8-2012

Broadband in Schools: Effects on Student Performance and Spillovers for Household Internet Adoption

Rodrigo Belo
Carnegie Mellon University, rbelo@cmu.edu

Broadband in Schools: Effects on Student Performance and Spillovers for Household Internet Adoption

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Engineering and Public Policy

Rodrigo Belo

B.S., Computer Science and Engineering, IST - Technical University of Lisbon
M.S., Engineering and Public Policy, Carnegie Mellon University

Carnegie Mellon University
Pittsburgh, PA

August 2012

For my father.

Abstract

This work comprises studies on the effects of broadband Internet in schools at three different levels: student performance, household Internet adoption, and individual computer and Internet use patterns and skill acquisition. I focus in the case of Portugal, where by 2006 the Portuguese government had completed a major initiative that upgraded the Internet connection of all public schools, replacing the previously existing connections by broadband.

In the first study I focus on the direct effects the introduction of broadband in middle schools had in students' performance. I find that high levels of broadband use in schools are detrimental for 9^{th} grade national exam scores. For the average broadband use in schools, exam scores reduce about 0.97 of a standard deviation from 2005 to 2009. I also find suggestive evidence that the way schools allow students to use the Internet affects students' performance. In particular, students in schools that block access to websites such as YouTube perform relatively better.

In the second study, I look at spillover effects of providing broadband to schools in home Internet adoption. I assess the magnitude of these effects using household level data on home Internet penetration and Internet traffic in all schools in Portugal. I find that school broadband use contributes directly to a higher adoption rate in households with children. During 2008 and 2009 school Internet use increased the probability of adopting Internet by 20% in households with children, while no statistically significant effect was found in households without children.

In the third study I focus on Information and Communication Technology (ICT) skills and on the dynamics of computer and Internet use inside the household. I provide empirical evidence that the presence of children or young adults in the household does contribute to an increase in the likelihood of having a computer or Internet at home, but does not contribute to an increase in use patterns and skills. Moreover, I find that the presence of children and young adults is associated with lower levels of computer and Internet use and skills.

Acknowledgments

I would like to acknowledge the financial support provided by Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the Carnegie Mellon Portugal Program under Grant SFRH/BD/35680/2007 and Research Contract NGN/56.

I would like to thank Professor Pedro Ferreira and Professor Rahul Telang, my advisors, for all their support and guidance, for the lively discussions, and for all the time dedicated to this research.

I would like to thank the members of my Ph.D. committee: Professor Pedro Ferreira (chair), Professor Rahul Telang, Professor Marvin Sirbu, and Professor Chris Forman, for their interest in my research, for insightful discussions and for their questions and suggestions that contributed greatly to improve the quality of this work.

I would like to show my appreciation to all Professors in the Carnegie Mellon Portugal Program in Technology Change and Entrepreneurship, from whom I learned immensely, and for their availability whenever needed.

I would like to show my gratitude to all my colleagues in the Ph.D. Program in Technology Change and Entrepreneurship, and in a special way to my cohort, Ana, Andreia, Carla, and Cristina, for their company, support and friendship.

I would like to leave a special word to my family. I would like to thank my brother Gil, my sister Cláudia, my father Salvador and my mother Maria Manuela, that have always encouraged me, and that supported each other during these last few difficult years.

Finally, I would like to thank my girlfriend, Sónia Pedrosa, who has been present during the good and bad times, and that helped me in virtually everything along these five years. I would like to thank my friends, for making this process easier, for all the good moments spent together, and for their words of encouragement. Without forgetting many others, I leave a special word of appreciation to Alexandre Mateus, Alexandre Ribeiro, Marco Almeida, João Paulo Bogalho and Fernando Martinho.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The use of information and communication technologies (ICTs) in schools, namely computers and broadband Internet access, has once again revived the idea that technology will significantly change education, namely by empowering teachers and students with access to online information, educational resources, and interactive learning tools that provide timely feedback and that engage students in learning activities. Accordingly, a growing number of federal government and state programs have been subsidizing the deployment of ICTs in schools. One such example is the case of the E-Rate program in the US, that comprised an investment of $2.25 billion per year in subsidies to schools and libraries for Internet and communications technology (Hudson, 2009). In 2003 more than 80% of students in the elementary and secondary education in the US used computers at school, and by 2008 the rate of students per instructional computer was 3.1 (Snyder and Dillow, 2011). This trend is visible in many other countries, and similar programs have been implemented in places such as Israel (e.g., Angrist and Lavy, 2002), the Netherlands (e.g., Leuven et al., 2007), Romania (e.g., Malamud and Pop-Eleches, 2011), and Portugal (e.g., Belo et al., 2012c).

Providing ICT resources to schools can have important implications not only to schools — by changing their working processes, the way classes are conducted, the resources available, and the way students use these resources — but also to the neighboring communities, especially in rural areas where contact with the Internet is often limited. In these areas the introduction of broadband in schools can be considered as a catalyst for household Internet adoption. On the one hand, the new infrastructure that is often put into place to meet schools' needs can also be used to serve households. On the other hand, students get acquainted with the technology and signal the value of Internet to other family members who can, as a consequence, also adopt.

However, evaluating the benefits of these investments is hard, not only because it is difficult to collect robust and consistent measures of impact, but also due to endogeneity, which often casts doubts on the causality of the relationships under analysis (see Webbink, 2005, for a detailed explanation of the endogeneity problem in studies that try to assess the relationship between education inputs and students performance). Thus, well-founded, good quantitative analyses of the impacts of this technology are still in need (see Hauge and Prieger, 2010, for a review of demand-side programs to promote Internet adoption and their evaluation).

I focus in the case of Portugal, where by 2006 the Portuguese government had completed a major initiative that upgraded the Internet connection of all public schools, replacing the previously existing ISDN connections by broadband ADSL. My goal is to quantify the effects of providing broadband in schools, both in terms of students performance and in terms of household Internet adoption and use in the neighboring communities.

In my studies I take advantage of how broadband technology works to alleviate concerns with endogeneity. In particular I use the fact that the DSL throughput decreases with distance between the schools and the ISP central office (CO) to derive an exogenous instrumental variable that explains usage at schools and is somewhat independent of regional characteristics as well as of my dependent variables. Using this technical knowledge is fundamental to help establish causality.

In my first study I focus on the direct effects the introduction of broadband in middle schools had in students' performance. The main idea is that Internet, as a new resource students can use, brings about lots of opportunities for learning, but also for distraction. I find that high levels of broadband use in schools are detrimental for 9^{th} grade national exam scores. For the average broadband use in schools, exam scores reduce about 0.97 of a standard deviation from 2005 to 2009. I also find suggestive evidence that the way schools allow students to use the Internet affects students' performance. In particular, students in schools that block access to websites such as YouTube perform relatively better.

In my second study, I look at spillover effects of providing broadband to schools in home Internet adoption. I develop a structural model that provides insight on how Internet use at school affects home Internet penetration and how Internet penetration affects school Internet use. I identify three potential sources of school to home spillovers: (1) change in children's utility, i.e., children that use Internet at school might not feel the need to use it at home, or might want to use it even more at home; (2) children information transmission to the adult, i.e., children learn about the technology in school and bring that knowledge home, which makes their parents aware of the value of having Internet; and (3) neighborhood level effects, i.e., adults observe their neighbors adopting Internet at home, learn about the technology, and decide to adopt. I assess the magnitude of these effects using household level data on home Internet penetration and Internet traffic in all schools in Portugal. I find that school broadband use contributes directly to a higher adoption rate in households with children. During 2008 and 2009 school Internet use increased the probability of adopting Internet by 20% in households with children, while no statistically significant effect was found in households without children.

In my third study I focus on ICT skills and on the dynamics of computer and Internet use inside the household. More specifically, I assess whether there is transmission of ICT-related knowledge among household members. I make use of a detailed individual level and household level survey on computer and Internet use and skills. I provide empirical evidence that the presence of children or young adults in the household does contribute to an increase in the likelihood of having a computer or Internet at home, but does not contribute to an increase in use patterns and skills. Moreover, I find that the presence of children and young adults is associated with lower levels of computer and Internet use and skills. A potential explanation for this that children and young adults monopolize the use of the computer and Internet, leaving older adults with less time to use the technology. Another explanation is the possibility that adults with children have less time to spend on the Internet. Finally, I can not rule out the possibility that selection bias is driving these results, despite the efforts to control for this phenomenon.

Chapter 2

The Effects of Broadband in Schools on Students' Performance

Abstract: The introduction of broadband in schools provides a new resource for learning but also an opportunity for distraction. Consequently, broadband use in schools can either increase or reduce students' performance. This paper provides a model that shows how these two effects trade off. We use a rich panel of data with information on broadband use and students' grades from all middle schools in Portugal to learn how broadband use affects performance. We use a first-differences specification to control for school-specific unobserved effects. We also use a proxy for the quality of broadband as an instrument to control for unobserved time-varying effects. We show that high levels of broadband use in schools are detrimental for grades in the 9^{th} grade national exams. For the average broadband use in schools, grades reduce about 0.97 of a standard deviation from 2005 to 2009. We also show evidence suggesting that broadband has a negative impact on exam

scores regardless of gender, subject or school quality. We also find that the way schools allow students to use the Internet affects students' performance. In particular, students in schools that block access to websites such as YouTube perform relatively better. Although test scores do not measure all the effects that broadband in schools have on the performance of students throughout life, our results show that different policies that schools may enact with respect to Internet use may result in different outcomes and therefore the introduction of Internet in schools is a task that deserves careful planning.

2.1 Introduction

The role of ICTs on our economy can hardly be overemphasized. There is a great amount of literature in both the IT and the economics communities on how computers and Internet affect firm productivity (Brynjolfsson and Hitt, 1996; Forman et al., 2005). However, the role of technologies on education is also an important policy and managerial issue. It has received much less attention in IS literature though. Predominantly, this has been a domain of research for economists and sociologists. As we will show below, even in this stream of literature, the role of ICTs in eduction is hardly settled. In this paper, we bring an interesting and detailed dataset to propose a convincing method to tease out how Internet and broadband affects students' grades in schools.

ICTs are perceived by many as potential powerful tools to improve the quality of education. They facilitate real time access to information, provide a more hands-on learning experience and foster new learning methods that promote more interaction and feedback, ultimately increasing students' interest and performance (e.g., Underwood et al., 2005).

Governments around the world are heavily subsidizing computers and now broadband access in schools. However, the Internet also offers significant opportunities for students to indulge in leisure and entertainment activities. Without effective monitoring and controls by schools, students may predominantly use broadband to play games, chat and watch movies. This can distract them from traditional study which can ultimately hurt the productivity of learning at school. In fact, some studies indicate that children spend considerable amounts of time playing computer games (Malamud and Pop-Eleches, 2011). It is also quite likely that teachers may find it hard to effectively use ICTs as part of the curriculum. Despite the large investments in computers and Internet access in schools, there are only a few studies that examine the impact of Internet in schools on students' performance.[1] Moreover, these studies provide mixed results on whether ICTs indeed help students. Thus there is little understanding of how broadband can help learning.

In this paper, we first provide a model for how broadband use in schools contributes to students' performance. We then provide empirical evidence on the impact of *actual usage* (as opposed to the simple existence of a broadband connection in schools) of broadband in schools on students' performance drawing from the case of Portugal. Actual usage is measured by the amount of information exchanged with the Internet over ADSL connections. Performance is measured by scores obtained in the 9^{th} grade national exams. We collect a panel of data on broadband use and school performance in more than 600 Portuguese middle schools, between 2005 and 2009. We use a first differences model to account for

[1]To the best of our knowledge Goolsbee and Guryan (2006) is the only study that directly measures the impact of school Internet availability on students' performance. They find no evidence that wiring classrooms affects students' grades.

school-specific unobserved effects. Still, the school performance may be endogenous to broadband use. We overcome this by instrumenting the schools' broadband use with the distance between the school and the provider's Central Offices (COs), which proxies the quality of the ADSL connection. Distance has some unique and desirable properties for a good instrument providing us confidence in the results obtained.

Our estimates indicate that more broadband use is detrimental for students' test scores. We find that, on average, grades declined about 0.97 of a standard deviation between 2005 and 2009 due to broadband use. We find that there is little difference across genders (although boys seem to be slightly more affected) and across math vs language (math grades seem to be slightly more affected). In addition, schools are equally affected by Internet use regardless of their performance prior to the deployment of broadband.

To explore the distraction effect of Internet in more detail, we conduct a survey to understand how Internet is utilized in schools. In particular we focus on the policy of schools regarding blocking or allowing applications and services such as Facebook, YouTube and file-sharing, which are likely to cause distraction. We find evidence that schools that allow these activities perform worse and the effect of Internet is significantly more negative when schools allow YouTube use. This result suggests that without proper monitoring and control, broadband access in schools may be more harmful than helpful.

2.2 Related work

Economists have been interested in how school resources like class size, school hours, teacher training, peer group and so on affects student performance. However, teasing out these effects is quite challenging. Concerns about endogeneity cast doubts on the causality of the relationship between education inputs and students performance (see Webbink, 2005, for a detailed explanation of the endogeneity problem in these studies).

