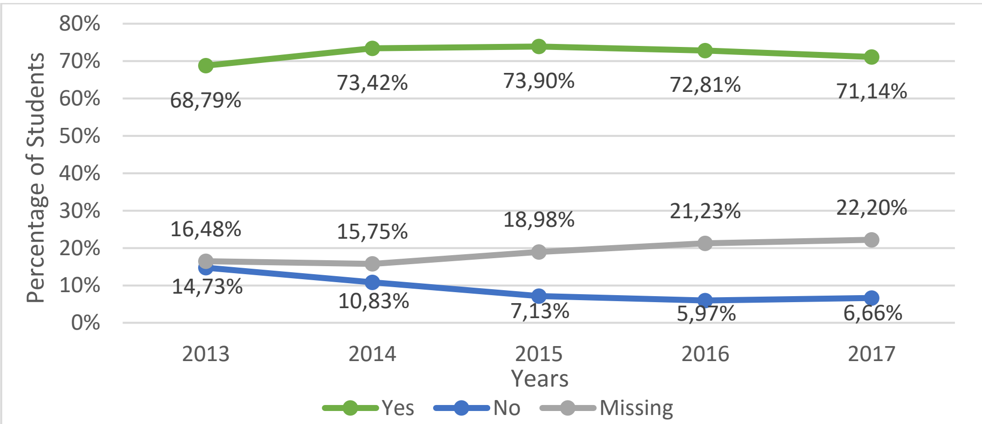
4.1.2 Kindergarten frequency

The second variable concerns the kindergarten attendance. Currently, kindergarten is not compulsory in Italy, although the government planned to introduce it by 2023. The description, summarized in Table 4.3, shows a strong prevalence of students who attended kindergarten, and a net reduction of students without preschool frequency over the years. As shown in Fig. 4.2, the percentage of attending students raised from 68,79% in 2013 to 71,14% in 2017, reaching a peak of 73,9% in 2015. The percentage of non-attending students followed an even more marked trend, falling from 14,73% to 6,66% during the same period. However, the variable is characterized by an increasing and significative presence of missing values.

Table 4.3: Kindergarten attendance frequency table over the period 2013-2017.

Years	Yes	%	No	%	Missing	%
2013	350.337	68,79%	75.023	14,73%	83.958	16,48%
2014	385.802	73,42%	56.893	10,83%	82.773	15,75%
2015	384.387	73,90%	37.082	7,13%	98.703	18,98%
2016	381.824	72,81%	31.283	5,97%	111.321	21,23%
2017	320.037	71,14%	29.945	6,66%	99.893	22,20%
Overall	1.822.387	72,05%	230.226	9,10%	476.648	18,85%

Fig. 4.2: Percentage of students by kindergarten frequency from 2013 to 2017.



4.1.3 Students and parents' birthplace

Table 4.4 differentiates students in the sample according to their place of birth. In general, almost all of the students in the five years were born in Italy. Only a small percentage, around 6%, comes from a foreign country. In particular, immigrant students have been divided into European and non-European students labeled as "other". The percentage of non-immigrant students increased by one percentage point from 2013 to 2017. During the same period, both

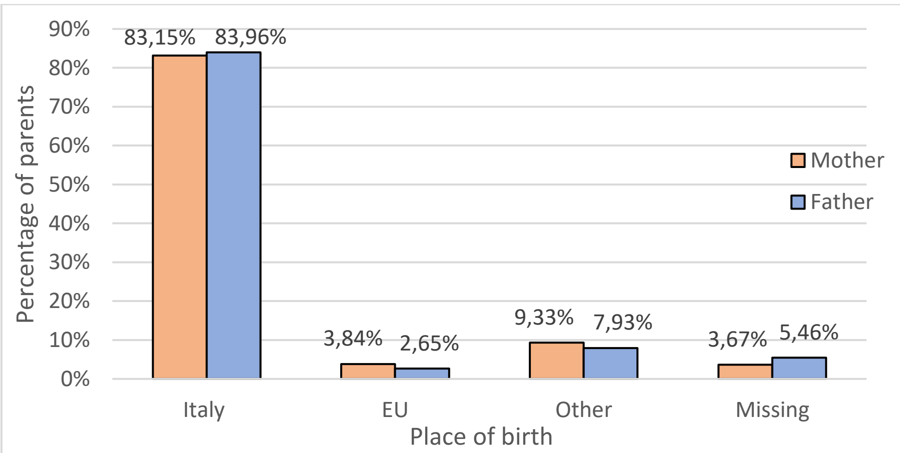
European and non-European students observed a slight percentage decrease. However, non-European students maintained a higher presence than European ones during the whole period.

Table 4.4: Students birthplace frequency table over the period 2013 to 2017.

Years	Italy	%	EU	%	Other	%	Missing	%
2013	470.839	92,44%	11.568	2,27%	22.332	4,38%	4.579	0,90%
2014	489.022	93,06%	11.593	2,21%	22.727	4,33%	2.126	0,40%
2015	488.848	93,98%	9.835	1,89%	20.333	3,91%	1.156	0,22%
2016	495.805	94,54%	8.914	1,70%	18.889	3,60%	820	0,16%
2017	421.341	93,66%	6.843	1,52%	17.431	3,87%	4.260	0,95%
Overall	2.365.855	93,54%	48.753	1,93%	1.017.12	4,02%	12.941	0,51%

Following the structure introduced before, Table 4.5 shows data on parents' place of birth. In particular, the description first focuses separately on father's and mother's birthplace, and then data are aggregated to show information at the couple level. A very large percentage of students, around 84%, reported having an Italian father, while more than 10% of students have a foreign-born father. In addition, most of the immigrant fathers are non-European, and only around 3% of them come from an EU country. Data on mother's birthplace are similar in proportion to those for the father. However, as shown in Fig. 4.3, mothers are more likely to come from a foreign country than fathers: around 13% of students reported having a foreign mother against 10,5% of the foreign fathers.

Figure 4.3: Comparison between father's and mother's birthplace.

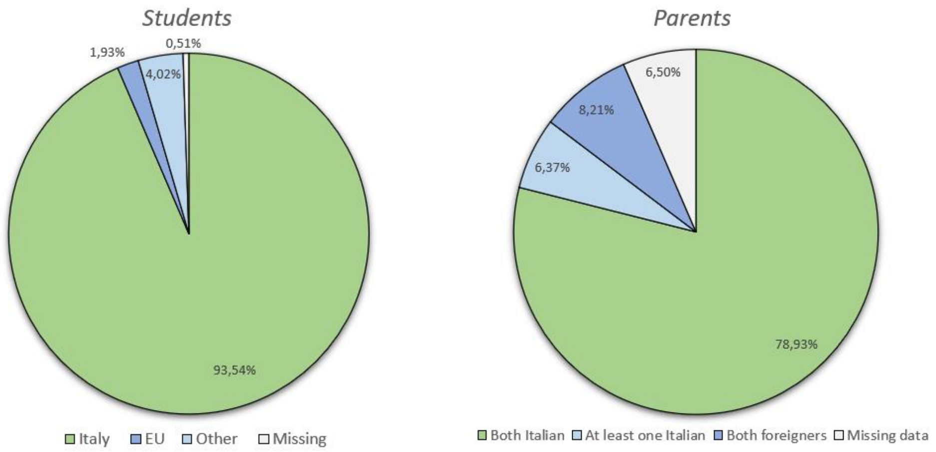


Looking at the couple level, around 79% of students have both Italian parents. Near 6,4% reported having only one Italian parent, while 8,2% of students have both foreign parents. However, the percentage of students with both Italian parents experienced a slight decrease over time, in opposition to the small percentage increase of students with at least a foreign parent.

Table 4.5: Parents birthplace frequency table over the period 2013 to 2017.

Years	Both Italian	%	Only one Italian	%	Both foreigners	%	Missing data	%
2013	401.113	78,75%	30.024	5,89%	40.944	8,04%	37.237	7,31%
2014	417.414	79,44%	33.421	6,36%	43.142	8,21%	31.491	5,99%
2015	413.468	79,49%	33.336	6,41%	42.022	8,08%	31.346	6,03%
2016	417.357	79,58%	34.673	6,61%	42.634	8,13%	29.764	5,68%
2017	346.873	77,10%	29.602	6,58%	38.855	8,64%	34.545	7,68%
Overall	1.996.225	78,93%	161.056	6,37%	207.597	8,21%	164.383	6,50%

Figure 4.4: Percentage of students (left-side) and parents (right-side) by birthplace - dataset.



4.1.4 Parents' educational background

I analyzed student's parents according to their educational background, distinguishing between high- and low-educated. According to the definition, high-educated people are those having more than average knowledge, usually with an education provided by a university. In this regard, parents with a university degree or higher qualification are considered as high-educated, while parents with a high-school diploma or lower qualification are considered as low-educated. In general, some high-schools may also provide a knowledge level considered as "above the average".

First, the description focuses on father's and mother's separately. Considering the whole dataset, around 67% of students reported having a low-educated mother, and 68% reported a low-educated father. As shown in Fig. 4.5, mothers tend to be more well-educated than fathers. Also, the percentage of students with a high-educated parent increased over the period, both in terms of father and mother. The percentage of students with a high-educated mother raised from 11,9% in 2013 to 15,6% in 2017. However, students with a high-educated father experienced a smaller increase over the same period, from 10,4% to 12,4%.

Figure 4.5: Comparison between high-educated parents from 2013 to 2017.