Some of the more recent studies overcome the endogeneity problem in different ways and find a positive impact of class size (e.g., Krueger, 1999; Angrist and Lavy, 1999), school hours (e.g., Lavy, 1999) and peer group effects (e.g., Sacerdote, 2001).[2] The impact of other characteristics, such as teacher training and computer use, either remains non-significant or exhibits mixed results (e.g., Angrist and Lavy, 2002; Webbink, 2005; Barrera-Osorio and Linden, 2009).

Most studies look at students' test scores in a standardized test as an outcome measure (e.g., Angrist and Lavy, 2002; Goolsbee and Guryan, 2006; Leuven et al., 2007; Machin et al., 2007). Even though test scores have some obvious limitations, they are used mainly because they are reliably measured, and provide a tangible and standard way to measure student performance. Test scores are also a barometer used by policy makers and administrators to assess a school's performance which affects teacher benefits, school subsidies and parents' demand. As a consequence, schools, teachers and students all have incentives to improve test scores.

Research on the contribution of ICTs to students' performance has produced mixed results. Early studies on the use of computers in the classroom report positive effects on students' performance, but are often criticized either because they fail to account for endo-

[2] All these studies take advantage of an exogenous source of variation to overcome the endogeneity problem. For example, Krueger (1999) use an experimental setting; Angrist and Lavy (1999) take advantage of a maximum class size rule; Lavy (1999) taps on variations on the allocation of school hours; and Sacerdote (2001) uses random dorm assignments.

geneity or because they report effects with small magnitudes (see Cuban and Kirkpatrick, 1998; Webbink, 2005, for critical assessments of these studies).

Angrist and Lavy (2002) exploit a randomization (determined by a lottery) in the timing of school computerization in Israel. They find no effect on students performance, except for a negative effect in math exam scores for 8^{th} graders. Goolsbee and Guryan (2006) study the impact of subsidizing schools' Internet access in the U.S. and find no evidence that more classrooms with Internet has an impact on students' performance, as measured by the Stanford Assessment Test (SAT). Leuven et al. (2007) exploit a discontinuity in a subsidy given to schools in the Netherlands. Using a differences-in-differences framework, they find that this subsidy had a negative impact on students' performance, especially on girls. Malamud and Pop-Eleches (2011) exploit a discontinuity in a subsidy provided in Romania in 2008. This subsidy would allow low-income families to acquire a home computer. They find that the students of families that used this subsidy (households that indeed bought a home computer) had significant lower school grades in math, English and Romanian. They also find that these students had higher scores in tests of computer skills and in self-assessment tests of computer fluency. Vigdor and Ladd (2010) use fixed-effects to estimate the impact of home computer and Internet access on students' performance in North Carolina. They use a panel on the state's public school students between 2000 and 2005 and find a small but statistically significant negative effect of home computer access on students' math and reading test scores. They also report a decrease in 3% of a standard deviation in male reading test scores associated with home Internet access.

An exception to this recent trend of non-significant or negative results is provided by

Machin et al. (2007) who use rule changes in UK to find evidence of a positive effect of ICT investment on educational outcomes in elementary schools.

Similarly, the effects of computer-aided learning or softwares on students' performance is also ambiguous (see Rouse and Krueger (2004), Banerjee et al. (2007) Barrow et al. (2009)). In some cases the effects are positive but in some other cases computer-aided learning tools make no difference. Their effectiveness also varies between math and readind and boys and girls.

In summary, the impact of ICTs is an empirically challenging question. Also, most studies published so far look at the impact of investment in ICTs on student's performance and not at the impact of actual usage of ICTs. Furthermore, most of these studies look at the availability of ICTs in general rather than the use of a specific technology. This paper looks at the impact of *actual broadband* use on a real school environment. We also examine the impact of specific applications and services. We provide a credible instrument to alleviate the endogeneity concerns. Overall, we find that broadband usage over the 2005-2009 period had an adverse effect on the performance of 9^{th} grade students in Portuguese schools.

2.3 Broadband in Portuguese Schools

2.3.1 Broadband Provision to Schools

In Portugal most elementary and secondary schools are public schools, funded either by the Central Government or the Local Government, with limited autonomy to manage their resources. The provisioning of Internet to schools has been managed by FCCN - the Portuguese National Foundation for Scientific Computation.

Portuguese government has taken many initiatives to connect schools in computers and Internet. For example, by mid 2001, all elementary and middle schools in the country had been connected with an ISDN connection and had atleast one computer in schools. (FCT, 2001). In 2004, the same Ministry launched another major initiative, this time aimed at replacing all the existing ISDN connections by broadband ADSL. By 2006, Most schools (>95%) received a DSL modem from FCCN and an ADSL connection of at least 1 Mbps over the copper line that connects them to the Central Office (CO) of Portugal Telecom (PT), the ISP from which FCCN buys connectivity to the Internet (Figure 2.1). This was a decision by the central government and schools did not have a say on whether they wanted a broadband or not. In other words, some (most) schools might not have been prepared to receive broadband at the time. Anyhow, the Ministry covered all up-front capital costs to deploy broadband to schools, as before. City Halls foot the broadband monthly bill for elementary schools while the Ministry covers these costs for the remainder of the schools.

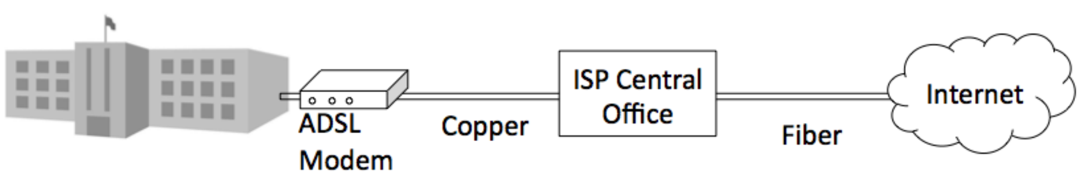


Figure 2.1: Broadband schools' connection to the Internet. Schools connect through a copper line to the ISP's central office. From there, the ISP ensures connectivity to the Internet backbone through fiber.

There is no information about whether some schools had already purchased broadband from the market by the time this intervention took place, but the schools' tight budgetary constraints must have allowed only a small fraction of them to do so, if at all. More importantly, FCCN strongly encouraged schools to use the broadband connection provided by the Government, after all traffic over this broadband connection is free of charge to schools, so even if some schools had bought a DSL connection before, they had a strong incentive to shut it down and use only the FCCN's connection. Therefore, the broadband use over the Internet connection provided by FCCN seems to be a good measure for the school's overall broadband use.

2.3.2 Internet Use at School

We conducted preliminary informal interviews with teachers in eight different schools to learn more about how Internet is used in schools. Some teachers are comfortable with using ICTs in the classroom and consider the Internet a good tool to capture the students' interest and to improve the learning process.[3] Other teachers look at the Internet as just another resource that students can possibly use for learning. However, not all teachers felt that Internet always provides easy to use information.[4] Differences in skills and in the attitude of teachers towards the Internet translate into significant differences on how much students use the Internet in the classroom.

School-specific Internet access policies may also explain part of the differences in the pattern of Internet use across schools. While some schools provide an open wireless network that any computer can tap into, such as students' laptops, other schools disallow access

[3] Some of the teachers interviewed referred that students engage more in discussions and are more motivated when Internet is used in class.

[4] One of the teachers interviewed pointed out that he had difficulty in explaining to students that Wikipedia is not a reliable source of information and that they should always check their sources.

to their wireless network to all but school computers. Some schools block access only to a restricted set of web sites (mainly adult content sites), while other schools block access to a whole range of sites considered inappropriate in the school context.[5] All these factors influence how students use the Internet at school and, consequently, their incentive to bring their laptop to school. Students in some schools bring their laptops several days a week to school and use them pervasively, which facilitates using a wide range of applications such as social media and video streaming, while in other schools students seldom make use of their own laptops.

The time that students spend at school is also heterogeneous. In some schools students usually stay at school after class time, while in other schools most students leave school right after classes. Most students that stay at school after hours often do so to use the school's computers and the Internet, most likely, in some unsupervised way.

Finally, students that do not have Internet at home are likely to exhibit different usage patterns than those who do. On the one hand students that only access the Internet at school might develop a more mature approach to use it because they learned how to navigate the Internet under the teachers' supervision. Students that have Internet at home might know better how to use it for recreational purposes and carry that practice to school. However, it might also be that students that use the Internet at home for recreational activities do not need to do so at school and thus indulge in learning activities while at school. All in all, there is a wide variation across schools in terms of how students use the Internet. Teacher knowledge and attitude towards the use of ICTs in the classroom, school's

[5] Video, chat, social network and adult content sites are among the categories most often blocked. In the later section of the paper, we provide more details on this.

Internet access policies, time spent at school after classes and the number of students that access the Internet, both at school and elsewhere, are some of the factors that contribute to such a variation.

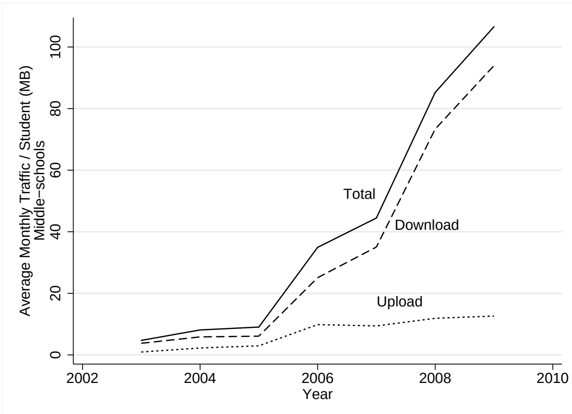
2.4 Data

School traffic data were obtained from the monitoring tools set up by FCCN. From the ISDN project, we obtained data for all ISDN sessions between November 2002 and January 2005 for all schools in the country. From the ADSL project, we obtained monthly reports that include download and upload traffic per school between November 2005 and June 2009. School traffic is measured at the school's edge router and consists of all traffic exchanged between the school and the Internet. For our measure of school broadband use, we average out the total monthly traffic (upload plus download) over the entire academic period.[6]

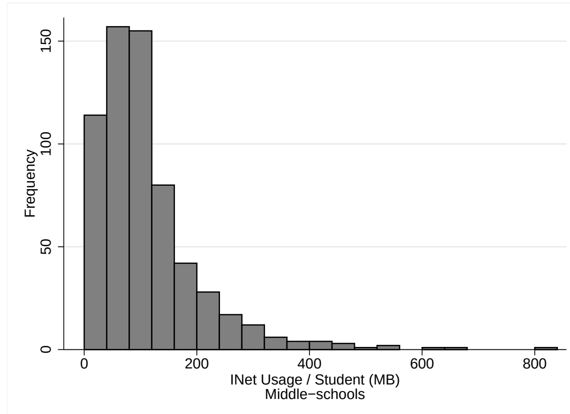
Internet use in schools grew significantly since the introduction of ADSL in late 2005. Before 2006, Internet use was virtually zero, compared to usage levels in 2008 and 2009 (see Figure 2.2(a)), probably because the ISDN connections could not carry more traffic. Inbound traffic is the major contributor for this increase; outbound traffic remains relatively little across most schools. Broadband use per student exhibits high variability across schools (see Figure 2.2(b) for a histogram). In 2009, students used 111 MB at school per month on average, which corresponds to watching almost one hour of YouTube video (at 260 Kbps), browsing 350 webpages (at 320 KB per page), or exchanging 850 emails (at

[6]We use as academic year the period between September and June.

130 KB per email).[7] There is significant heterogeneity in usage (large standard deviation (95 MB)).



(a) Internet traffic between 2003 and 2009.



(b) Monthly average Internet use per student in 2009.

Figure 2.2: Middle school Internet traffic and monthly average Internet use per student in 2009.

Performance is measured by the school's average score at the 9th grade national exams. The Ministry of Education publishes anonymous disaggregated data at the exam level since 2005, including information on exam score, course, gender, and age of the examinee. 9th graders are examined in two subjects, Portuguese and math, and their exam scores constitute part of their final score on these subjects and might determine whether the student graduates. Therefore, students have clear incentives to perform well in the 9th grade national exams.[8]

Figure 2.3 shows 9th grade average exam scores in terms of 2005 standard deviations.[9] Average exam scores have increased from 2005 to 2009 (14.0%), which may reflect a positive

[7] Average webpage size was obtained from *http://code.google.com/speed/articles/web-metrics.html*. We use the average email size of one of the authors as reference, as we found no reliable information on this statistic.

[8] Even though this is a standardized exam, it is not necessarily a multiple choice or binary response only exam. The students have to write detailed answers.

[9] 9th grade exam scores are published in a 1-5 scale (with increments of 1).

impact of broadband on students' performance. Alternative explanations for this rise include unobserved factors, such as exams becoming easier with time, particularly in 2008.