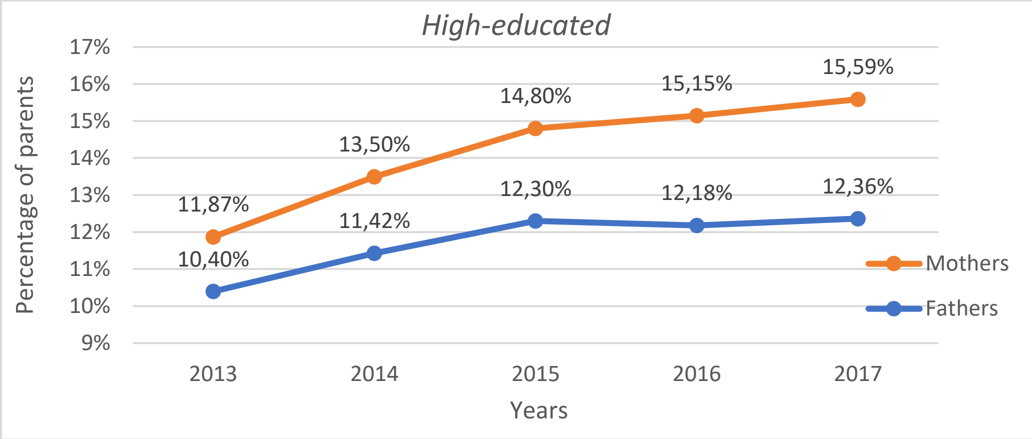
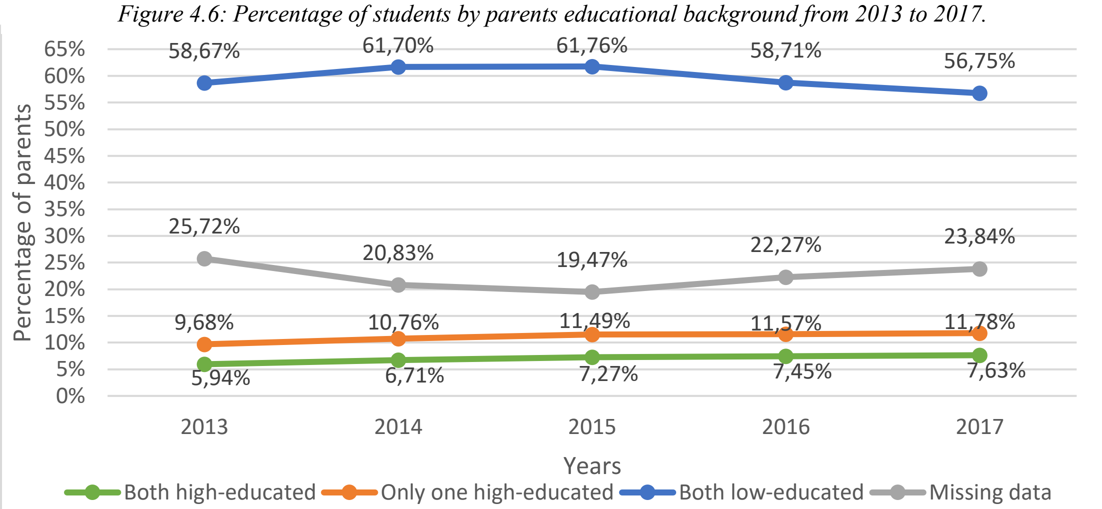


Table 4.6 presents data on parents' educational background at the couple level. A very large percentage of students, around 59,6%, reported having both low-educated parents, while 18% of students reported at least one educated parent. As the percentage of high-educated increased both for mothers and fathers over the period, we expect an increase also at the couple level. Indeed, as shown in Fig. 4.6, the percentage of students with both or at least one educated parent roughly raised by two percentage point from 2013 to 2017. In contrast, students with low-educated parents have declined since 2015. However, the presence of missing data is significantly high.

Table 4.6: Parents educational background frequency table over the period 2013-2017.

Years	Both high-educated	%	Only one high-educated	%	Both low-educated	%	Missing data	%
2013	30.234	5,94%	49.292	9,68%	298.794	58,67%	130.998	25,72%
2014	35.233	6,71%	56.532	10,76%	324.224	61,70%	109.479	20,83%
2015	37.819	7,27%	59.784	11,49%	321.272	61,76%	101.297	19,47%
2016	39.048	7,45%	60.683	11,57%	307.886	58,71%	116.811	22,27%
2017	34.333	7,63%	52.986	11,78%	255.302	56,75%	107.254	23,84%
Overall	176.667	6,98%	279.277	11,04%	1.507.478	59,60%	565.839	22,37%

Figure 4.6: Percentage of students by parents educational background from 2013 to 2017.



4.1.5 Parents' job situation

According to the definition, high-skilled workers are characterized by advanced education and possession of knowledge and skills to achieve complex tasks. In contrast, low-skilled labour refers to workforce characterized by lower educational attainment and by performing routine tasks, typically related to manual activities. Following this classification, parents were split in high-skilled employed, low-skilled employed and unemployed. The variable concerning parents job situation is composed of 9 classes, each containing different types of works. Therefore, some jobs considered as medium- or high-skilled (e.g. school or high-school teachers) are merged within the same class with jobs that may be considered as low-skilled (e.g. regular employee). Furthermore, the distinction between high- and low-skilled is not always clear-cut, and for simplicity, I did not introduce the medium-skilled category.

Looking separately at parents' data, it emerges a significantly higher unemployment rate for mothers (around 34,3%) compared to fathers (around 5,5%). However, the percentage of students with an unemployed mother decreased from 34,3% to 32,2% over the period. In particular, 89% of unemployed mothers are classified as "housewife". Also, as shown in Fig. 4.7, the percentage of fathers with a high-skilled position is almost double than that of mothers. This result is ambiguous: although the percentage of well-educated mothers is higher, this does not translate into high-skilled employment. A first explanation may be an existing gender gap in the workplace, facilitating men in reaching high-skilled position. However, both the percentage of high-skilled fathers and mothers raised by roughly a percentage point over the period.

Figure 4.7: Comparison between father's and mother's job situation.

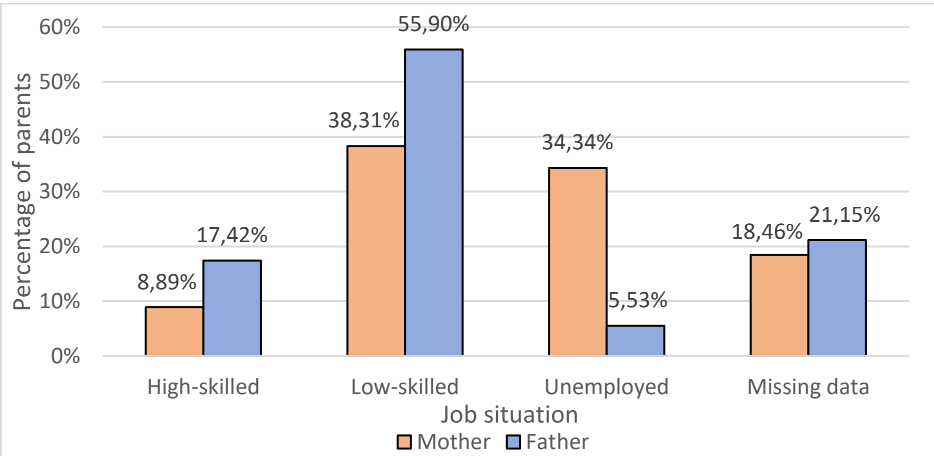


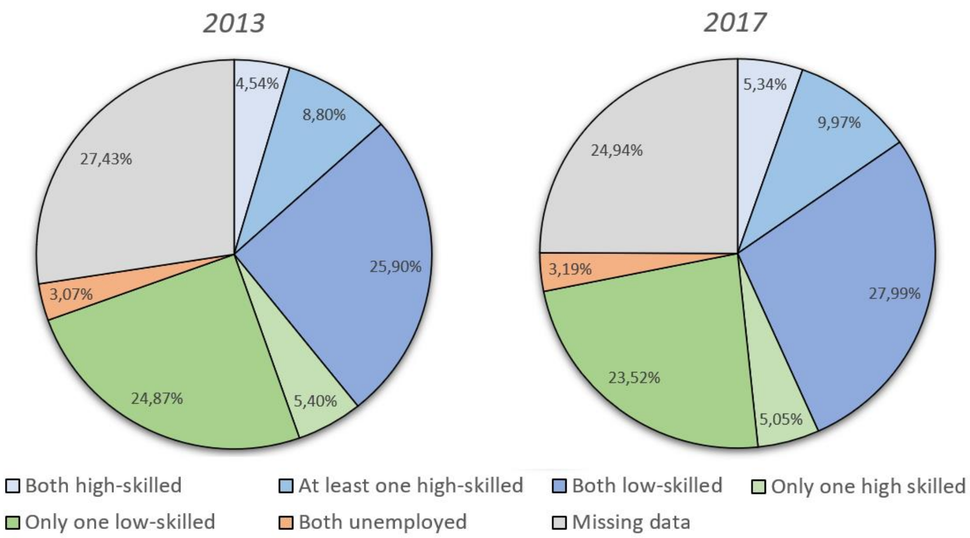
Table 4.7 summarizes parents' data at the couple level by job situation. Considering the entire dataset, around 42,4% of students reported having both parents employed. Most of

them reported both low-skilled parents, while just the 14,8% have at least a high-skilled parent. Near 30,5% of students have only one parent employed, and only 5,5% of them is high-skilled. The percentage of students with both parents unemployed remained around 3,2% over the period. Data at the couple level are treated as "missing" if at least one of the two data regarding parents is absent. Again, the presence of missing data is significantly high in each year.

Table 4.7: Parents job situation frequency table over the period 2013-2017.

Years	Both employed			Only one employed		Both unemployed	Missing data
	Both high-skilled	One high-skilled	Both low-skilled	High-skilled	Low-skilled		
2013	23.132 4,54%	44.805 8,80%	131.903 25,90%	27.507 5,40%	126.646 24,87%	15.618 3,07%	139.707 27,43%
2014	26.334 5,01%	50.835 9,67%	144.003 27,40%	29.829 5,68%	137.949 26,25%	18.736 3,57%	117.782 22,41%
2015	27.290 5,25%	53.194 10,23%	149.584 28,76%	29.698 5,71%	135.012 25,96%	17.882 3,44%	107.512 20,67%
2016	27.958 5,33%	52.778 10,06%	147.191 28,07%	28.031 5,35%	127.003 24,22%	14.939 2,85%	126.528 24,13%
2017	24.018 5,34%	44.842 9,97%	125.930 27,99%	22.724 5,05%	105.818 23,52%	14.340 3,19%	112.203 24,94%
Overall	128.732 5,09%	246.454 9,74%	698.611 27,62%	137.789 5,45%	632.428 25,00%	81.515 3,22%	603.732 23,87%

Figure 4.8: Percentage of parents at the couple level by job situation in 2013 and 2017.



4.2 Descriptive statistics for test scores

This section aims to describe test scores differentiating by several variables to understand how they affect students' achievement. Students' performance is measured through the

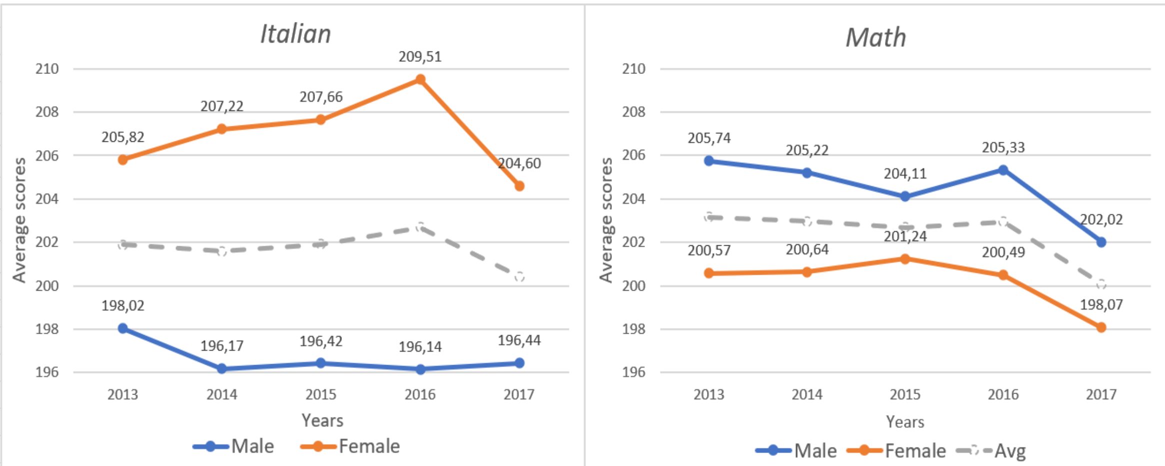
INVALSI national test scores. Following most European countries, the Italian Ministry of Education developed an evaluation system for students based on standardized tests. The INVALSI test was introduced in 2005, aiming to identify strength points and weaknesses of the education system, as well as areas that need improvement and existing inequalities between different schools or regions. However, the INVALSI test changed over the years adapting to new legislations, introducing improvement in type and quality of questions, and modifying administration methods. For this reason, test scores have been standardized with mean 200 and standard deviation 40 to allow comparison between different years.

4.2.1 Test scores by students' characteristics

Test scores analysis helps us to identify variables that affect student's performance and which categories of students may have an advantage or disadvantage performing Italian or Math. The following analysis focuses on the student-level characteristics described in the previous section such as gender, kindergarten frequency, and birthplace.

The main result that emerges differentiating test scores by students' gender is a performance gap between males and females. In particular, boys tend to do better in Math, while females outperform males in Italian scores. As shown in Fig. 4.9, the gender gap is much more evident for Italian tests. However, in contrast to the growth in the previous four years, female students experienced a significant deterioration in Italian scores between 2016 and 2017. Math scores have also declined over the last year, both for males and females.

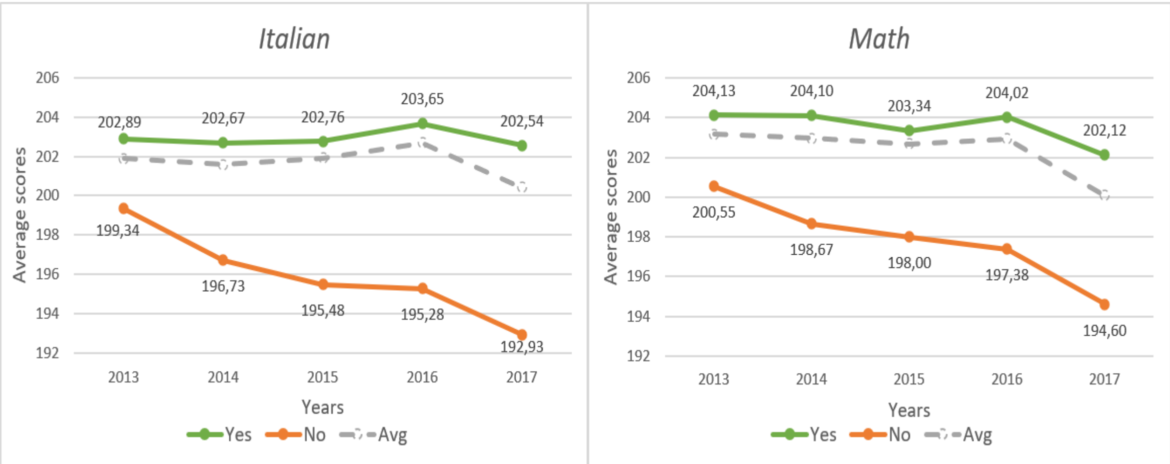
Figure 4.9: Italian and Math test scores differentiated by students gender over the period 2013-2017.



The existence of a gender gap is consistently documented in educational literature. In general, literature agrees that female students perform better than males in most of the school subjects, particularly on verbal scores. However, the presence of a females' disadvantage in Math is very important, because Math test scores serve as a good predictor for future income. In particular, according to the Programme for International Student Assessment (PISA), Italy resulted in one of the OECD countries with the largest math gender gap in 2015. Several explanations have been proposed for the existence of a gender gap: (i) girls tend to have higher levels of math anxiety and lower levels of confidence in math skills; (ii) boys have and develop superior spatial skills that give them a math advantage; (iii) gender-biased environments and social-conditioning such as stereotypes or parental expectations affect performance.

Test scores are differentiated also by kindergarten frequency. As shown in Fig. 4.10, students that attended pre-primary school perform systematically better than students that did not attend, suggesting a positive impact of kindergarten frequency on students' achievement. Moreover, the performance of non-attending students declined between 2013 and 2017, both for Italian and Math. This results in an increasing gap between students that attended or not kindergarten.

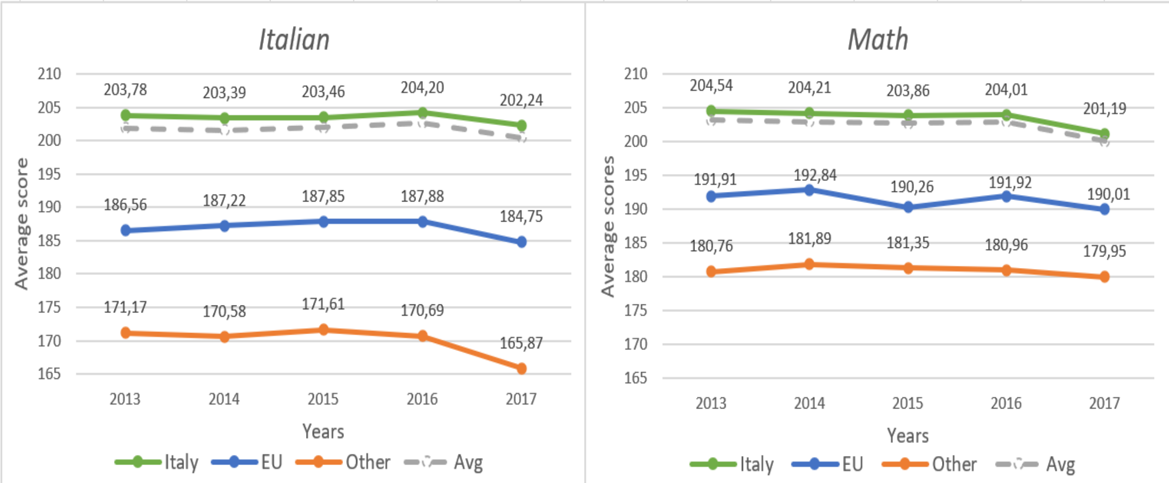
Figure 4.10: Italian and Math test scores differentiated by kindergarten frequency over the period 2013-2017.



The literature focusing on early childhood education support the positive impact of kindergarten participation on school achievement, both in the short- and long-term. In particular, the PISA report of 2011 found that students who attended pre-primary school outperformed students without preschool frequency, by about a year of achievement. The main explanation is that social relationships and interactions within early childhood, with both adults and other children, may produce benefits for cognitive, language and social development. This positive effect on students' performance supports the Italian government's decision to introduce compulsory kindergarten by 2023.

Figure 4.11 compares students' performance differentiating by place of birth, over the period 2013-2017. As shown, immigrant students consistently perform worse than native students, both in Italian and Math scores. Between immigrants, students from a European country seems to perform better than non-European. However, the gap between foreign and native students is lower in Math scores.

Figure 4.11: Italian and Math test scores differentiated by students birthplace over the period 2013-2017.



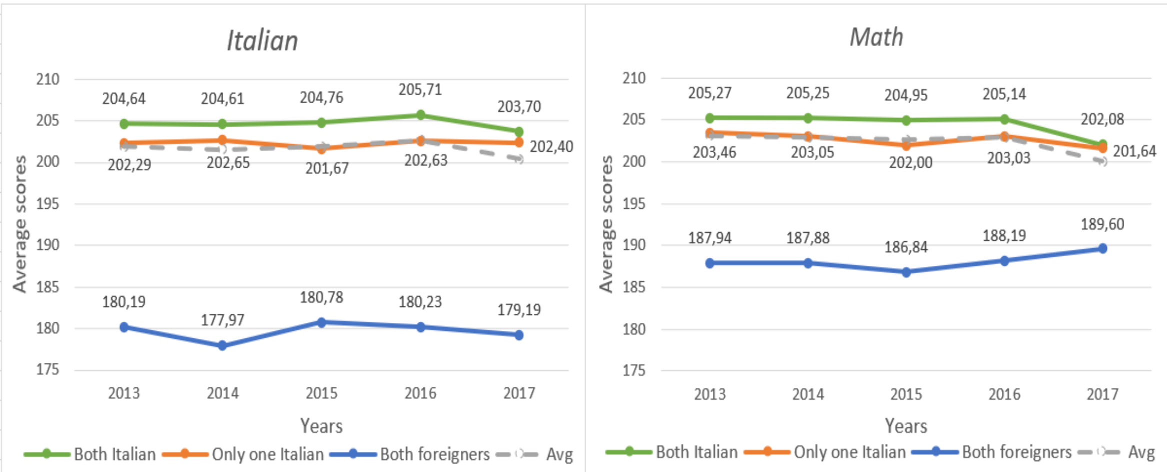
The performance gap is partly explained by differences in the socio-economic profile. Immigrant students often come from a disadvantaged socio-economic background when compared to native students. However, according to the literature, lower average performance is also associated with other factors such as language barriers, concentration in disadvantaged schools, or late school entry compared to native students. Immigrant students may find difficult adapting to the new language, explaining the additional negative effect on verbal scores. Also, immigrants tend to be concentrated in poor neighborhoods and their children are more likely to end up in lower-performing school with lower educational resources. The cultural proximity between the origin and the host country reduces the effect of barriers facing immigrants, explaining the higher performance of European students. Many immigrants may also face discrimination and racism at their schools, resulting in a restrictive learning environment.