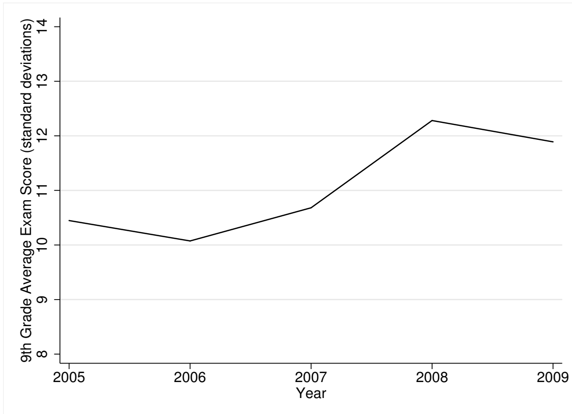


Figure 2.3: 9$^{\text{th}}$ grade average exam scores between 2005 and 2009.

Finally, we obtained GPS coordinates for all the COs of PT (Portugal Telecom), the historic operator in Portugal, which held a market share of about 70% in the Portuguese broadband market during the years of our study (ANACOM, 2010). Furthermore, we obtained from PT the average monthly traffic rate per CO for residential Internet access. Regional data were provided by the Portuguese National Statistics Institute. These data include population density (2001 census data; at the civil parish level), average earnings and regional dropout rates (2005; at the municipality level) across municipalities. Table 2.1 presents summary statistics of these variables for schools in our sample.[10] School enrollment was obtained from the Ministry of Education.[11]

[10]Portugal has a population of 10.6 million. The country is divided into 308 municipalities, which are further divided into 4,261 civil parishes. Schools in our sample cover 204 municipalities and 547 civil parishes.

[11]We were able to obtain student enrollment data only for 2007. We use these 2007 values for the whole time period as the number of students in a school is unlikely to change much from year to year.

Table 2.1: Summary statistics.

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Avg. Grade 2009 (s.d.)	628	11.59	1.135	7.888	15.21
Avg. Grade 2008 (s.d.)	628	11.97	1.117	7.898	15.88
Avg. Grade 2005 (s.d.)	628	10.20	1.008	7.185	13.80
INet Usage 2009 / Stu. (MB)	628	111.2	95.32	4.22e-04	800.5
INet Usage 2008 / Stu. (MB)	628	86.70	97.42	0.123	1,766
Students	628	579.3	239.2	72	1,412
Pop. Density	628	1,820	2,868	5.800	20,648
Earnings 2005	628	787.0	186.8	532.8	1,487
Mandatory Educ. (%)	628	39.14	13.73	10.38	80.05

2.5 Framework

We first introduce a simple model that explains how the time students spend using the Internet at school affects their performance. Let P represent students' performance. Let I represent the time they spend using the Internet at school. Let S represent the time they spend at school without using the Internet, otherwise hereinafter called traditional study time at school. Let $T = I + S$ represent the total time students spend at school. We assume that the total time students spend at school remains unchanged with the introduction of Internet in the school.

The performance of students depends on the effectiveness of the time they spend using the Internet at school and on the effectiveness of the time they dedicate to traditional study at school. Therefore, define $P = f(I, S)$, where f is a production function. All else being equal, more of one input cannot reduce output, thus we have $f_I \geq 0$ and $f_S \geq 0$.

The effect of Internet use in school on students' performance is given by

$$\frac{dP}{dI} = f_I + f_S S_I = f_I + f_S(T_I - I_I) = f_I + f_S(0 - 1) = f_I - f_S.$$

At school, time on the Internet substitutes traditional study time without the Internet. The productivity of Internet time at school (f_I) trades off with the productivity of traditional study time (f_S) and thus performance can either increase or decrease when Internet is introduced in schools.

Furthermore, split Internet time at school into learning time, L, and distraction time, D, and make $I = L + D$. We also have $\partial L/\partial I \geq 0$, that is, all else being equal, more time on the Internet does not reduce learning time. Likewise for distraction and thus $\partial D/\partial I \geq 0$. These statements, together with $I = L + D$, imply that $\partial L/\partial I \leq 1$.

Consider now that the students' performance depends on the effectiveness of the time they spend learning on the Internet at school and on the effectiveness of the time they dedicate to traditional study at school. Therefore, define $P = g(L, S)$, where g is a production function. As before, we have $g_L \geq 0$ and $g_S \geq 0$.

In this case, and using the fact that $T = S + I$ is constant, the effect of Internet use at school on students' performance is given by

$$\frac{dP}{dI} = g_L \cdot \frac{\partial L}{\partial I} - g_S.$$

The productivity of learning with the Internet (g_L) weighted by how Internet time is devoted to learning ($\partial L/\partial I$) trades off with the productivity of traditional study time at school (g_S). Note that $g_L \cdot \partial L/\partial I \geq 0$ and $g_S \geq 0$ and thus, again, the introduction of

Internet in schools can either increase or decrease performance. In fact,

$$sgn\,[dP/dI] = sgn\left[\frac{g_L}{g_S}(\partial L/\partial I) - 1\right].$$

The impact of Internet at school on students' performance (dP/dI) is positive when the relative productivity of learning time on the Internet at school to the productivity of traditional study time at school (g_L/g_S), weighted by how Internet time is devoted to learning ($\partial L/\partial I$), is greater than one. One may expect that learning with the Internet may be more productive than traditional study ($\frac{g_L}{g_S} > 1$). Even then, our model highlights that the impact of Internet is critically affected by how Internet time is devoted to learning. Even if $g_L > g_S$, only if $\partial L/\partial I$ is large, that is, only if students are largely using the Internet for learning purpose, we could expect their performance to improve.

Consider a CES production function

$$p = [\beta L^r + (1-\beta)S^r]^{1/r},$$

with $0 \leq \beta \leq 1$ and $r \leq 1$. Differentiating with respect to I yields

$$sgn\,[dP/dI] = sgn\left[\gamma(L/S)^{r-1} \cdot \partial L/\partial I - 1\right],$$

where $\gamma \equiv \beta/(1-\beta)$. In this case, $\gamma(L/S)^{r-1}$ is the relative productivity of learning time on the Internet to traditional study time. For the case of a linear production function ($r = 1$), the effect of Internet use in school is given by $\gamma(\partial L/\partial I) - 1$. Furthermore, if students devote a constant share of the time they spend on the Internet at school to learning activities, call it α ($\alpha \equiv \partial L/\partial I$), then the effect of Internet use in school is constant and given by

$$\frac{dP}{dI} = \gamma\alpha - 1. \tag{2.1}$$

Or, in other words, the impact of Internet depends on how effective it is relative to standard study and how much time students actually devote to learning activities.

2.6 Empirical Specification

2.6.1 First-Differences Model

School performance is assumed to depend on broadband use, on socio-economical factors, such as average earnings, population density and percentage of people with mandatory level of education, and on school-specific unobserved factors, such as the quality of teachers and the comfort and size of the classrooms. Therefore, school performance can be expressed by the following structural equation

$$P_{it} = \delta + \omega I_{it} + \mathbf{x}_i\beta + \mathbf{w}_{it}\theta + c_i + u_{it} \tag{2.2}$$

where P_{it} represents the performance of school i at time t; $\omega \equiv (\gamma\alpha - 1)$ is the effect of Internet use on school performance (see Equation 2.1), our parameter of interest; I_{it} represents broadband use; $\mathbf{x}_i$ and $\mathbf{w}_{it}$ are row vectors with time-fixed and time-varying school- and region-specific control variables. We include, as time-invariant control variables, school size (measured by the number of students in each school), population density, earnings in 2005, and the percentage of people with mandatory level of education in 2001 in the municipality where the school is located. As time-varying control, we use average Internet traffic rate per person (in Mbps per capita) at the school's closest ISP's Central Office (CO). This variable is used as a proxy for home Internet use in the region where the school is located. β and θ are parameter vectors; c_i is an unobserved time-constant school specific effect; and u_{it} is a random error term.

This is the classic fixed-effects specification. Specifying a separate dummy for each school in the form of c_i allows for controlling for school-specific unobserved factors. Alternatively, we can write this as a differences model as

$$\Delta P_{it} = \phi + \omega \Delta I_{it} + \Delta \mathbf{w}_{it}\theta + \Delta u_{it}. \tag{2.3}$$

where ϕ captures the average change in exam scores over the period of analysis. For example, a $\phi > 0$ captures the fact grades increased because, suppose, exams became easier. Δ represents the difference between period t and 2005 (e.g., $\Delta P_{it} \equiv P_{it} - P_{i2005}$). Differences in broadband use in schools over one single academic year are likely to have little impact on that year's exam scores, if at all. Therefore, we use differences to 2005 to capture the accumulated effect of broadband use on performance. We also assume Internet use to be zero in 2005 and thus substitute ΔI_{it} by I_{it}.[12]

We pool all the differences into one regression and include year-dummies to control for

[12]Broadband was brought to schools during the second half of 2005. Thus it is safe to assume there was no broadband use for most of 2005, and in fact Internet use in 2005 is negligible when compared to 2008 and 2009 levels. In any case, we have used the exact differences in our regressions and obtained similar qualitative results.

period-specific variation. Additionally we cluster the standard errors at the municipality level to account for possible correlation among observations in the same municipality.

Note that the term $\mathbf{x}_i\beta$ in equation (2.2) gets differenced out because it corresponds to time constant factors. However, to account for the possibility that some school-specific variables in $\mathbf{x}_i$ might also drive the change in performance and in broadband use, we include the baseline values of $\mathbf{x}_i$ as additional controls:

$$\Delta P_{it} = \phi + \omega \Delta I_{it} + \mathbf{x}_i\beta + \Delta \mathbf{w}_{it}\theta + \Delta u_{it}. \tag{2.4}$$

This is equivalent to adding an extra term $d_{2005} \cdot \mathbf{x}_i\beta$ to our structural equation where d_{2005} is an indicator variable such that $d_{2005} = 1$ for the year 2005. This allows for the possibility that $\mathbf{x}_i$ influences not only performance but also its rate of change.[13]

2.6.2 Identification

Despite the first-differences setting and the controls in $\mathbf{x}_i$, potential unobserved *time-varying* factors may lead to both increased broadband use and better (or worse) exam scores, resulting in inconsistent estimates for ω. For example, a change in the resources available to a school[14], internal organization or technical savviness, might have influenced both broadband use and exam scores during the period of analysis. The school-specific dummies do not capture these time-varying unobserved effects and therefore our estimates might become inconsistent.

[13]Our results are similar whether or not we include $\mathbf{x}_i$ as controls. We leave these controls in the differences equation for generality.

[14]During the period of analysis students were awarded laptops, under a parallel Governmental program. This may have changed both broadband usage patterns and scores.

We ensure identification by exploiting the variation in the quality of broadband connections across schools as an exogenous source of variation in our setup. Schools that benefit from a better connection to the Internet are more likely to use it more and therefore more likely to register more traffic. With ADSL technology, a greater distance between the costumer's premises and the ISP's Central Office (CO) results in a lower maximum transfer bitrate. Therefore, schools further away from the CO are likely to get less throughput on their connection. Such lower throughput leads to degraded performance decreasing the attractiveness of the broadband connection at the school and thus lowering the amount of traffic exchanged with the Internet. Consequently, we use line-of-sight distance between each school and its closest CO as a proxy for the quality of the school's broadband connection.[15]

Distance is an attractive choice for the instrument because one expects that the distance between the schools and the CO would be fairly randomly distributed; schools and COs have been around for much longer than broadband. The population in Portugal is fairly densely distributed. Therefore, unlike the US where one would worry about rural schools being systematically farther from the CO than urban schools, Portugal is more homogeneous: most schools are within 2 Km from a CO (see Figure 2.4) and there is little difference in distance of a urban school vs the rural school.

In Table 2.2 we provide the correlation matrix with distance and socio-economic characteristics for middle schools. Distance does not seem to be correlated with any of the socio-

[15]Line-of-sight distance is calculated from information on the GPS coordinates of both schools and the ISP's COs. We obtain similar results when using walking distance between the schools and CO, as calculated by Google Maps.

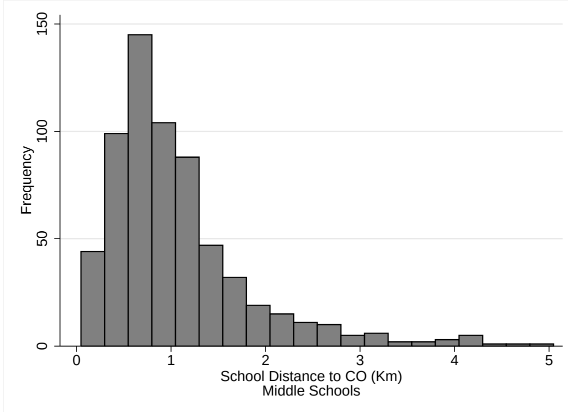


Figure 2.4: Middle Schools' distances to the closest CO.

economic characteristics or with grades before the deployment of broadband in schools. This strengthens our intuition that distance to CO seems to be fairly independent of specific regional characteristics. Figures 2.5, 2.6 and 2.7 in the appendix offer more details on relationship between distance and demographic characteristics.