4.2.2 Test scores by family background

Several studies prove that family background is strongly correlated with students' performance. Also, family may play an even more important role than school does, as it is the primary environment to which children are exposed. Family socio-economic background is principally measured by parental occupation, education level and birthplace. This section aims to identify how these variables impact students' achievement in school.

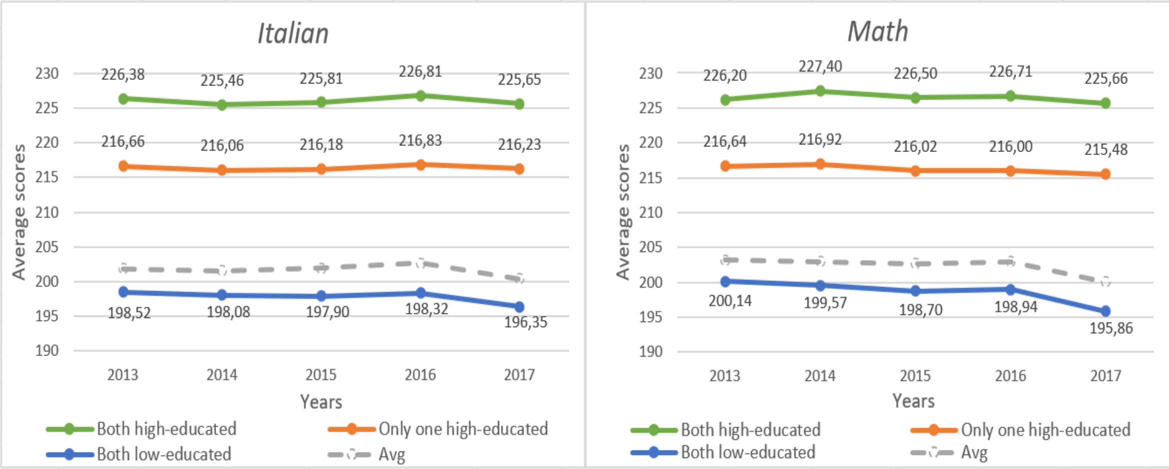
Test scores differentiated by parents' birthplace show results similar to those obtained in the students' case. Looking at data at the single-parent level, students with an immigrant mother or father underperform students with a native parent, both in Italian and Math scores. In particular, the disadvantage of tests performance is lower for students with a European rather than non-European parent. As shown in Fig. 4.12, data at the couple level show a wide gap between students with Italian and immigrant parents. The possible phenomena that explain this gap are similar to those presented in the previous section for students' birthplace. However, the presence of a single native parent has a strong positive effect on test scores: students with at least an Italian parent perform equal or better than the average of the sample. Also, those students do not perform better in Math rather than verbal scores, in contrast to students with both foreign parents. The presence of a native parent in the family seems to cancel the "language penalty" experienced by foreign students.

Figure 4.12: Italian and Math test scores differentiated by parents birthplace over the period 2013-2017.



Test scores suggest a strong positive effect of parental education on students' performance at school. Observing data at the level of single parent, students with a high-educated father or mother scored well above the average of the sample, between 221 and 223 points for both school subjects. However, as shown in Fig. 4.13, data at the couple level highlight an incremental positive effect in having both highly educated parents, raising average scores by about four points. Furthermore, even the presence of an educated parent alone leads students to score higher than the average, reinforcing the hypothesis of a positive impact.

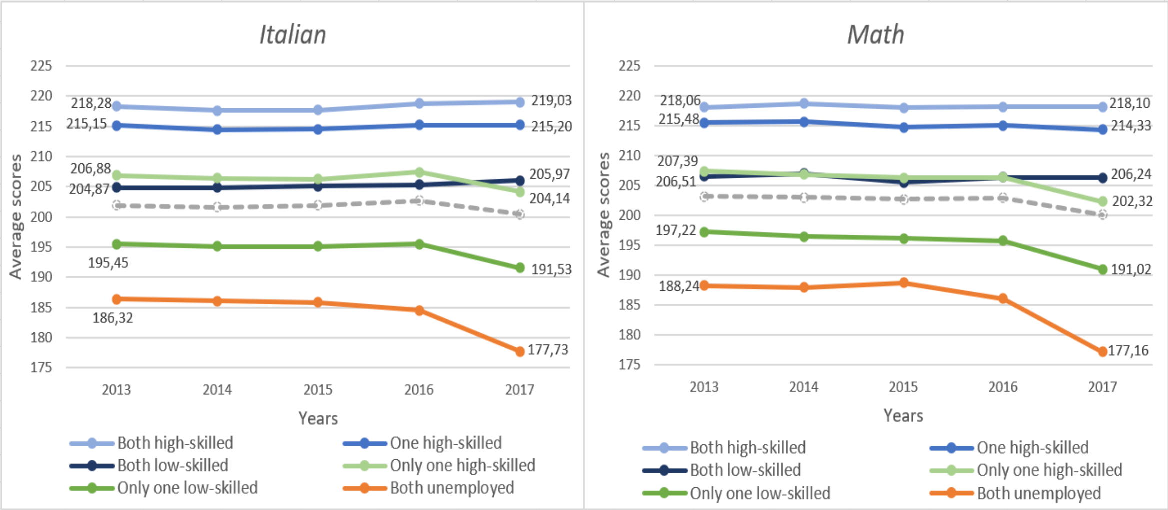
Figure 4.13: Italian and Math test scores differentiated by parents' educational background over the period 2013-2017.



In most studies, family educational background is identified as the strongest correlate of students' performance in school. High-educated parents may affect children's achievement both directly and indirectly through the school choice. They are more involved in children education, participating in their learning (e.g. assisting them in doing homework) and provide more activities enhancing children's academic development (e.g. encouraging them to read). In general, well-educated parents provide better educational opportunities and a learning environment that positively affect students' performance. Indirectly, they consider more the quality of schools and teachers and, in general, pay more attention to ensure good education tools for their children.

Test scores differentiated by parental job situation show that students with a high-skilled father perform well above the average, students having a low-skilled father perform on average, while students with an unemployed father have the worst test scores. This situation is similar when looking at mothers' employment. However, in contrast to fathers data, students with a low-skilled mother score from 4 to 6 points higher than the average. Also, students having an unemployed mother have better scores than students with an unemployed father, suggesting a higher negative effect of paternal job loss. Figure 4.14 summarizes test scores looking at the couple of parents. The blue lines identify the case where both parents are employed, while the green lines correspond to a single working parent. As shown, high-skilled parents seem to have a great impact on students' achievement. This hypothesis is reinforced by the fact that students with a single employed but high-skilled parent perform as the students with both low-skilled parents. Also, students with both or only one employed parent experienced a decrease in test scores between 2016 and 2017, while students with both employed parents slightly raised their performance during the same period.

Figure 4.14: Italian and Math test scores differentiated by parents job situation over the period 2013-2017.



Possible explanations are similar to those applied to parents' educational background, as high-education is strictly linked to high-skilled employment. However, according to the literature, the overall effect of parents' employment on children's learning is not clear: on the one hand, family income may play a major role providing better learning opportunities and resources, on the other hand, employed parents spend less time with their children. High-skilled employment leads to a higher family income, enabling parents to pay for high-quality schooling and additional resources (e.g. extra lessons). In contrast, parents with lower income focus on satisfying basic needs and pay less attention to children education. Also, unemployment may affect stress level, attitudes, and role models within the family, inducing negative spillover effects on children's school performance. The empirical research provides evidence on different implications of maternal rather than paternal unemployment, suggesting that fathers occupation loss has more severe effects. This effect has two possible explanations: (i) fathers' unemployment strongly lowers family income, as the father's salary usually represents the largest part of it; (ii) mothers tend to dedicate more time than fathers to children's education.

4.2.3 Test scores by administrative regions and cheating

Italy is characterized by strong differences in infrastructure, income, culture and quality of life travelling down the peninsula. In particular, we observe a significant economic inequality in term of household income between the "rich" North and the "poor" South. This inequality is mainly explained by lower industrialization and employment opportunities in Southern regions. Also, the Southern population is less educated and has higher school dropout rates than the North. As introduced before, family income and education are recognized as key factors influencing children's academic achievement, suggesting a significant "North-South gap" in students' test scores. However, as shown in Fig. 4.15 and 4.16, the gap between North and South is not evident until 2017. Regions such as Sicilia, Molise, Puglia and Calabria performed equally or even better than Northern regions in both Italian and Math between 2013 and 2015. Starting in 2016, Southern regions experienced a negative trend in test scores, losing more than 10 points on average each year, and leading to the territorial divide in 2017. In contrast, Northern regions slightly improved performance over the period.

Figure 4.15: Italian test scores differentiated by administrative region in 2013,2015 and 2017.

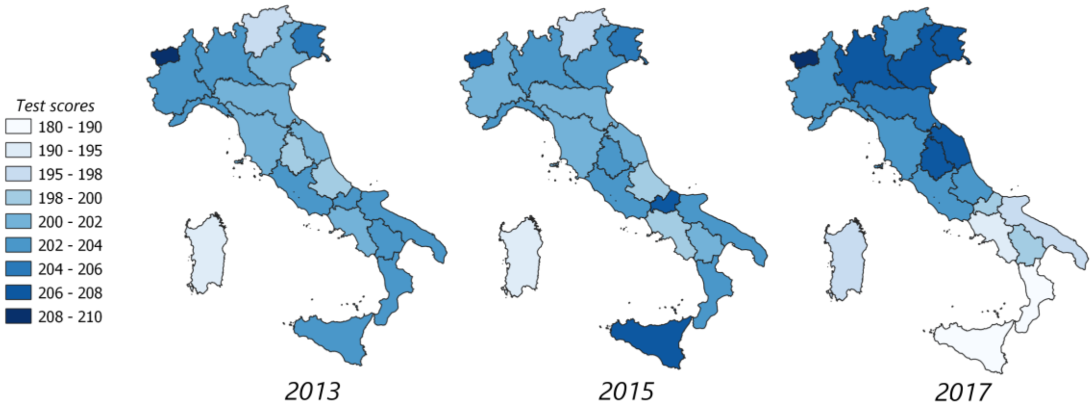
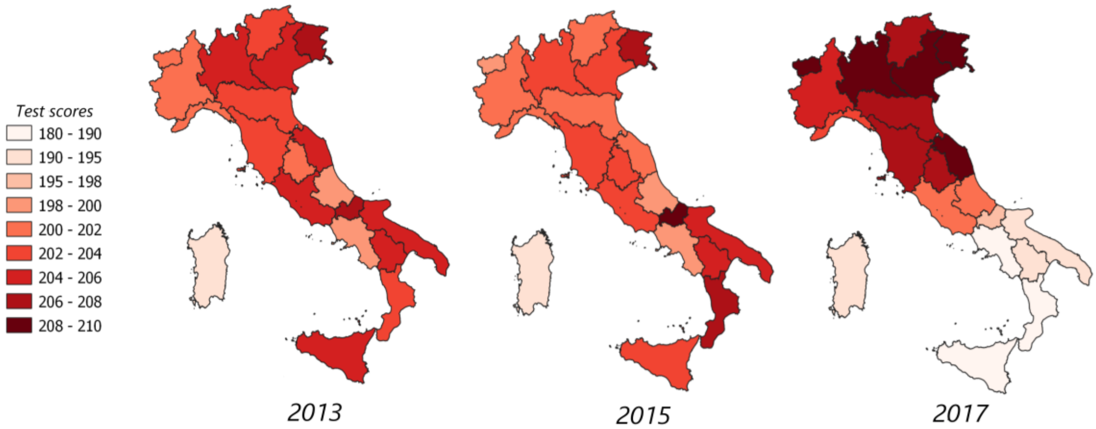


Figure 4.16: Math test scores differentiated by administrative region in 2013,2015, and 2017.

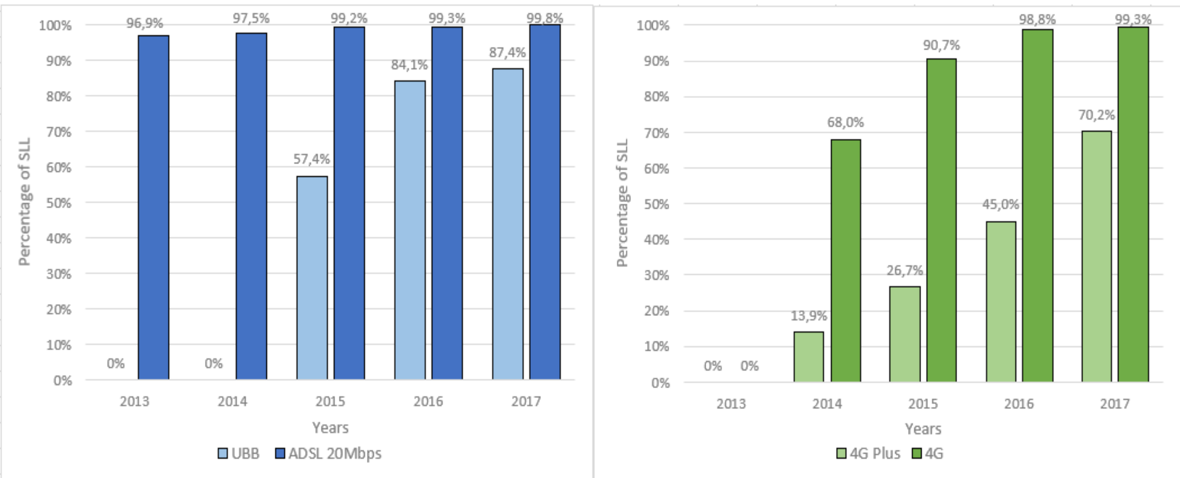


A possible explanation for this trend may be related to the *cheating* phenomenon. Cheating is a recognized event that may occur among students (e.g. copying answers) or because of teachers (e.g. suggesting answers or cheating in the test correction). The INVALSI institute measures cheating with a statistical indicator that considers the homogeneity of responses, the number of missing answers, and scores variability. According to INVALSI, cheating occurs more in Southern regions, particularly for eighth-grade students where the test contributes to the evaluation of the final exam. Starting in 2013, INVALSI administered tests in five different versions, changing the order of the answers to contrast misconduct between students. However, the presence of cheating is more closely related to teachers' behavior. The reasons for the teacher cheating are unclear: fear of judgement, attempt to make a good impression, or boycott. For simplicity, test scores used in the analysis do not consider the corrective factor for cheating. Since 2017, the INVALSI test is no longer part of the national eighth-grade exam. Although this event may have reduced the cheating phenomenon, it is not considered as a "shock" that can explain such anomalies in the observed data.

4.3 Descriptive statistics for broadband penetration

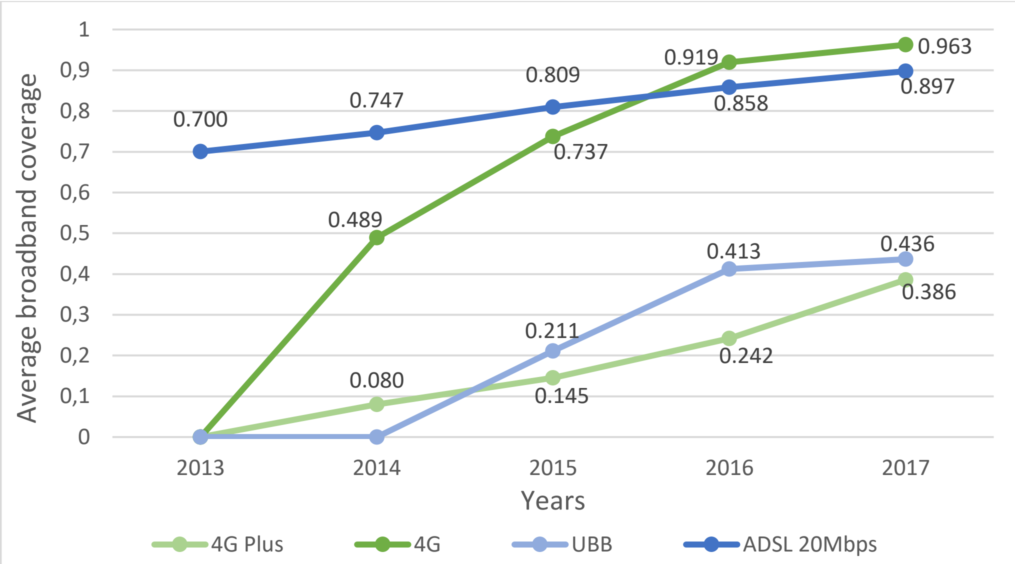
Data on the availability and coverage of broadband technologies are provided at the SLL-level. According to the ISTAT definition, *Local Labour Systems* are sub-regional geographical areas where the bulk of the labour force lives and works. The key criterion to define a SLL is the proportion of commuters who cross the SLL boundary while going at work. Figure 4.17 summarizes data on fixed and mobile broadband availability in SLLs, over the period 2013-2017. As shown, 20Mbps ADSL technology was widespread and present in almost all SLLs over the considered period. Ultrafast broadband (UBB) diffusion started in 2015 and the percentage of SLLs that adopted fiber-based technologies raised by 30 percentage point, from 57,4% to 87,4% in just three years. 4G diffusion started in 2014 and was faster than UBB: in 2016, almost all SLLs were connected with the fourth-generation mobile broadband.

Figure 4.17: Fixed (left-side) and mobile (right-side) broadband adoption in Italian SLLs over the period 2013-2017



However, the presence of a specific technology does not give information on the extent of the diffusion within the SLL. Broadband coverage is defined as the proportion of households who can get broadband access. For completeness, Fig. 4.18 shows the average broadband coverage for both fixed and mobile technologies over the period 2013-2017. In conclusion, not only fiber-based UBB and 4G adoption but also the actual coverage of these technologies increased over the considered period.

Figure 4.18: Average broadband coverage in Italian SLLs over the period 2013-2017.



5. Methodology and Results

This chapter includes the description of the estimation models used in the empirical analysis and the results obtained. The models aim to estimate the impact of fiber-based broadband, also known as ultra-fast broadband (UBB), and 4G mobile broadband on students' academic achievement. In particular, this analysis mainly uses two estimation models: *ordinary least squares* (OLS) and a *fixed-effects model*. OLS is a statistical method that identifies the relationship between one or more independent variables and a dependent variable, estimating the unknown parameters in a linear regression model. The parameters of the linear function are chosen by minimizing the sum of the squares of the differences between the observed values contained in the dataset and those predicted by the linear function. The estimated parameters identify the strength of the relationship between the dependent variable and the explanatory variables. The fixed-effect model refers to a regression model that allows controlling for time-invariant unobserved characteristics that can be correlated with the regressors and may affect the outcome of the analysis. Fixed effects estimator reduces omitted variable bias by measuring changes within groups across times. The unobserved group-specific heterogeneity can be removed from the data by subtracting group means from individual observation within each group. The statistical software used for the econometric analysis is Stata (version 13.0).

5.1 Empirical specification

The effect of broadband penetration on students' performance is estimated through the following equation:

$$Y_{i,t,s} = \beta_0 + \beta_1 BB_{s,t} + \beta_2 X_{i,s,t} + \alpha_s + \tau_t + \varepsilon$$

The dependent variable $Y_{i,t,s}$ represents the language or mathematics test scores of the individual student i, in the school s, at time t. The variable $BB_{s,t}$ is a dummy variable for whether or not the school s had a certain broadband connection available at year t. $X_{i,s,t}$ is a vector of control variables related to the dependent variable. The vector of controls comprises the student- and family-level characteristics described in the previous chapter such as student's gender, kindergarten frequency, family birthplace, parental education and parental employment. Those variables are included between regressors to separate their effects from the explanatory variable as they may have a strong impact on students' performance (as suggested by the previous test scores analysis). The terms α_s and τ_t are respectively school and year fixed effects. School fixed effect controls for school-specific unobserved factors (e.g. quality of teachers). Year fixed effect controls for factors changing each year, capturing the influence of time trends on test scores. Basically, the software Stata specifies a dummy variable for each year and school in the sample and runs the fixed effect regression. The coefficient β_0, β_1, and β_2 are the unknown parameters to be estimated by the model, while ε is the econometric error. The parameter β_1 is the coefficient of interest, which measures the effect of broadband on students' test scores.

The study also focuses on finding the interaction effects of broadband on other variables. In a regression, interaction effects exist when the effect of an independent variable on a dependent variable depends on the values of another independent variable. This effect is introduced in the model by interacting two independent variables. For example, I used the following specification to estimate the interaction effect between broadband and students' gender:

$$Y_{i,t,s} = \beta_0 + \beta_1 BB_{s,t} + \beta_2\left(BB_{s,t} \times Male_i\right) + \beta_3 X_{i,s,t} + \alpha_s + \tau_t + \varepsilon$$

This model helps us to verify if broadband affects male students differently compared to females. As shown, the specification includes the interaction between the dummy "broadband" and a dummy variable that takes the value 1 for male students. In this case, the coefficient β_1 measures the effect of broadband on female students, while β_2 gives us the differential effect between males and females. Therefore, the overall impact of broadband on male students is given by the sum of the two coefficients $\beta_1 + \beta_2$. In this study, the interaction effect of broadband is estimated not only for students' gender, but also for parents' educational background and job situation.