Table 2.2: Cross-Correlations for Middle Schools.

Variables	Dist.	Avg. Grade '05	Stud.	Pop. Dens.	Earn. '05	Mand. Educ.
Avg. Grade 2005	-0.092					
Students	0.034	0.161				
Pop. Density	-0.030	0.055	0.323			
Earnings 2005	-0.023	0.126	0.095	0.496		
Mandatory Educ. (%)	-0.106	0.262	0.401	0.521	0.579	
Avg. CO Traffic 2005 (Mbps)	0.121	-0.040	0.090	-0.047	-0.124	0.011

We also test whether distance explains grades before the deployment of broadband in schools by regressing average school grades in 2005 on distance and other covariates for middle schools. Table 2.3 presents the results.

Distance to CO is statistically and economically insignificant in Table 2.3 suggesting that school grades are not affected by distance.[16] Thus schools that perform better or worse are not systematically located closer or further from the CO. These facts suggest

[16]This coefficient on distance would yield a reduction of 0.08 of a standard deviation in grades per kilometer, but again distance is not statistically significant in this regression.

Table 2.3: Average score in 2005 as a function of distance and other controls (OLS).

	(1)
VARIABLES	Avg. Grade 2005
Distance (Km)	-0.0760
	(0.0473)
Students (x 1000)	0.299
	(0.238)
Pop. Density (x 1000)	-0.0366**
	(0.0168)
Earnings (x 1000)	-0.0930
	(0.223)
Mandatory Educ. (%)	0.0238***
	(4.90e-03)
Avg. CO Traffic (Mbps)	-30.88
	(44.43)
Constant	9.310***
	(0.197)
Observations	538
R-squared	0.098

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

that distance from the CO is a viable instrument for our analysis. More details on the appropriateness of distance as an instrument are provided in the Appendix.

More importantly, notice that since we use school fixed effects, we need distance to be uncorrelated with Δu_{it} in Equation (2.4) and not necessarily with u_{it}. In other words, our strategy allows us to control for various school unobserved effects increasing the robustness of our instrument. With distance as an instrument, we estimate a two stage least squares (2SLS) specification as follows:

$$\begin{aligned} \Delta P_{it} &= \phi + \omega \Delta I_{it} + \mathbf{x}_i \beta + \Delta \mathbf{w}_{it} \theta + \Delta u_{it} \quad (2.5) \\ \Delta I_{it} &= \varrho + \eta Distance_i + \mathbf{x}_i \varphi + \Delta \mathbf{w}_{it} \vartheta + \epsilon_{it} \end{aligned}$$

2.7 Results

2.7.1 Estimates without the instrument

We first estimate Equation (2.4) without accounting for endogeneity concerns. However, notice that we still control for school unobservable effects via first differences. The results are presented in Table 2.4. Estimates with and without covariates are very similar. Broadband use is measured as average use per student in chunks of 100 MB. Results show a very small and statistically insignificant relationship between change in exam scores and change in broadband use. Not only the standard errors are high, the estimates are economically insignificant. Control variables are also statistically and economically insignificant, which is reasonable given that we are using school fixed effects. In short, the OLS produces insignificant coefficients.

2.7.2 Correcting for Endogeneity

We estimate our Instrumental Variable (IV) specification as given by equation (2.5). The results are presented in Table 2.5. Columns (1) and (2) present results without covariates while columns (3) and (4) present results with all covariates.

The first stage of the IV specification is presented in columns (1) and (3). The estimate on distance is highly significant and negative in all specifications. This suggests that our instrument works as expected. A one kilometer increase in the distance between a school and the CO leads to about 9.1 MB (11.3 MB with covariates) decrease in total usage per student. We provide additional details on the effectiveness of our instrument

Table 2.4: Changes in 9th grade performance as a function of broadband use (OLS).

VARIABLES	(1) OLS	(2) OLS
INet Usage / Student (100 MB)	0.0110	0.0207
	(0.0329)	(0.0369)
Students (x 1000)		0.104
		(0.122)
Pop. Density (x 1000)		-0.0149
		(0.0137)
Earnings (x 1000)		-0.0142
		(0.182)
Mandatory Educ. (%)		1.83e-03
		(2.33e-03)
Δ Avg. CO Traffic (Mbps)		0.751
		(8.319)
Constant	-0.446***	-0.540***
	(0.0340)	(0.150)
Observations	2,535	2,111
R-squared	0.503	0.507
Year Dummies	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

in the Appendix. Other estimates are sensible as well. Number of students, earnings, and educational level all affect Internet usage negatively. However, the estimates are quite small. Recall that most of the control variables are pegged at 2005 levels.

Our key focus is on the results of the second stage which are presented in columns (2) and (4) of Table 2.5. The key estimate of interest is how the growth in usage of broadband per student affected grades. The estimates are negative, large and significant (at the 5% level with and without covariates). The sign on the estimate is now unequivocally negative, pointing clearly to the adverse effect of broadband on performance. Moreover, this effect seems to be reasonably large. The estimate with covariates (-.868) suggests that a unit (100 MB) increase in broadband use at the student level leads to a decrease of about 0.87 standard deviations in exam scores. The average broadband use per student in 2009 was

Table 2.5: Aggregate change in 9^{th} grades as a function of broadband use (IV).

VARIABLES	(1) 1st Stg	(2) 2nd Stg	(3) 1st Stg	(4) 2nd Stg
INet Usage / Student (100 MB)		-0.926**		-0.868**
		(0.439)		(0.373)
Students (x 1000)			-1.172***	-0.967**
			(0.141)	(0.429)
Pop. Density (x 1000)			-2.64e-04	-0.0152
			(6.38e-03)	(0.0162)
Earnings (x 1000)			-0.490***	-0.491
			(0.125)	(0.300)
Mandatory Educ. (%)			-6.16e-03***	-2.40e-03
			(2.23e-03)	(3.32e-03)
Δ Avg. CO Traffic (Mbps)			4.490	1.816
			(5.368)	(10.44)
Distance (Km)	-0.0914***		-0.113***	
	(0.0180)		(0.0283)	
Constant	0.428***	-0.137	1.776***	0.933
	(0.0262)	(0.145)	(0.182)	(0.603)
Observations	2,533	2,533	2,111	2,111
R-squared	0.150	0.178	0.325	0.264
Year Dummies	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

about 111 MB and the average grade in 2005 was about 10 standard deviations. Therefore, broadband growth between 2005 and 2009 resulted in an average decrease of 0.97 standard deviations or 9.7% in the average exam score.

We also test whether this effect changes over time by interacting the school Internet use with year dummies. This is equivalent to running separate difference regressions for each year, except that we force the covariate coefficients to be the same for all years (see Table 2.6).Despite being significant only for 2008 and 2009, the Internet use coefficient is negative for all the years (both with and without covariates) and seems to be decreasing in magnitude since 2007. These results might suggest that the adverse effect of broadband use may wear off with time. The estimates on other control variables seem to be similar to OLS specification. Given the school fixed effects, most of the control variables are insignificant.

Table 2.6: Year-specific change in 9th grades as a function of broadband use (IV).

VARIABLES	(1) 2nd Stg	(2) 2nd Stg
INet / Student (100 MB) * 2006	-0.322	-0.128
	(0.819)	(0.837)
INet / Student (100 MB) * 2007	-1.985	-1.525
	(1.378)	(0.995)
INet / Student (100 MB) * 2008	-0.973**	-0.972**
	(0.474)	(0.457)
INet / Student (100 MB) * 2009	-0.684*	-0.703*
	(0.399)	(0.372)
Students (x 1000)		-0.957**
		(0.440)
Pop. Density (x 1000)		-0.0144
		(0.0162)
Earnings (x 1000)		-0.446
		(0.298)
Mandatory Educ. (%)		-3.18e-03
		(3.52e-03)
Δ Avg. CO Traffic (Mbps)		2.934
		(10.63)
Constant	-0.337	0.681
	(0.269)	(0.661)
Observations	2,533	2,111
R-squared	0.081	0.209
Year Dummies	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In summary, our results seem to suggest that broadband use in school is generally detrimental for students' performance, at least within a few years after its introduction into the school's environment. If one believes that distraction activities on the Internet (for example, listening to music, playing games and watching movies) are inherently bandwidth intensive, then our instrument provides a consistent reason for the observed behavior. Schools which are closer to the CO, allow higher throughput and thus make it easier for students to indulge in distractive activities, lowering their exam scores. We will explore the distraction hypothesis in more detail in section 2.8.

The deployment of broadband in schools can certainly provide significant benefits and our results do not suggest that schools should not have broadband. There are many other benefits broadband may accrue which we do not measure. However, our results seem to suggest that merely connecting schools to broadband may not be enough. Various other measures need to be implemented in parallel in order to increase the productivity of investments in school broadband. We discuss the implications of our results in detail in later sections.

2.7.3 Impact Across Gender

Our specification does not allow us to estimate α and γ in equation 2.1 separately. However, distinct groups of students might use broadband to perform different activities that affect them differently. For example, we can expect that students who tend to perform more distracting activities (lower α) become more adversely affected with increased broadband use.

According to a survey administered by the Portuguese Telecom Regulator (ANACOM)to 659 students[17] boys and girls tend to perform different sets of activities on the Internet (see Table 2.7).

For instance, a higher percentage of boys reports using MySpace, watching YouTube videos and TV, listening to online radio and music, and playing online games than girls do. Girls are also more likely to look for scientific information online. Most of these differences

[17] From these, 652 students (332 girls and 320 boys) answered the question regarding activities performed on the Internet.

Table 2.7: Internet activities by gender (%).

Activity	Male	Female	Diff.
Search for Scientific Info	67.5	74.1	-6.6**
Chat	89.4	88.2	1.1
General Information	59.4	57.8	1.5
Email	93.1	89.5	3.7**
VOIP	14.1	9.3	4.7**
Radio	48.4	42.5	6.0*
TV	27.8	13.9	14.0***
MySpace & YouTube	75.9	61.7	14.2***
News & Magazines	43.8	23.8	20.0***
Music	75.6	52.7	22.9***
Games	71.9	34.9	36.9***

*** p<0.01, ** p<0.05, * p<0.1 (t-tests eq. var.)

are considerable and statistically significant. Thus, according to our framework, if we characterize many of these activities as distracting (YouTube, chat, games), we should expect a stronger adverse effect of broadband use on boys' performance. We test this hypothesis by calculating separate average scores for boys and girls and by running separate regressions of performance on broadband use for each of them.

Table 2.8 shows the results from separate IV regressions for 9$^{\text{th}}$ graders. For brevity we do not report the first stage of IV regression. Both boys and girls seem to be affected by broadband Internet use (columns (1) and (2)), but boys seem to be slightly more affected both in terms of magnitude[18] and statistical significance. Although not conclusive, these estimates are in line with our hypothesis that boys should be more affected than girls given that they perform more distracting activities on the Internet (lower α).

[18] Although not statistically different, the effect is 9% larger for boys than for girls.

Table 2.8: Change in 9th grade performance as a function of broadband use by gender and by course.

VARIABLES	(1) Male	(2) Female	(3) Port.	(4) Math
INet Usage / Student (100 MB)	-0.950**	-0.870*	-0.769**	-0.995**
	(0.461)	(0.473)	(0.381)	(0.488)
Students (x 1000)	-0.728	-1.294**	-0.865*	-1.092**
	(0.526)	(0.535)	(0.462)	(0.552)
Pop. Density (x 1000)	-0.0264	-4.72e-03	-0.0135	-0.0178
	(0.0220)	(0.0158)	(0.0160)	(0.0214)
Earnings (x 1000)	-0.353	-0.744**	-0.478*	-0.531
	(0.445)	(0.308)	(0.273)	(0.432)
Mandatory Educ. (%)	-3.13e-03	-1.36e-03	-3.50e-03	-1.47e-03
	(4.13e-03)	(4.31e-03)	(3.86e-03)	(4.47e-03)
Δ Avg. CO Traffic (Mbps)	5.076	-0.461	11.47	-7.134
	(12.01)	(12.61)	(9.557)	(14.83)
Constant	0.812	1.221*	0.316	2.235***
	(0.775)	(0.734)	(0.610)	(0.795)
Observations	2,087	2,087	2,111	2,111
R-squared	0.133	0.268	0.365	0.454
Year Dummies	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.7.4 Impact on Different Courses

The 9th grade score combines scores in math and Portuguese. Since we have information on scores in each of these courses, we now split the data between math and Portuguese and examine how these scores are affected by broadband usage. Literature does not provide a clear guidance on whether computer or broadband should affect math or languages. Angrist and Lavy (2002) find a negative effect in math exam scores for 8th graders. Malamud and Pop-Eleches (2011) find that families that acquire computers had significant lower school grades in math, English and Romanian. Rouse and Krueger (2004) find that use of a specific software designed to improve language or reading skills (FastForWord) improves some aspects of students' language skills. Banerjee et al. (2007) and Carrillo and Ponce

(2011) report that the use of computer-assisted programs improve performance in math but not in language.