This study mainly focuses on how ultra-fast broadband, referring to fiber-based connections, impact on students' performance. However, students at school do not only have Internet access through school's fixed connection. Students can also surf the Internet using the mobile network of their cell phones. For this reason, the previous estimation models were also used to evaluate the impact of 4G mobile broadband on students' test scores. Finally, the interaction effect between the two broadband technologies was evaluated using the following specification:

$$Y_{i,t,s} = \beta_0 + \beta_1 UBB_{s,t} + \beta_2 (UBB_{s,t} \times 4G_{s,t}) + \beta_3 X_{i,s,t} + \alpha_s + \tau_t + \varepsilon$$

In this case, the coefficient β_1 represents the effect on students that only have a fixed fiber-based connection available, while $\beta_1 + \beta_2$ is the overall effect on students that have both technologies available.

As the broadband availability occurs at the school level, students' behavior is likely correlated within schools. In this case, classical standard errors may be inappropriate and could lead to overestimating the precision of the effects, because tests scores within each school are not independently distributed. To account for this aspect of the study, we allowed clustered standard errors at the school level.

In conclusion, Table 5.1 summarizes the relevant variables used in the empirical specifications:

Table 5.1: Summary of the variables employed in the empirical analysis.

Variable	***Type***	***Description***
$WLE\_ITA\_200_{i,t,s}$	Dependent	The variable represents Italian test scores standardized with a mean of 200 and a standard deviation of 40 for student *i*, in school *s*, at time *t*.
$WLE\_MAT\_200_{i,t,s}$	Dependent	The variable represents Math test scores standardized with a mean of 200 and a standard deviation of 40 for student *i*, in school *s*, at time *t*.
$UBB_{s,t}$	Explanatory	Dummy variable representing schools where an ultra-fast broadband connection is available at time t. Since UBB was introduced in 2015, the dummy always assumes a null value before that year.
$4G_{s,t}$	Explanatory, Interaction	Dummy variable representing schools covered by 4G mobile broadband at time t. Since 4G was introduced in 2014, the dummy always assumes a null value before that year.
$Male_i$	Control, Interaction	Dummy variable that is equal to 1 if student's gender is male, and equal to 0 if it is female.
$Kindergarten\_freq_i$	Control	Dummy variable that is equal to 1 if the student *i* attended kindergarten, and 0 otherwise.
$Italian\_student_i$	Control	Dummy variable that is equal to 1 for students that were born in Italy, and 0 otherwise.
$Italian\_father_i$	Control	Dummy variable that is equal to 1 for students with an Italian born father, and 0 otherwise.
$Italian\_mother_i$	Control	Dummy variable that is equal to 1 for students with an Italian born mother, and 0 otherwise.

High_ed_father$_i$	Control	Dummy variable that is equal to 1 for students with a high-educated father, and 0 otherwise.
High_ed_mother$_i$	Control	Dummy variable that is equal to 1 for students with a high-educated mother, and 0 otherwise.
Both_high_ed$_i$	Control, Interaction	Dummy variable that is equal to 1 for students whit both parents high-educated, and 0 otherwise.
Atleast_high_ed$_i$	Interaction	Dummy variable that is equal to 1 for students with at least a high-educated parent, and 0 otherwise.
High_skill_father$_i$	Control	Dummy variable that is equal to 1 for students with an employed high-skilled father, and 0 otherwise.
High_skill_mother$_i$	Control	Dummy variable that is equal to 1 for students with an employed high-skilled mother, and 0 otherwise.
Both_high_skill$_i$	Control, Interaction	Dummy variable that is equal to 1 for students with both employed high-skilled parents, and 0 otherwise.
Low_skill_father$_i$	Control	Dummy variable that is equal to 1 for students with an employed low-skilled father, and 0 otherwise.
Low_skill_mother$_i$	Control	Dummy variable that is equal to 1 for students with an employed low-skilled mother, and 0 otherwise.
Both_low_skill$_i$	Control	Dummy variable that is equal to 1 for students with both employed high-skilled parents, and 0 otherwise.
Atleast_both_low$_i$	Interaction	Dummy variable that is equal to 1 for students with at least a high-skilled parent or both low-skilled parents, and 0 otherwise.
Both_unemployed$_i$	Interaction	Dummy variable that is equal to 1 for students with both unemployed parents, and 0 otherwise.
Post_student$_i$	Control	Dummy variable that is equal to 1 for repeating students or for students with a delayed school entry, and 0 otherwise.

5.2 UBB results

This section presents the regressions results focusing on the relationship between ultra-fast broadband and students' performance. The analysis verifies also if this impact is homogeneous or heterogeneous between different group of students. To do so, we estimate the interaction effect between the variable UBB and other variables regarding students' gender, parental education and parental employment. Finally, we exploit the geographical dimension of the dataset by using three separate regressions for students from North, Center and South Italy.

Table 5.2 provides results when regressing test scores against the explanatory variable UBB, controlling for students' characteristics and family background. Columns (1), (2), (3) present the results for Italian scores, while columns (4), (5), (6) for Math scores. In particular, columns (1) and (4) refer to the OLS specification, columns (2) and (5) include year fixed-effects, and columns (3) and (6) also add school fixed-effects.

Table 5.2: Regressions results on ultra-fast broadband for Italian and Math test scores, using OLS and FE models.

	WLE_ITA_200			*WLE_MAT_200*		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	FE1	FE2	OLS	FE1	FE2
UBB	-0.349**	0.948**	0.076	-1.455***	1.344***	1.304***
	(0.149)	(0.423)	(0.347)	(0.173)	(0.513)	(0.394)
Male	-9.639***	-9.635***	-9.609***	5.457***	5.464***	5.475***
	(0.071)	(0.071)	(0.070)	(0.089)	(0.088)	(0.087)
Kindergarten_freq	3.037***	3.065***	3.855***	2.170***	2.221***	2.913***
	(0.343)	(0.342)	(0.275)	(0.415)	(0.413)	(0.275)
Italian_student	6.272***	6.306***	6.266***	3.284***	3.351***	3.326***
	(0.216)	(0.216)	(0.208)	(0.210)	(0.210)	(0.202)
Italian_father	5.687***	5.670***	5.725***	2.856***	2.822***	3.261***
	(0.174)	(0.174)	(0.160)	(0.174)	(0.174)	(0.162)
Italian_mother	3.738***	3.709***	3.910***	1.698***	1.639***	2.177***
	(0.157)	(0.157)	(0.141)	(0.164)	(0.164)	(0.143)
High_ed_father	13.709***	13.699***	12.672***	13.067***	13.045***	12.153***
	(0.152)	(0.152)	(0.148)	(0.172)	(0.172)	(0.157)
High_ed_mother	14.319***	14.330***	13.500***	13.378***	13.397***	12.777***
	(0.131)	(0.131)	(0.127)	(0.140)	(0.140)	(0.130)
Both_high_ed	-5.929***	-5.940***	-5.698***	-4.481***	-4.503***	-4.349***
	(0.212)	(0.212)	(0.206)	(0.238)	(0.238)	(0.225)
High_skill_father	11.941***	11.941***	9.906***	12.015***	12.025***	9.478***
	(0.214)	(0.214)	(0.168)	(0.235)	(0.234)	(0.167)
High_skill_mother	6.678***	6.692***	5.393***	6.870***	6.900***	5.042***
	(0.199)	(0.199)	(0.164)	(0.221)	(0.220)	(0.172)
Both_high_skill	-2.890***	-2.899***	-2.587***	-2.912***	-2.930***	-2.442***
	(0.230)	(0.230)	(0.222)	(0.245)	(0.244)	(0.232)
Low_skill_father	7.254***	7.258***	6.018***	7.514***	7.532***	5.808***
	(0.196)	(0.196)	(0.157)	(0.209)	(0.209)	(0.152)
Low_skill_mother	5.743***	5.749***	4.796***	6.077***	6.091***	4.589***
	(0.170)	(0.169)	(0.142)	(0.186)	(0.185)	(0.146)
Both_low_skill	0.764***	0.768***	0.570***	1.224***	1.234***	0.941***
	(0.163)	(0.163)	(0.155)	(0.170)	(0.170)	(0.158)
Post_student	-23.996***	-23.999***	-23.583***	-22.829***	-22.844***	-22.499***
	(0.161)	(0.160)	(0.150)	(0.165)	(0.164)	(0.153)
Constant	177.68***	178.00***	179.23***	178.85***	179.18***	180.49***
	(0.447)	(0.463)	(0.379)	(0.478)	(0.497)	(0.374)
FE Year	NO	YES	YES	NO	YES	YES
FE School	NO	NO	YES	NO	NO	YES
Observations	1 445 203	1 445 203	1 445 203	1 455 758	1 445 758	1 445 758
R^2	0.131	0.131	0.169	0.098	0.099	0.155

Standard errors are clustered by school and found in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

The presented regressions help us to verify if the control variables impact on test scores consistently with the expectations from the descriptive analysis in Chapter 4. In general, our findings strongly agree with the existing literature and all the estimated coefficients have the expected sign. The results show a gender-related performance gap: females outperform males in Italian scores, while male students have better results in Math. Kindergarten frequency has a positive impact on students' achievement for both school subjects. Students with a family immigration background perform worse than Italian students in both subjects. However, this gap is lower in Math scores, probably due to lower language barriers. Parental education confirms as the strongest correlate of students' performance in school: students with at least a high-educated parent score on average 12-20 points higher in both Italian and Math scores. Parental job situation positively affects students' performance in proportion to the employment level of parents. As expected, fathers' employment has a more severe effect for both high- and low-skilled job.

Focusing on the explanatory variable UBB, the results show a positive and significant relationship between ultra-fast broadband and Math scores, while results for Italian scores lost statistical significance after including school fixed-effects. Moreover, the effect of UBB on Math scores remained highly significant across all specifications. The sign of the estimated coefficients become positive only when year fixed-effects are included in the regression, suggesting a negative time trend in test scores. Our results suggest that UBB connection only impacts on students' achievement in Math. However, the impact is small: the presence of UBB raises test scores by just 0.0326 standard deviations. Consistent with our findings, the existing literature generally agrees on a more effective impact of new technologies in school on scientific rather than verbal subjects (Angrist and Lavy [2002], Rouse and Krueger [2004], Barrow et al. [2009]). Moreover, Silva et al. (2014) and Grimes and Townsend (2017) also found a small but positive relationship between broadband and Math scores. However, in contrast to our results, this effect was also present in verbal scores.