We estimate Equation 2.5 for math and Portuguese separately. First stages yield consistent estimates as before. The results are presented in Table 2.8.

We get large, negative and statistically significant estimates for both math and Portuguese (columns (3) and (4)), consistent with Malamud and Pop-Eleches (2011).[19]

2.7.5 Low Performance vs. High Performance Schools

We also study which schools suffer the most with the introduction of broadband. We split our sample of schools into quartiles based on their 9th grade average exam score in 2005, thus just prior to the deployment of broadband. Table 2.9 shows the descriptive statistics for schools in the 1st and 4th quartiles. Notice that the average distance to CO is very similar for schools in first and fourth quartiles, confirming the validity of our instrument. In 2005 average grade in the 4th quartile is 32% higher than in the 1st quartile. This difference reduces to 19% and 18% in 2008 and 2009 respectively.

Table 2.9: Descriptive statistics for schools in the 1st and 4th Quartiles in 2005.

Variable	1st Quart.	4th Quart.	Diff.
Avg. Grade 2005 (s.d.)	8.97	11.87	-2.90***
Avg. Grade 2008 (s.d.)	11.19	13.32	-2.12***
Avg. Grade 2009 (s.d.)	10.87	12.86	-1.99***
Students	549.6	582.9	-33.36
Pop. Density	2017.7	2221.0	-203.3
Earnings	786.43	833.14	-46.71**
Mandatory Educ. (%)	36.49	44.46	-7.97***
Distance (Km)	1.10	0.95	0.15*

*** p<0.01, ** p<0.05, * p<0.1 (t-tests eq. var.)

We interact broadband use and distance with each of the quartile dummies in our IV

[19]Although not statistically different, the adverse effect is 29% larger for math than for Portuguese.

setting. Table 2.10 shows the results obtained. None of the quartile interaction variables displays a statistically significant coefficient. Moreover, Wald tests suggest that there is no difference across these coefficients. If anything, we see that the coefficient of the 4th quartile is more negative, possibly indicating a slight approximation of schools in extreme quartiles. Overall, these results suggest that broadband affects exam scores across all types of schools, independently of how good they were prior to the deployment of broadband.

Table 2.10: Change in 9th grade performance as a function of broadband use per quartile (IV).

VARIABLES	(1) IV 2nd Stage
INet Usage * 1st Q.	-4.75e-03
	(4.87e-03)
INet Usage * 2nd Q.	-8.64e-03
	(0.0116)
INet Usage * 3rd Q.	-4.15e-03
	(7.02e-03)
INet Usage * 4th Q.	-0.0301
	(0.0345)
Students (x 1000)	-1.268
	(1.107)
Pop. Density (x 1000)	-0.0434
	(0.0271)
Earnings (x 1000)	-0.874
	(0.843)
Mandatory Educ. (%)	7.03e-03
	(5.08e-03)
Δ Avg. CO Traffic (Mbps)	-7.625
	(11.77)
Constant	1.259
	(1.024)
Observations	2,111
Quartile Dummies	Yes
Year Dummies	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.8 Distraction Hypothesis: Additional evidence

To better understand how distraction and learning with the Internet at school affect students' grades we need to acknowledge that different schools put in place different strategies to benefit from the Internet that ultimately result in different usage patterns and learning experiences. In particular, some schools restrict access to distracting websites and applications such as Facebook and YouTube (i.e., schools with higher α), while other schools allow full access to the Internet. If the impact of broadband on school performance is negative due to these distractive activities, we should see the effect of such policies on school performance and broadband use.

We designed and deployed a survey to middle schools in Portugal in order to better understand current Internet access policies and practices. The survey consisted of 27 questions and was administered over the phone to school ICT managers between December 10^{th} 2010 and January 17^{th} 2011.[20] A total of 344 answers were obtained (response rate of 55%). Schools who completed the survey are similar to schools that did not in terms of grades, size and distance to the CO, but are different[21] in terms of Internet usage, population density, income levels and basic education levels (see Table 2.11[22]).

Among other questions, the survey asked whether the school blocks access to specific websites or applications. Respondents indicated a subset of the following options as sites or applications blocked in the school: YouTube, Facebook, Hi5, MySpace, Chat Applications,

[20]The role of ICT manager is well defined in each school, and corresponds to the person that is responsible for the maintenance of the school's computers and network. This role is usually attributed to one of the ICT teachers in the school.

[21]We use 95% confidence interval t-tests to test whether the two groups have the same mean.

[22]The '*' symbols correspond to significance levels in equal variance t-tests for the difference between surveyed and non-surveyed schools. See table footer.

Table 2.11: Summary statistics. Surveyed vs. non-surveyed schools.

VARIABLES	(1) No Survey	(2) Survey
Avg. Grade 2005 (s.d.)**	10.28	10.13
	(1.026)	(0.990)
INet Usage 2009 / Stu. (MB)***	100.2	120.2
	(81.41)	(104.7)
INet Usage 2008 / Stu. (MB)**	77.30	94.46
	(65.81)	(116.8)
Students	589.0	571.2
	(231.9)	(245.1)
Pop. Density***	2,142	1,553
	(3,211)	(2,523)
Earnings 2005**	802.8	773.9
	(183.0)	(189.2)
Mandatory Educ. (%)***	40.84	37.74
	(13.74)	(13.59)
Distance (Km)*	1.025	1.111
	(0.716)	(0.815)
Observations	284	344

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ (t-tests eq. var..)

Online Games, Other Video Sites, File Sharing Applications, Other Sites. This question seems to be the one that best proxies distraction activities with the Internet at school. The other questions in the survey covered mostly IT resources and skills.

We examine if these policies have an impact on school performance. Such policies possibly proxy for the school attitude towards technology use. By explicitly capturing these in our analysis, we are controlling for these (hence) unobserved differences across schools. Our focus is to extend our earlier model by examining how the marginal effect of broadband is conditioned by school policies.

Schools seem to be quite heterogeneous in terms of what content and activities they allow. We will focus on two measures. First, we examine if the schools that block all activities perform differently. Second, we examine the role of YouTube. We focus on

YouTube in particular because not only it may be a distracting activity, it is also bandwidth intensive. Thus the marginal effect of Internet use in schools that allow YouTube should capture the effect of distraction. We first present some summary statistics in Table 2.12.

Table 2.12: Summary statistics by blocking policy: *No Blocks* and *Allow YouTube*.

VARIABLES	(1) No Block	(2) Block	(3) Allow YouTube	(4) Block YouTube
Avg. Grade 2005 (s.d.)	10.36	10.09	10.12	10.24
	(0.847)	(1.005)	(0.977)	(1.132)
INet Usage 2009 / Stu. (MB)	120.4	120.2	121.6	103.1
	(104.0)	(104.8)	(106.8)	(68.87)
INet Usage 2008 / Stu. (MB)	83.41	96.05	95.65	77.54
	(65.83)	(122.7)	(120.3)	(48.49)
Students	607.4	565.2	563.6	662.7
	(255.4)	(242.6)	(246.0)	(207.5)
Pop. Density	1,766	1,526	1,553	1,636
	(2,028)	(2,595)	(2,561)	(2,031)
Earnings 2005	766.9	775.8	769.8	833.0
	(168.1)	(192.7)	(182.7)	(254.5)
Mandatory Educ. (%)	39.54	37.46	37.38	42.33
	(12.51)	(13.71)	(13.47)	(13.90)
Distance (Km)	1.164	1.102	1.103	1.209
	(0.883)	(0.802)	(0.807)	(0.893)
Observations	48	298	320	26

The differences in school policies seem to independent of school characteristics. The average 2005 grades across schools are quite similar. Still, schools in slightly higher income and more educated regions are more likely to block YouTube. The Internet use in schools that allow YouTube is substantially higher as expected.

Our hypothesis is that Internet use is more harmful in schools that do not restrict access to distracting websites or applications. To test this hypothesis we add the indicator *No Blocks* to our IV setup along with the interaction between Internet use and the *No Blocks* indicator.[23] We also assume that the Internet usage policies in these schools have

[23]We use the interaction between the predicted Internet use and the *No Blocks* indicator as a second instrument.

not changed over time.[24] Since we are assuming *No Blocks* to be a time-fixed school characteristic, it would be differenced out along with the other time-fixed covariates. By including it in the differences equation we are allowing it to drive the *change* in school performance, along with all other time-constant covariates (see Section 2.6.1). Table 2.13 show the results obtained.[25]

The *No Blocks* indicator provides preliminary evidence that schools that do not block any type of content perform worse (column (1)). However, we do not find evidence that this is related to Internet use given that the estimate on the interaction term is statistically insignificant (column (2)). One of the reasons for this result might be that not all websites are bandwidth-intensive and thus it is hard to establish a relationship between the time students spend in distraction activities and the amount of bytes consumed. From all the web sites and applications considered in our survey, YouTube seems to be the one for which a linear relationship between Internet use and distraction time is more likely to hold. Social network sites, chat applications and online games are relatively low-bandwidth intensive, so students can spend a lot of time with them without consuming many bytes. File-sharing applications might also be bandwidth intensive, but students can share files in the background as they perform other activities. Hence, we build an indicator called *Allow YouTube* to identify laxer schools in terms of Internet access policies.

As before, we use our IV setup to regress change in average grade on Internet use, our

[24]Several schools reported that they have been blocking more sites over time, taking advantage of a filtering service provided by the ISP for this purpose. Thus, our estimates are more conservative and should be interpreted as a lower-bound.

[25]Some covariates are missing for some of the middle schools that were surveyed and therefore the number of observations in these regressions falls short of 344. First stage estimates will be provided upon request.

Table 2.13: Change in 9th grade performance as a function of broadband use and site blocking policy (IV).

VARIABLES	(1) No Block	(2) No Block	(3) Allow Youtube	(4) Allow Youtube
INet Usage / Student (100 MB)	-0.446*	-0.444*	-0.388	0.187
	(0.245)	(0.245)	(0.236)	(0.341)
INet * No Blocks		-0.0705		
		(0.185)		
No Blocks	-0.209*	-0.159		
	(0.119)	(0.129)		
INet * Allow YouTube				-0.562**
				(0.242)
Allow YouTube			-0.283**	0.0442
			(0.117)	(0.178)
Students (x 1000)	-0.681*	-0.693*	-0.647*	-0.617*
	(0.373)	(0.384)	(0.365)	(0.361)
Pop. Density (x 1000)	0.0137	0.0140	0.0155	0.0145
	(0.0148)	(0.0151)	(0.0152)	(0.0150)
Earnings (x 1000)	-0.372	-0.382	-0.356	-0.342
	(0.312)	(0.318)	(0.323)	(0.315)
Mandatory Educ. (%)	-1.67e-03	-1.70e-03	-1.91e-03	-1.25e-03
	(4.06e-03)	(4.08e-03)	(3.95e-03)	(3.88e-03)
Δ Avg. CO Traffic (Mbps)	-9.123	-8.997	-11.47	-11.87
	(12.79)	(12.91)	(12.01)	(11.84)
Constant	0.522	0.533	0.707	0.338
	(0.496)	(0.506)	(0.531)	(0.545)
Observations	1,148	1,148	1,148	1,148
R-squared	0.425	0.425	0.447	0.456
Year Dummies	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

regional co-variates, the *Allow YouTube* indicator, and the interaction between Internet use and this indicator. Again, we use the interaction between the predicted Internet use and the *Allow YouTube* indicator as a second instrument. Columns (3) and (4) in Table 2.13 show the results obtained.[26] Schools that allow YouTube perform worse than the average: the magnitude of the *Allow YouTube* coefficient in column (3) corresponds to a decrease in grades of about 0.28 of a standard deviation. Most importantly, including the interaction

[26]First stage estimates will be provided upon request.

effect now shows that the Internet use in schools that allow YouTube leads to a large adverse effect on grades (column (4)): a decrease of 0.55 standard deviations per 100 MB. Put another way, the effect of Internet is significantly worse when schools allow *YouTube*. This provides support to our argument that when Internet is being used for bandwidth intensive distracting activities, it leads to adverse effect on student performance.

In sum, we find evidence that the way schools allow students to use Internet connectivity affects students' performance and students do relatively worse in schools that enact laxer access policies that do not control the opportunities for exaggerated distraction.

2.9 Conclusion and Discussion

Using a comprehensive dataset on broadband use in every middle school in Portugal, we find evidence that broadband hurts student performance. Our analysis shows that on average broadband is responsible for a decline of 0.97 of a standard deviation in grades in the 2005-2009 window. We also use a unique and interesting instrument to account for endogeneity. All tests on the robustness of our the instrument suggest that it is a credible instrument. In this regard, our paper makes a significant methodological contribution. We also find all types of schools (low vs high performing) are equally affected by broadband regardless of their performance in 2005. A technology like broadband may not always be used productively, hence its availability in poor performing schools might not necessarily translate into better grades.