Table 5.3 displays outcomes for the specifications that include the interaction effect between UBB and students' gender. For Italian scores, the estimated coefficients show an insignificant effect of UBB on female students and a negative correlation with males' performance. As shown, the interaction effect is statistically significant and negative. This gender-related difference is also evident in Math scores: the presence of UBB raises females scores by 1.619 points and males scores by just 0.894 points. In general, we can affirm that UBB impacts less effectively on male students. This evidence was also found in Belo et al. (2014). According to Belo, boys and girls might use the Internet connection to perform different activities. In particular, boys tend to perform more distracting activities on the Internet than girls, such as watching video, listening to music or playing video-games. Although this may explain the adverse effect on male students, as opposed to their results, our correlation between UBB and Math scores remained positive and statistically significant.

Table 5.3: Regressions results with interaction effect between ultra-fast broadband and students' gender.

		WLE_ITA_200		*WLE_MAT_200*	
		(1)	(2)	(3)	(4)
	UBB	-0.110	0.111	1.249***	1.619***
		(0.305)	(0.311)	(0.360)	(0.369)
	UBB × Male	-	-1.187***	-	-0.725***
			(0.101)		(0.105)
	Constant	237.006***	236.679***	231.317***	231.117***
		(0.330)	(0.332)	(0.364)	(0.367)
Dummy	Student-level	YES	YES	YES	YES
	Family Birthplace	YES	YES	YES	YES
	Parental education	YES	YES	YES	YES
	Parental employment	YES	YES	YES	YES
	Observations	2 517 645	2 517 645	2 521 459	2 521 459
	R^2	0.210	0.210	0.183	0.183

Standard errors are clustered by school and found in parentheses. All results have both year and school fixed effects.
* p<0.10, ** p<0.05, *** p<0.01

Table 5.4 shows our results on the interaction effect between UBB and parental education. In particular, columns (1) and (3) interact the variable UBB with a dummy that identifies students with both high-educated parents, while columns (2) and (4) focus on the effect on students with at least a high-educated parent.

Table 5.4: Regressions results with interaction effect between ultra-fast broadband and parental education.

		WLE_ITA_200		*WLE_MAT_200*	
		(1)	(2)	(3)	(4)
	UBB	-0.003	-0.168	1.170***	0.995***
		(0.327)	(0.328)	(0.376)	(0.377)
	UBB × Both_high_ed	0.937***	-	1.343***	-
		(0.223)		(0.245)	
	UBB × Atleast_high_ed	-	1.003***	-	1.269***
			(0.153)		(0.162)
	Constant	238.525***	237.095***	231.747***	232.616***
		(0.358)	(0.392)	(0.384)	(0.437)
Dummy	Student-level	YES	YES	YES	YES
	Family Birthplace	YES	YES	YES	YES
	Parental education	YES	YES	YES	YES
	Parental employment	YES	YES	YES	YES
	Observations	1 960 882	1 982 645	1 961 393	1 983 158
	R^2	0.212	0.212	0.189	0.189

Standard errors are clustered by school and found in parentheses. All results have both year and school fixed effects.
* p<0.10, ** p<0.05, *** p<0.01

Our findings show a positive and significant relationship between UBB and language scores of students with both high-educated parents. The correlation remains positive and highly significant looking also at students with at least a high-educated parent. This group of students seems to benefit from UBB technology in a better way compared to students without a well-educated family background. Looking at Math outcomes, all students experience a positive and significant effect of UBB, whatever the parental education is. However, the positive relationship is stronger for students with both or at least one high-educated parent. On average, UBB increases Math scores by 2.513 points for students with both high-educated parents, and by 2.264 for students with at least a high-educated parent. In conclusion, UBB impacts very differently on students with diverse parents' educational background. Having even one high-educated parent allows students to better appropriate from the benefit of UBB in both school subjects.

Table 5.5 provides the results of the interaction effects between UBB and parental employment. In particular, columns (1) and (4) include the interaction between the variable UBB and a dummy that identifies students with both employed high-skilled parents. Columns (2) and (5) use a dummy variable that identifies students with at least one high-skilled parent or both low-skilled parents, while columns (3) and (6) refers to students with both unemployed parents.

Table 5.5: Regressions results with interaction effect between ultra-fast broadband and parental employment.

		WLE_ITA_200			*WLE_MAT_200*		
		(1)	(2)	(3)	(4)	(5)	(6)
	UBB	-0.029	-0.788**	0.124	1.066***	0.393	1.216***
		(0.325)	(0.341)	(0.324)	(0.373)	(0.390)	(0.371)
	UBB × Both_high_skill	0.993***	-	-	1.345***	-	-
		(0.232)			(0.254)		
	UBB × Atleast_both_low	-	1.327***	-	-	1.236***	-
			(0.147)			(0.152)	
	UBB × Both_unemployed	-	-	-2.411***	-	-	-1.732***
				(0.335)			(0.347)
	Constant	238.17***	238.40***	237.67***	232.01***	232.20***	232.07***
		(0.351)	(0.358)	(0.328)	(0.381)	(0.394)	(0.360)
Dummy	Student-level	YES	YES	YES	YES	YES	YES
	Family Birthplace	YES	YES	YES	YES	YES	YES
	Parental education	YES	YES	YES	YES	YES	YES
	Parental employment	YES	YES	YES	YES	YES	YES
	Observations	1 923 022	1 946 700	1 923 022	1 923 563	1 947 254	1 923 563
	R^2	0.213	0.213	0.213	0.187	0.187	0.187

Standard errors are clustered by school and found in parentheses. All results have both year and school fixed effects.

* p<0.10, ** p<0.05, *** p<0.01

As shown, the impact of UBB on students' performance depends on the job situation of parents. The presence of a fiber-based connection improves the performance of students with both employed high-skilled parents: on average, Italian and Math scores increase respectively by 0.993 and 2.411 points. The relationship between UBB and test performance remains positive also considering students with at least a high-skilled parent or both low-skilled: Italian scores increase on average by 0.539 points, while Math scores raise by 1.236 points. In contrast, UBB has a highly significant negative effect on both language and Math scores for students with both parents unemployed. As shown, the size of the effect increases as the employment situation of parents improves, expanding the existing gap in performance between students from different social status. However, the positive effect of UBB extends to a very large group of students: students with at least an employed parent, even if low-skilled, are positively affected by the presence of UBB technology in their Math scores. For Italian scores, the positive effect extends to all students with both employed parents, even if both low-skilled.

Table 5.6 displays the correlation between UBB and test scores differentiating students by geographical area. In particular, column (1) refers to students from North, column (2) to students from Center, and column (3) to students from South Italy.

Table 5.6: Regressions results on ultra-fast broadband for Italian and Math scores differentiated by geographical area.

		(1) North	(2) Center	(3) South
WLE_ITA_200	*UBB*	-0.146	-0.937	-0.091
		(0.307)	(0.772)	(0.703)
	Constant	238.168***	233.722***	237.510***
		(0.414)	(0.658)	(0.754)
	Observation	1 103 736	471 045	942 864
	R^2	0.227	0.189	0.208
WLE_MAT_200	*UBB*	0.282	0.207	1.279
		(0.331)	(0.819)	(0.829)
	Constant	232.151***	228.392***	230.912***
		(0.459)	(0.731)	(0.813)
	Observation	1 107 825	470 943	942 691
	R^2	0.187	0.149	0.202
Dummy	Student-level	YES	YES	YES
	Family Birthplace	YES	YES	YES
	Parental education	YES	YES	YES
	Parental employment	YES	YES	YES

Standard errors are clustered by school and found in parentheses. All results have both year and school fixed effects.
* p<0.10, ** p<0.05, *** p<0.01

All the estimated coefficients are insignificant for both school subjects and for all geographical areas. In particular, the UBB coefficients for Center and South Italy are characterized by a high standard deviation. In general, these outcomes confirm the null effect of UBB on language scores but reverse our expected positive impact on Math scores.

5.3 UBB and 4G results

As mentioned above, students at school may have Internet access through the fixed connection of the school and the mobile connection of their smartphone. Furthermore, these two broadband technologies are typically found also at home, where children usually study or do their homework. For this reason, it is crucial to understand how UBB and 4G interact with each other influencing students' performance. This section presents regressions results focusing on the relationship between the availability of a 4G connection and test scores, both for Italian and Math. Lastly, we run a regression that includes the interaction effect between the two broadband technology.

Table 5.7 provides our results on the relationship between 4G and students' test scores. In particular, columns (2) and (4) introduce the interaction between 4G and students' gender.

Table 5.7: Regressions results with interaction effect between 4G and students' gender.

		WLE_ITA_200		*WLE_MAT_200*	
		(1)	(2)	(3)	(4)
	4G	-0.454*	0.445	-0.237	-0.155
		(0.270)	(0.278)	(0.285)	(0.297)
	4G × Male	-	-1.767***	-	-0.163
			(0.115)		(0.120)
	Constant	238.007***	236.365***	231.321***	231.262***
		(0.330)	(0.334)	(0.364)	(0.369)
Dummy	Student-level	YES	YES	YES	YES
	Family Birthplace	YES	YES	YES	YES
	Parental education	YES	YES	YES	YES
	Parental employment	YES	YES	YES	YES
	Observations	2 517 645	2 517 645	2 521 459	2 521 459
	R^2	0.210	0.210	0.182	0.182

Standard errors are clustered by school and found in parentheses. All results have both year and school fixed effects.
* p<0.10, ** p<0.05, *** p<0.01

The presence of a 4G connection does not appear to affect Math scores, even separating its effect between male and female students. Instead, the outcome of the regression for Italian scores shows a little negative relationship with 4G. However, this negative correlation is more severe and significant for male students. These findings show again that an increase in connectivity tend to negatively affect boys rather than girls.

Table 5.8 displays the results of the interaction between 4G and parental education. The observed outcomes are similar to those found for UBB. Students with at least a high-educated parent benefit from the presence of a 4G connection in both school subjects. However, this relationship is stronger and highly significant in Math scores. The presence of a 4G connection increase Math scores by 1.753 points for students with both high-educated parents, and by 1.336 points after including in the group also students with just one high-educated parent. In contrast, students with a low-educated family background are not affected by the 4G technology. An increase in connectivity seems to benefit more the performance of students with a high family education, increasing the existing gap with students with a low parental education.

Table 5.8: Regressions results with interaction effect between 4G and parental education.

		WLE_ITA_200		*WLE_MAT_200*	
		(1)	(2)	(3)	(4)
	4G	-0.380	-0.433	-0.262	-0.410
		(0.282)	(0.284)	(0.292)	(0.294)
	4G × Both_high_ed	0.307	-	1.753***	-
		(0.262)		(0.288)	
	4G × Atleast_high_ed	-	0.353**	-	1.336***
			(0.175)		(0.189)
	Constant	238.507***	237.052***	231.796***	232.673***
		(0.358)	(0.393)	(0.384)	(0.437)
Dummy	Student-level	YES	YES	YES	YES
	Family Birthplace	YES	YES	YES	YES
	Parental education	YES	YES	YES	YES
	Parental employment	YES	YES	YES	YES
	Observations	1 960 882	1 982 645	1 961 393	1 983 158
	R^2	0.212	0.212	0.189	0.189

Standard errors are clustered by school and found in parentheses. All results have both year and school fixed effects.
* p<0.10, ** p<0.05, *** p<0.01

Table 5.9 contains the outcomes of the regressions including the interaction effect between 4G and parental employment. The estimated coefficients show a strong positive relationship between 4G availability and test scores of students with both employed high-skilled parents. Although the effect is smaller, this relationship remains positive and strongly significant also looking at students with at least one high-skilled parent. However, students with both parents unemployed are negatively affected by the technology. Despite these findings are evident for both school subjects, the availability of 4G seems to impact more on students' performance in Math. In general, these results are very similar to those found in the regressions for ultra-fast broadband. Students with a high parental job situation are more advantaged by the increase in connectivity than others. However, the 4G technology positively affects a very large proportion of students, including those with both employed low-skilled parents.

Table 5.9: Regressions results with interaction effect between 4G and parental employment.

		WLE_ITA_200			*WLE_MAT_200*		
		(1)	(2)	(3)	(4)	(5)	(6)
	4G	-0.338	-0.801***	-0.211	-0.157	-0.846***	-0.005
		(0.282)	(0.306)	(0.281)	(0.293)	(0.317)	(0.291)
	4G × Both_high_skill	0.644**	-	-	1.534***	-	-
		(0.264)			(0.288)		
	4G × Atleast_both_low	-	0.848***	-	-	1.274***	-
			(0.162)			(0.169)	
	4G × Both_unemployed	-	-	-2.048***	-	-	-1.608***
				(0.383)			(0.405)
	Constant	238.17***	238.33***	237.86***	232.03***	232.34***	232.06***
		(0.351)	(0.363)	(0.328)	(0.381)	(0.401)	(0.360)
Dummy	Student-level	YES	YES	YES	YES	YES	YES
	Family Birthplace	YES	YES	YES	YES	YES	YES
	Parental education	YES	YES	YES	YES	YES	YES
	Parental employment	YES	YES	YES	YES	YES	YES
	Observations	1 923 022	1 946 700	1 923 022	1 923 563	1 947 245	1 923 563
	R^2	0.213	0.213	0.213	0.187	0.187	0.187

Standard errors are clustered by school and found in parentheses. All results have both year and school fixed effects.
* p<0.10, ** p<0.05, *** p<0.01

In conclusion, Table 5.10 presents the results when including both UBB and 4G availability between regressors, as well as their interaction. The outcomes confirm an insignificant relationship between UBB and language scores. The previously found statistical significance between 4G and Italian scores vanishes after including UBB in the regression. Looking at Math scores, the relationship between UBB and test scores remained positive and significant even in the presence of 4G technology. The results do not report evidence of an interaction effect between the two technologies. Having both UBB and 4G available does not seem to have an additional effect on students' test scores.

Table 5.10: Regressions results with interaction effect between UBB and 4G.

		WLE_ITA_200 (1)	*WLE_MAT_200* (2)
	UBB	-0.170	1.257**
		(0.555)	(0.635)
	4G	-0.431	-0.235
		(0.290)	(0.307)
	UBB × 4G	-0.101	-0.009
		(0.506)	(0.564)
	Constant	237.007***	231.318***
		(0.330)	(0.364)
Dummy	Student-level	YES	YES
	Family Birthplace	YES	YES
	Parental education	YES	YES
	Parental employment	YES	YES
	Observations	2 517 645	2 521 459
	R^2	0.210	0.183

Standard errors are clustered by school and found in parentheses. All results have both year and school fixed effects.
* p<0.10, ** p<0.05, *** p<0.01

6. Conclusions

In this study, we made a first step in evaluating the effect of broadband Internet on students' performance. In particular, the study looks for a correlation between the availability of ultra-fast broadband and the language and Math scores of Italian eighth-grade students, as measured by a standardized national test.

We find that the availability of ultra-fast broadband connection positively affects Math scores. However, the effect size we estimate is not large: the presence of UBB increases test scores by just 0.0326 standard deviations. Nevertheless, we can distinguish it from zero due to the strong statistical significance of the estimated coefficient. This evidence persists even after introducing the presence of another broadband technology, mobile 4G, between regressors. The study also finds differences in the effect size among different groups of students. Looking at students' gender, we show that the positive effect of UBB on Math scores is reinforced for females and weakened for males when compared to the pooled estimates. Also, male students are negatively affected by UBB in Italian scores. A possible explanation is that boys tend to perform more distracting activities on the Internet than girls. Furthermore, students were grouped based on their social status, looking at parental education and parental employment. Students with highly educated or high-skilled employed parents are more positively affected than others by the availability of UBB, increasing the existing achievement gap in both school subjects. However, the positive effect of UBB extends to a very large group of students: children with at least an employed parent, even if low-skilled, are still positively affected by the presence of UBB in Math scores. For Italian scores, the positive effect extends to all students with both employed parents, even if both low-skilled.

In general, the implications from these findings suggest that ultra-fast broadband may support students in their learning activities at home or at school, especially for scientific subjects, but we should not expect large positive impacts on grades or test scores. Also, students' achievement can be evaluated in several ways besides scores on standardized tests, and all benefits from fast Internet may not be directly measurable by academic outcomes. In conclusion, there are some limitations in our study that deserve consideration and suggest caution in the interpretation of the results. First of all, we used information on the availability of UBB connection in a certain geographical area, ignoring the real coverage in that area. Also, we do not know which schools are actually connected to the Internet using fiber-based technology, and we do not have information on the actual usage of broadband by students at school or at home. Furthermore, there may be an endogeneity problem: the adoption of UBB by schools or households is often an exogenous choice. The presence of an omitted variable that jointly influence the outcome of interest and the explanatory variable lead to an estimation bias. To overcome this problem, we suggest instrumental variables methods to estimate the causal effect.

7. References

Bresnahan, T., & Trajtenberg, M. (1995). "General purpose technologies: 'Engines of growth'?". *Journal of Econometrics*, 65, 83–108.

Czernich, N., Falck, O., Kretschmer, T., & Woessmann, L. (2011). "Broadband infrastructure and economic growth". *The Economic Journal*, 121(552):505–532

Rohman, I. K., & Bohlin, E. (2012). "Does broadband speed really matter as a driver of economic growth? Investigating OECD countries". *International Journal of Management and Network Economics*, 2(4), 336–356.

Forman, C., Goldfarb, A., & Greenstein, S. (2012). "The internet and local wages: A puzzle". *American Economic Review*, 102(1):556–75.

Akerman, A., Gaarder, I., & Mogstad, M. (2015). "The skill complementarity of broadband internet". *Quarterly Journal of Economics*, 130(4), 1781–1824.

Bai, Y. (2016). "The faster, the better? The impact of internet speed on employment". Working paper available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2772691.

Angrist, J., & Lavy, V. (2002). "New evidence on classroom computers and pupil learning". *Economic Journal*, pages 735-765.

Fuchs, T., & Woessmann, L. (2004). "Computers and student learning : bivariate and multivariate evidence on the availability and use of computers at home and at school". CESifo Working Paper No. 1321.

Rouse, C. E., & Krueger, A. B. (2004). "Putting computerized instruction to the test: a randomized evaluation of a "scientifically based" reading program". *Economics of Education Review*, 23(4):323-338.

Banerjee, A., Cole, S., Duflo, E. & Linden, L. (2007). "Remedying Education: Evidence from Two Randomized Experiments in India". *Quarterly Journal of Economics*, 122(3): 1235-1264.

Barrow, L., Markman, L., & Rouse, C. E. (2009). “Technology's Edge: The Educational Benefits of Computer-Aided Instruction”. *American Economic Journal: Economic Policy*, 1(1): 52-74.

Barrera-Osorio, F., & Linden, L. L. (2009). “The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia”. Policy Research Working Paper 4836, Impact Evaluation Series No. 29, The World Bank.

Grimes, D., & Warschauer, M. (2008). “Learning With Laptops: A Multi-Method Case Study”. *Journal of Computing Research*, 38(3): 305-332.

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): 211-240.

Goolsbee, A., & Guryan, J. (2006). “The Impact of Internet Subsidies in Public Schools”. *The Review of Economics and Statistics*, 88(2): 336-347.

Belo, R., Ferreira, P., & Telang, R. (2014). “Broadband in School: Impact on Student Performance”. *Management Science*, 60 (2): 265-282.

Bulman, G., & Fairlie, R. W. (2016). “Technology and Education: Computers, Software, and the Internet”. Working Paper 22237. National Bureau of Economic Research.

Grimes, A., & Townsend, W. (2017). “The Effect of Fibre Broadband on Student Learning”. Motu Economic and Public Policy Working Paper 17-03.

Hazlett, T. W., Schwall, B., & Wallsten, S. (2016). “The Educational Impact of Broadband Subsidies for Schools Under E-Rate”. SSRN Scholarly Paper ID 2796048. Working paper available at: https://papers.ssrn.com/abstract=2796048.

Silva, S., Milkman, M., & Badasyan, N. (2014). “The Impact of Brazils 'Broadband at School Program' on Student Achievement”. TPRC Conference Paper. Working paper available at: http://ssrn.com/abstract=2413976.

Tajuddin, M., & Rohman, I. (2018). “The Impact of Broadband Penetration on the Student Performance in Primary and Secondary School in Malaysia”. In Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance, Galway, Ireland, April 2018 (ICEGOV’18), 7 pages.