We conduct a survey to explore the distraction effect of Internet in more detail. Some schools block many applications and services which can be characterized as distractive (such as music, movies, chat, online gaming). Using these additional data, we find that schools that block access to all such activities perform better that the schools which do not. More interestingly, we focus on YouTube access, which is a bandwidth intensive application. In fact, schools which allow YouTube, typically also consume more bandwidth. We find strong evidence that indeed Internet use is significantly more adverse in schools that allow access to YouTube.

Our study, applied to the case of Portugal, shows that the introduction of broadband in schools does not necessarily contribute to increase students' performance, at least in the few years after its deployment. While we do not have direct measurement, our results suggest that the introduction of broadband in the school environment must be complemented with policies aimed at effectively embedding the Internet in the education system that promote productive use of the Internet in ways that complement traditional study. This may be particularly true for students in early high school who, without proper monitoring, may be more likely to engage in distracting activities. Recall that broadband was provided to all schools as a central policy decision possibly without giving enough time to schools to think and plan ahead. Benefiting from the Internet requires active engagement from schools, teachers and students to bring everyone on board to correctly exploit the opportunities offered by the technology.

While we use a very detailed dataset, our study is not without limitations. We do not know precisely the kind of activities students engage in with the Internet. We have also looked at the effects of blocking access to certain applications but have not carefully analyzed the full spectrum of policies that schools can use to benefit from Internet. Similarly, broadband may still be beneficial for students in ways that test scores do not capture, whose effects our study cannot appreciate. For example, broadband deployment in schools allows students to be exposed to new sets of technologies that they will most likely use later both in their professional careers to increase their productivity and in their personal lives to facilitate, for example, communication with friends and family. However, these kinds of benefits are extremely difficult to measure and our study fails to take them into account. Nevertheless, we emphasize that in any country education policy today is largely shaped by schools' performance, and everyone in the educational system, from students, to teachers and schools, parents and educators, have clear incentives to improve students' performance. In this regard our paper is the first of its kind to provide concrete evidence of how the introduction of broadband in schools affects student performance.

Appendices

2.A Robustness tests for Distance as an Instrument

The distance between a school and the CO that serves it is a good instrument because the speed of the ADSL connection reduces with the length of the copper wire (see Tanenbaum, 2002, chapter 2). Our first stage regressions show that this is the case. Also, grades in 2005 seem to be unaffected by distance, after controlling for region and school-specific characteristics (see Table 2.3).

There is, however, the concern that end-users may not be able to appreciate differences in the quality of ADSL connections for short distances between schools and COs, rendering our instrument invalid for schools that are very close to the CO. Also, ADSL speeds may have been capped by the provider, which would render the quality of ADSL connections similar for all schools close to the CO. We test these hypotheses by introducing distance threshold dummies in the first-stage regression.

Table 2.14 shows that none of the distance thresholds is significant in 2009.[27] This shows that usage reduces with distance for schools close and far away from the CO alike. This is consistent with the hypotheses that ADSL connections have not been capped, at least not at a rate that schools do use, and that users perceive differences in the quality of the ADSL connection even across schools that are close to the CO.

Table 2.14: Distance threshold regressions for schools with 9^{th} grade students.

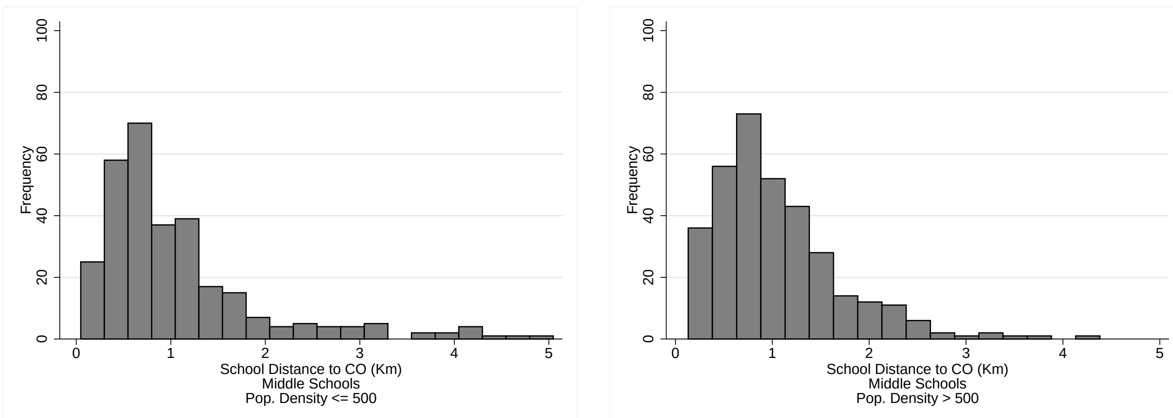
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Distance (Km)	-16.56***	-19.53***	-17.66**	-16.80***	-16.63***	-30.40*
	(4.315)	(5.775)	(7.259)	(5.467)	(4.094)	(16.27)
Students	-0.186***	-0.187***	-0.186***	-0.186***	-0.186***	-0.187***
	(0.0203)	(0.0209)	(0.0207)	(0.0206)	(0.0207)	(0.0208)
Pop. Density	1.10e-04	9.33e-05	1.17e-04	1.11e-04	1.11e-04	1.14e-04
	(9.80e-04)	(9.83e-04)	(9.81e-04)	(9.82e-04)	(9.82e-04)	(9.75e-04)
Earnings 2005	-0.0968***	-0.0972***	-0.0967***	-0.0968***	-0.0968***	-0.0979***
	(0.0217)	(0.0220)	(0.0220)	(0.0220)	(0.0220)	(0.0219)
Mandatory Educ. (%)	-0.850***	-0.858***	-0.849***	-0.850***	-0.851***	-0.844***
	(0.314)	(0.299)	(0.303)	(0.304)	(0.300)	(0.316)
Δ Avg. CO Traffic (Mbps)	257.1	286.5	254.3	254.6	257.7	250.8
	(585.4)	(585.6)	(581.3)	(587.6)	(580.9)	(587.1)
Dist. > 0.5 Km	-0.299					3.706
	(11.31)					(11.98)
Dist. > 1 Km		6.324				12.54
		(9.041)				(12.64)
Dist. > 2 Km			3.340			13.28
			(16.60)			(20.54)
Dist. > 3 Km				1.077		20.05
				(24.22)		(32.02)
Constant	345.6***	347.1***	346.1***	345.6***	345.5***	351.7***
	(27.18)	(26.58)	(26.53)	(26.37)	(26.17)	(28.62)
Observations	527	527	527	527	527	527
R-squared	0.368	0.369	0.368	0.368	0.368	0.369

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

There is also a concern that distance to CO and regional co-variates such as population

[27]Regressions for 2008 yield similar results.

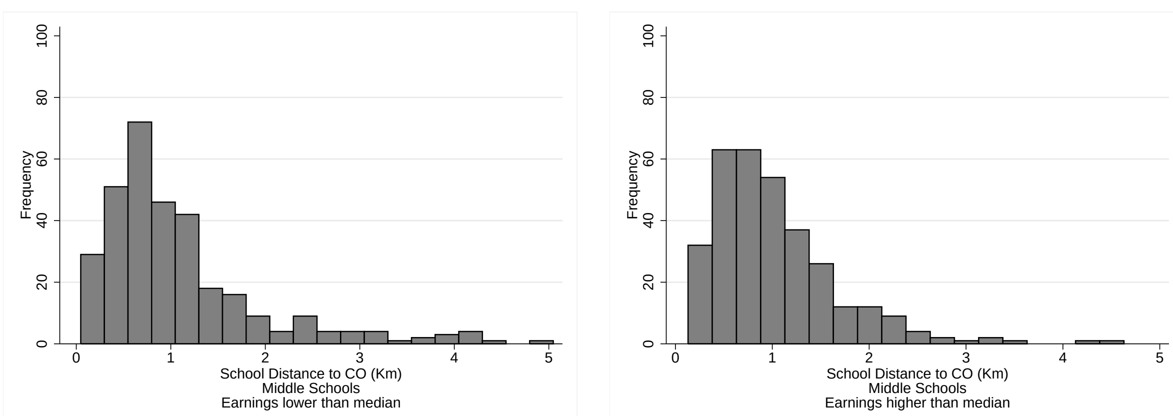
density, earnings and mandatory education are correlated. Table 2.2 shows that this is not the case. Furthermore, Figure 2.5 shows that the distance to CO for schools in both high and low density areas ranges from a few meters to as much as 5 Km. Likewise for earnings and mandatory education as Figures 2.6 and 2.7 report.



(a) Pop. Density ≤ 500.

(b) Pop. Density > 500.

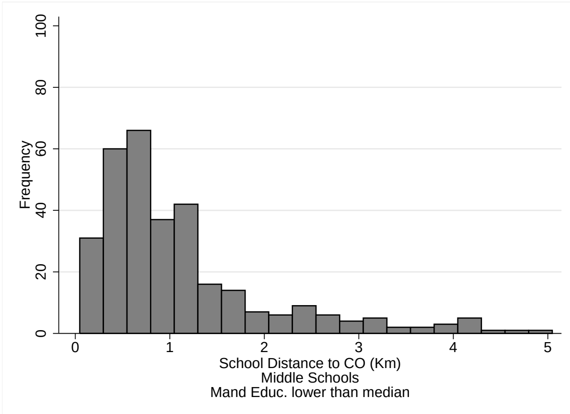
Figure 2.5: Middle School distances to the closest CO by Population Density.



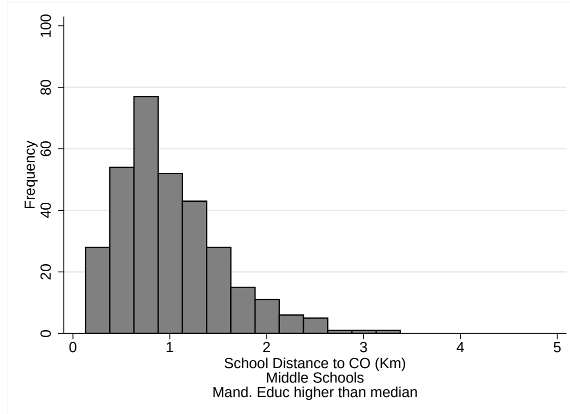
(a) Earnings lower than median

(b) Earnings higher than median

Figure 2.6: Middle School distances to the closest CO by Earnings.



(a) Mand Educ. lower than median



(b) Mand Educ. higher than median

Figure 2.7: Middle School distances to the closest CO by Educ Level.

2.B Testing for weak instrument

As mentioned by Staiger and Stock (1997), weak instruments may lead to a more severe bias than the bias introduced by OLS estimates when one of the regressors is endogenous. We follow Stock et al. (2002) to test whether distance to the ISP's CO is a weak instrument. We use the size-based definition of weak instruments to test whether the correlation between our instrument and the endogenous regressor is weak, in which case the conventional first-order asymptotics no longer hold. The F-statistic for the distance to the CO in the first-stage of our main regression is 15.89. The Stock et al. critical value for a test of size $r = 0.15$ and significance level $\alpha = 0.05$ is 8.96. Therefore, our instrument does not belong in the set of weak instruments, for this size and significance level.

Chapter 3

Spillover Effects of Broadband in Schools and the Critical Role of Children

Abstract: Providing broadband access to schools can contribute to increase household penetration because new infrastructure that is often put into place to meet schools' needs can also be used to serve households and also because students get acquainted with the technology and signal the value of Internet to other family members who can, as a consequence, also adopt. Using a household level survey data on home Internet penetration we find evidence that Internet use at school leads to higher levels of Internet penetration in the surrounding region, and that this spillover effect is mediated by children. We develop a structural model that provides insight on how Internet use at school affects home Internet penetration and how Internet penetration affects school Internet use. We assess the

magnitude of this effect using household level data on home Internet penetration and Internet traffic in all schools in Portugal. We address endogeneity issues by using two sets of instruments for schools' broadband use. Both approaches yield similar results providing us confidence in the results obtained. We find that school broadband use contributes directly to a higher adoption rate in households with children. During 2008 and 2009 school Internet use increased the probability of adopting Internet by 20% in households with children, while no statistically significant effect was found in households without children.

3.1 Introduction

The diffusion of technology has been an active area of research since the classic works of Griliches (1957), Rogers (1962) and Bass (1969). In particular, the diffusion of computers and the Internet has been widely researched (e.g., Brynjolfsson and Hitt, 2003; Chinn and Fairlie, 2006; Goolsbee, 2006; Greenstein and McDevitt, 2009; Rosston et al., 2010). More recently a number of studies have also looked at the determinants of broadband adoption, focusing both on the supply side (e.g., Glass and Stefanova, 2010) and on the demand side (e.g., Chinn and Fairlie, 2006; Whitacre, 2008; Lee and Brown, 2008). On the demand side, many different kinds of programs have been implemented to stimulate adoption of broadband Internet. These programs target different barriers to adoption, ranging from affordability to lack of perceived value (see Hauge and Prieger, 2010, for a review of demand-side programs aimed at encouraging broadband adoption). One such kind of programs subsidizes the acquisition of computers and broadband technology by schools (e.g., E-Rate[1]). One idea beyond these programs is that students get acquainted with the new technology at school and bring new knowledge home, which in turn generates adoption at the household level.

In this paper we study the effect of providing broadband Internet to schools on the penetration rate of Internet at home. The deployment of Internet in schools might have direct effects in the ICT skills of students and performance (e.g., Goolsbee and Guryan, 2006; Belo et al., 2010), but also indirect effects such as fostering the adoption of Internet at home through spillover effects (e.g., Goolsbee and Klenow, 2002).[2] On the other hand, these programs might decrease the demand for Internet at home because Internet in schools might act as a substitute for Internet at home. First, the use students make of the Internet at school might be enough. Second, adults at home might not find it beneficial to subscribe Internet service. Therefore, the correlation between Internet penetration at school and at home might go either way. Furthermore, a causal effect of Internet in school on Internet at home, while interesting to observe and to manage from a public policy perspective, might be inexistent. Even if otherwise, its magnitude might be trivial leaving policy makers with little room to claim that wiring schools accelerates Internet penetration.

We develop a structural model that captures how Internet use at school affects Internet adoption at home and how Internet penetration at home affects Internet use at school. We identify three potencial sources of spillovers from school to home. First, depending on whether Internet use at school and Internet use at home are complements or substitutes,

[1]See, for example Hudson (2004); Jayakar (2004) for more information on the E-Rate program.

[2]Depending on whether these second order effects have been accounted for or not when setting up the projects, they are called either externalities or spillovers. In this paper we treat them as spillovers. See Angelucci and Di Maro (2010) for a taxonomy of spillovers.

the need for children to have Internet at home can, respectively, increase or decrease with school Interet use. Second, under the null hypothesis of positive spillovers, the more children use Internet at school the higher the likelihood their parents will adopt Internet at home. Finally, our model allows us to test a third complementary spillover effect in which Internet use in schools might also influence Internet adoption by households without children simply because they are located in a region with more Internet penetration. These spillovers are referenced in the literature as community level spillovers (see Goolsbee and Klenow, 2002).

We use household level survey data to provide empirical evidence of spillover effects from using Internet in school. Home Internet penetration is measured by a binary variable obtained survey on the use of ICTs in Portugal. School Internet use is measured at the municipality level and corresponds to the amount of traffic exchanged between schools and the Internet in gigabytes (GB) per student. We collect household level information, including whether or not there is a child in the household, Internet use, household income, household size and locality type, for 2008 and 2009 and match them with school Internet use at the municipality level. We address endogeneity issues by instrumenting schools' broadband use with (1) municipality level variables, and with (2) the distance between the school and the closest Central Office (CO) of the ISP providing broadband to schools, which proxies the quality of the ADSL connection. Both approaches yield similar results providing us confidence in our findings.

We find evidence that children play a key role in home Internet adoption. We find that school broadband use contributes directly to a higher adoption rate in households with children. School Internet use contributed to an increase in home Internet adoption of 20% in households with children between 2008 and 2009. This corresponds to an increase of roughly 5% in the penetration of home Internet across the whole country[3]. The spillover effect is similar for wireless Internet adoption. We found no evidence of community level spillover effects. In other words, the fact that a household locates in a region with a higher penetration of home Internet does not contribute, per si, to increase the propensity to adopt.

The remainder of this paper is structured as follows. Section 3.2 provides a review of the relevant literature. Section 3.3 describes our structural model. Section 3.4 introduces the initiatives sponsored by the Portuguese Government to provide Internet to schools and describes the data we used. Section 3.5 presents the empirical strategy we followed, and Section 3.6 presents the empirical results obtained. Section 3.7 concludes pointing out limitations of our study and ideas for future research.

3.2 Related Work

The diffusion of general purpose technologies, such as computers and the Internet, has attracted significant attention from scholars and policy makers. Many authors have shown evidence of positive economic effects in computerization and broadband diffusion in the long-term[4] However the diffusion of these technologies is not uniform across the population,

[3]These statistics are obtained from our IV estimates. OLS estimates are, respectively, 11% and 3%

[4]For instance Brynjolfsson and Hitt (2003) study the impact of computerization in 587 large U.S. firms. They find that the long-term impact of computerization is 5 times greater than the corresponding short-term impact. This suggests that computerization is complemented with several other costly activities that drive down short-term estimates. Greenstein and McDevitt (2009) estimate the value created by the diffusion of broadband in the U.S. economy from 1999 to 2006. They estimate that broadband contributed

and it has been shown that spillovers play an important role in explaining the dynamics and heterogeneity in the diffusion of these technologies. Spillovers have been traditionally studied in the context of knowledge and technology transfer, both from the perspective of firm-to-firm (e.g., Jaffe et al., 1993) and university-to-firm (e.g., Monjon and Waelbroeck, 2003).

For instance, Goolsbee and Klenow (2002) find evidence of city-level spillovers in the adoption of computers. They use a survey of 11,000 households in the U.S. and find that spillovers are higher across family members and among people that use the computer more intensively. Goldfarb (2006) found that students that contacted with the Internet in University are more likely to use Internet later in their lifes and that people that contacted directly with these people are also more likely to have Internet. This is especially true for lower income households. Ward (2010) also studied Internet adoption in the U.S.. He found evidence of spillovers from three sources: local Internet usage, school funding and university exposure.

In parallel, research on the determinants of broadband adoption has shown that, in general, higher levels of income and education are associated with higher levels of Internet adoption, [5] and that availability and affordability do not seem to be enough to ensure

to an increase in revenue between $8.3 billion and $10.6 billion from Internet access between 1999 and 2006, which is associated to an increase in consumer surplus between $4.8 and $6.7 billion. Forman et al. (2012) find that investments in advancing Internet payoff only in counties that were already well-off. They aggregate firm-level data on Internet use and on the number of computers from 1995 to 2000 in the US and link them with income, population and education data at the county level.

[5]For example, Chinn and Fairlie (2006) examine a panel of 161 countries from 1999 to 2004 and find a positive correlation between technology penetration rates and income, human capital and telephone density. Prieger and Hu (2008) analyze the gap in Internet access between minority groups and white households. They find that even after controlling for income, education and other demographic indicators, the gap in DSL penetration between minority households (i.e., Hispanic and black households) and white households remains statistically significant. Whitacre (2008) analyzes the influence of household characteristics and infrastructure availability in the adoption of Internet in Oklahoma and concludes that

adoption.[6]

A wealth of demand-side programs aim at fostering broadband adoption by lowering barriers such as the lack of perceived value of Internet and the lack of digital literacy. (see Hauge and Prieger, 2010, for a review of demand-side programs aimed at encouraging broadband adoption). One such strategy is to promote contact with the Internet in public spaces such as schools and libraries, helping students and individuals to get acquainted with new technologies and the opportunities they offer. From a research perspective, however, the non-experimental nature of most of these programs hinder the identification and measurement of their effects, which can easily be confounded with unobserved factors. In most cases, researchers need to resort to sophisticated analytical methods to tease out the causal effects of these programs. Still, in many cases only limited statistical analyses can be performed because one is unable to characterize how outcomes would have been in the absence of these programs.

Despite these difficulties some studies have been able to measure program outcomes. One such example is Goolsbee and Guryan (2006) that studied the effects of subsidizing schools' Internet equipment acquisition on actual equipment aquisition. They take advantage of variance in subsidy rates awarded to schools based on poverty rates and urban/rural status, and find that schools that entitled to receive higher subsidy rates invest more in technology. Their identification assumption is that the exact subsidy rate schools receive is determined exogenously and that schools do not try to change their eligibility status.

household characteristics rather than infrastructure availability are the main drivers for Internet adoption.

[6]In fact, the top two reasons for not having Internet at home in EU households in 2010, were lack of perceived value (40%) and lack of digital literacy (32%), ahead of high access costs (23%), as registered by the Eurostat. In the US these statistics were 19%, 10%, and 24%, respectively, as reported in The National Broadband Plan by the FCC.

Our work is closest to Goolsbee and Klenow (2002) and Ward (2010). Both of these studies use region-level observables as instruments to overcome potential endogeneity problems.

3.3 Model

We provide a structural model to highlight the main aspects of household Internet adoption and its relationship to school Internet use.

Setup. Every household in a continuum of households with total mass N inlcudes two types of agents: adults, a, and children, c. Some households have no children, and the total mass of households with children is N_c. The utility of children, u_c, depends on how much Internet they use at home and at school, h_c and s_c respectively, and on how much time they spend in other leisure activities, l_c. We assume the utility function is increasing and concave in all its arguments ($\frac{\partial u_c}{\partial i} > 0; \frac{\partial^2 u_c}{\partial i^2} < 0, \forall i \in \{h_c, s_c, l_c\}$).

The *perceived* utility from using the Internet for an adult is based on her children's utility contingent on whether there is Internet at home. Let u_c^* and u_c^{**} represent the utility of the children when there is and there is not Internet at home, respectively, and let I be an indicator variable that represents the existence of Internert at home. The adult's utility depends also on her *a priori* belief about how much time she will spend on the Internet, h_a, on how much time will be spent on complementary leisure activities, l_a and on the utility derived from consuming an outside good q^O. Furthermore, the utility of the adult from using the Internet is affected by three factors: her *a priori* belief about the technology and its usefulness, θ, which is randomly drawn from a distribution common to all households, $G(\cdot)$, the knowledge transmitted by the children about how best to benefit from the Internet, which depends directly on the children's use of the Internet at school, s_c, and on a signal conveyed by the fraction of neighboring households that have already adopted the Internet, k, of the utility that one can derive from using the Internet.[7] Additionally, we assume that subscribing Internet provides some utility that does not depend on the effective usage: w. This utility is associated to the sheer fact that Internet is there and can be readily used if needed. This utility increases with θ, s_c, and k.

Adult i maximizes her utility, u_{ai}, defined as

$$
\begin{aligned}
u_{ai}(h_a, l_a, I, \theta, s_c, k) &= I \cdot u_{ci}^{*} + (1 - I) \cdot u_{ci}^{**} \\
&+ I \cdot [w + (\theta + \varphi_s s_c + \varphi_k k)(2h_a - h_a^2)] + (2l_a - l_a^2) + q^O
\end{aligned}
$$

subject to monetary and time constraints, $q^O + f \cdot I \leq y$ and $h_a + l_a \leq T_a$, where f represents a fixed fee for Internet access, y is the total monetary budget, and T_a represents the adult's time budget.

Note that θ, s_c and k are interchangeable. Therefore, we define $\gamma \equiv \theta + \varphi_s s_c + \varphi_k k$. Note also that the assessment the adult makes on the potential utility derived from using the Internet is based on the children's usage *before* Internet is adopted. Therefore, s_c corresponds to the intensity of Internet use at school *before* the household adopts Internet.

At any given point in time adults in households without Internet decide whether to

[7] In this setting, k can also represent a network-level effect: the higher the mass of neighboring households that have adopted the Internet, the higher the individual utility drawn from adopting it.

adopt Internet. This decision is based on the expected utility of the entire household, which is given by the sum of the adults' and children's utilities. For simplicity, we assume that households have only one adult and at most one child and that their utilities contribute equally to the total utility of the household. Thus, contingent on whether the household has adopted Internet, the household's utility will be:

$$u_{ai}(h_a, l_a, I, \gamma) = \begin{cases} u_{ci}^{*} + w + \gamma(2h_a - h_a^2) + (2l_a - l_a^2) + q^O & \text{if } I = 1 \\ u_{ci}^{**} + (2l_a - l_a^2) + q^O & \text{otherwise} \end{cases}$$

Utility maximization and adoption. A child's utility is maximized by allocating time to activities according to their relative contribution. Let h_c^*, s_c^* and l_c^* represent the optimal amounts of time spent on each activity if the child has Internet at home. Let h_c^{**} $(= 0)$, s_c^{**} and l_c^{**} represent the optimal amounts of time spent on each activity if the child has no Internet at home[8].

The adult also maximizes her utility, which incorporates the child's utility, yielding the optimal amounts of Internet use, h_a^*, leisure time, l_a^*, and of the outside good, q^{O*}:

$$h_a^* = \begin{cases} \frac{\gamma + T_a - 1}{\gamma + 1} & \text{if } I = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$l_a^* = \begin{cases} T_a - \frac{\gamma + T_a - 1}{\gamma + 1} & \text{if } I = 1 \\ T_a & \text{otherwise} \end{cases}$$

[8]For example, when all activities contribute equally, a child's utility is maximized by allocating similar amounts of time to each activity. Thus, for children with Internet at home, each activity is allocated one third of the time available. For children without Internet at home, $h_c^{**} = 0$ and s_c^{**} and l_c^{**} are allocated half of the time available each.

$$q^{O*} = \begin{cases} y - f & \text{if } I = 1 \\ y & \text{otherwise} \end{cases}$$

Normalizing T_a to 1 yields the following indirect utility function:

$$u_{ai}^*(y, f, T_a, I, \gamma) = \begin{cases} u_{ci}^* + w + \gamma(\frac{2\gamma}{\gamma+1} - (\frac{\gamma}{\gamma+1})^2) + (\frac{2}{\gamma+1} - (\frac{1}{\gamma+1})^2) + y - f & \text{if } I = 1 \\ u_{ci}^{**} + 1 + y & \text{otherwise} \end{cases}$$

At any given point in time, a household adopts Internet iff $u_a^*|_{I=1} - u_a^*|_{I=0} \geq 0$:

$$\begin{aligned} u_a^*|_{I=1} - u_a^*|_{I=0} &= \delta_c + w + \gamma + \frac{1}{\gamma+1} - 1 - f \\ &= \delta_c + \delta_a - f \end{aligned}$$

where $\delta_c \equiv u_c^* - u_c^{**} \geq 0$ and $\delta_a \equiv w + \gamma + \frac{1}{\gamma+1} - 1 \geq 0$.

Thus, a household adopts Internet if the sum of the differences in utility is higher than the Internet subscription fee. To avoid clutering the notation and to simplify empirical estimation later on, we assume $w \equiv \frac{\gamma}{\gamma+1}$. This assumption simplifies estimation because neglects the existence of second-order interaction effects among the terms of γ. Such effects would attenuate the marginal effect of s_c on the adult's utility. Thus, the components of γ we estimate in this paper will be lower bounds for the actual effects.

The minimum level of *a priori* utility beyond which a household decides to adopt

Internet is, represented by $\hat{\theta}$:

$$\begin{aligned} u_a^*|_{I=1} - u_a^*|_{I=0} &\geq 0 \Leftrightarrow \\ \hat{\theta} &\geq f - \delta_c - \varphi_s s_c - \varphi_k k \end{aligned}$$

In particular, for a household without children:

$$\hat{\theta} \geq f - \varphi_k k$$

Thus, the minimum level $\hat{\theta}$ beyond which a household decides to adopt Internet is lower for households with a child and is lower the greater the amount of Internet used at school. This leads us to state our first research hypothesis:

H_1: Households with children are more likely to adopt Internet.

School Internet use. Given the substitution effect between Internet use at school and Internet use at home, children without Internet at home are likely to use more Internet at school than children with Internet at home, $s_c^{**} > s_c^*$. Thus, schools located in areas with lower home Internet penetration are likely to register more traffic, which leads us to state our second research hypothesis:

H_2: Schools located in neighborhoods with lower levels of home Internet penetration are likely to register more Internet traffic.

Recall that s_c^* and s_c^{**} correspond to the use of Internet in schools that maximize the utility of children assuming no restrictions in Internet use in schools. However, in practice, several restrictions in schools are likely to affect the students' use of Internet downwards (e.g., Belo et al., 2010). Hence, we assume that Internet use in school, s_c, is capped at the school level to all students. This assumption does not mean that schools in areas with lower home Internet penetration cannot enact policies and practices that favor a more intense use Internet (H_2). It only captures the fact that Internet traffic in schools is not completely determined by the students behavior and that additional factors also play a role in determining how much students can use . This assumption will be revisited later in this paper because it will be useful for our identification strategy.

Internet penetration evolution. At each moment in time, the number of new adopters is determined by the distribution density function of θ evaluated at the minimum *a priori* utility level, $g(\hat{\theta})$.

$$\dot{k} = \frac{N - N_c}{N} g(\hat{\theta}_0) + \frac{N_c}{N} g(\hat{\theta}_1) \tag{3.1}$$

Assuming that θ is uniformly distributed over its support, $[\underline{\theta}, \bar{\theta}]$, $\dot{k} = 1/(\bar{\theta} - \underline{\theta}) \equiv \alpha$ is constant and thus $k(t) = \alpha t + C$. Plugging in the values for $t = 0$, before any adoption took place ($k = 0$),

$$\begin{aligned} k(0) &= \frac{N - N_c}{N} \int_{\hat{\theta}_0}^{\bar{\theta}} g(x)dx + \frac{N_c}{N} \int_{\hat{\theta}_1}^{\bar{\theta}} g(x)dx \\ &= \alpha \frac{N_c}{N} (\varphi_s s_c + \delta_c) - \alpha f + 1 \end{aligned} \tag{3.2}$$

Therefore,

$$k(t) = \alpha t + \alpha \frac{N_c}{N}(\varphi_s s_c + \delta_c) - \alpha f + 1 \tag{3.3}$$

Home Internet penetration increases linearly with time and with school Internet use if $\varphi_s > 0$.[9]

Putting it all together. Plugging in the household Internet penetration evolution (3.3) into the adoption decision of a household with children (3.1), we obtain:

$$\begin{aligned} u_a^*|_{I=1} - u_a^*|_{I=0} &= \delta_c + \delta_a - f \\ &= \delta_c + \theta + \varphi_s s_c + \varphi_k k - f \\ &= \delta_c + \phi_1 s_c + \phi_2 \frac{N_c}{N} s_c + \phi_3 \frac{N_c}{N} + \phi_4 t + u \end{aligned} \tag{3.4}$$

where $\phi_1 \equiv \varphi_s$, $\phi_2 \equiv \alpha\varphi_k\varphi_s$, $\phi_3 \equiv \alpha\varphi_k\delta_c$, $\phi_4 \equiv \alpha\varphi_k$, and $u \equiv \theta - (\alpha\varphi_k + 1)f + \varphi_k$.

For a household without children the difference in utilities with and without households is given by:

[9] In fact, given that the marginal utility of school Internet is higher the lower the Internet use at home ($\frac{\partial^2 u_c}{\partial h_c \partial s_c} < 0$), if s_c increases, δ_c decreases, meaning that the extra utility that children get from using Internet at home is not as big if Internet use in school is high. Thus, an increase in s_c is attenuated by the decrease in δ_c, which means that the direct effect of s_c is not identifiable. However, this attenuation effect tends to disappear as s_c increases: $lim_{s_c \to \infty} \frac{\partial \delta_c}{\partial s_c} = 0$. Moreover, this attenuation would bias the coefficient of s_c downwards and therefore any positive coefficient obtained using this model for s_c must be interpreted as a lower bound.

$$u_a^*|_{I=1} - u_a^*|_{I=0} \quad = \quad \phi_2 \frac{N_c}{N} s_c + \phi_3 \frac{N_c}{N} + \phi_4 t + u$$

Note that for households with children s_c shows up twice in the decision to adopt. The first appearance corresponds to the direct effect that the child exerts over the adult. The second, that shows up also in households without children, corresponds to the community level effect, and is proportional to the percentage of households with children. This leads us to state our third and fourth research hypotheses:

H_3 : Households with children located close to schools with more Internet use are more likely to adopt Internet.

H_4 : Households located in areas with more children are more likely to adopt Internet.

3.4 Context and Data

In Portugal most elementary and secondary schools are public schools, funded either by the Central Government or the Local Government, with limited autonomy to manage their resources. The provisioning of Internet to schools has been managed by FCCN - the Portuguese National Foundation for Scientific Computation. FCCN is a private foundation, under the tutelage of the Ministry of Science, Technology and Higher Education, that runs the National Research and Education Network (NREN). The NREN connects all schools, institutions of higher education and research labs in the country. The same institutional model is followed by a number of other European countries, each having its own NREN.

NRENs interconnect forming a trans-European NREN, called the GEANT network.

In 2004, this Ministry launched a major initiative, aimed at replacing all the existing ISDN connections in schools by broadband ADSL.[10]. This project was completed by January 2006, despite the fact that only less than 15% of the schools had migrated to ADSL before July 2005 (UMIC, 2007). Most schools (>95%) received a DSL modem from FCCN and an ADSL connection of at least 1 Mbps over the copper line that connects them to the ISP's Central Office (COs) from which FCCN buys connectivity to the Internet backbone (Figure 3.1).[11] The Ministry covered all up-front capital costs to deploy broadband to schools. City Halls foot the broadband monthly bill for elementary schools and the Ministry covers these costs for the reminder of the schools.

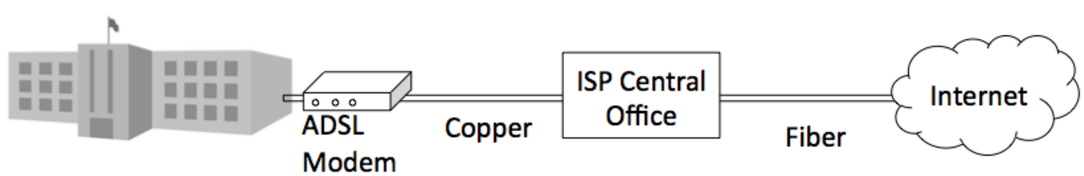


Figure 3.1: Broadband schools' connection to the Internet. Schools connect through a copper line to the ISP's central office. From there, the ISP ensures connectivity to the Internet backbone through fiber.

[10]Migration to ADSL was complemented with several other initiatives. One such initiative was ICTs training for teachers. Another initiative was the subsidization of 150-Euro laptops to students. This initiative, called "e-schools", might have boosted Internet use in many schools. A third initiative was to award up to 24 laptops to each and every school. Most schools use these laptops to bring Internet to the classroom. Some schools have a dedicated room in which these laptops remain and can be used as desktops.

[11]The remainder of the schools, where this speed could not be offered over copper, got a symmetric 256 Kbps ISDN connection to the Internet.

Household level data

Household data were obtained from a yearly survey administered to households in Portugal by the Portuguese National Statistics Institute (PNSI). PNSI administers this survey to track the use of Information and Communication Technologies (ICT) in the country. This survey is administered since 2003, and from year to year 3/4 of the sample are kept from the previous year while 1/4 are new households. Household identifiers are not available and thus we are not able to construct true a panel, but we know the municipality in which each household is located. Also, some of the survey questions change from one year to the following year, which prevents us from using some of the data available. For example, information about the household composition and income are available only since 2008. Roughly 4,000 households are surveyed every year. There are about 3.5 million households in Portugal.

Internet penetration in Portugal increased from just over 20% in 2003 to almost 50% in 2009. Broadband Internet grew as well: in 2003 it represented half of all Internet penetration, while in 2009 broadband was in virtually all households with Internet (see Figure 3.2).

Figure 3.3 depicts the evolution of broadband technologies over the period of analysis. Cable and DSL have been the dominant broadband technologies for a long time but starting in 2006 wireless Internet has grown considerably surpassing wired broadband in 2009. [12] In this paper, 3e are interested in learning whether the increase in broadband household

[12] Note that the sum of ADSL, cable and wireless is higher than household broadband penetration because some households have both wired and wireless Internet. While some people use wireless Internet as their primary way to access the Internet, other people use wireless only as a secondary way to connect.

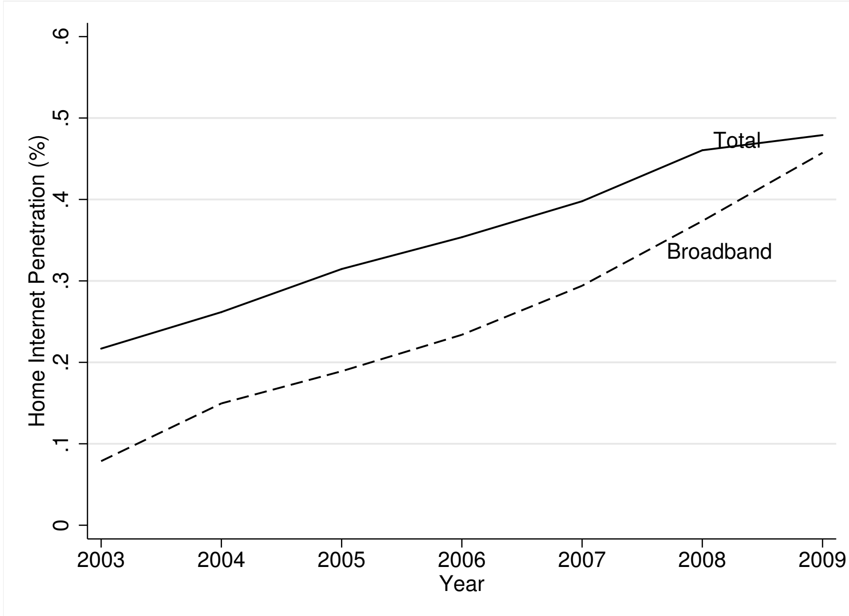


Figure 3.2: Home Internet penetration.

penetration has been influenced by the existence of Internet in nearby schools and, if so, how big is this effect.

As expected, computer and home Internet penetration rates are higher in more densely populated areas (see Table 3.1). Intermediate density areas are the ones with higher percentage of households with children. Also, consistent with our model and with H_1, households with children exhibit higher levels of computer and Internet adoption, and have higher income (see Table 3.2).[13]

Internet Use at Schools

School traffic data were obtained from monitoring tools set up by FCCN. We obtained monthly reports that include download and upload traffic per school between November

[13]This is stated assuming that household composition is exogenous. However, we are aware that selection might have played a role here. For instance, households with better financial conditions might have chosen to have children, and thus are also more likely to be able to afford computers and Internet.

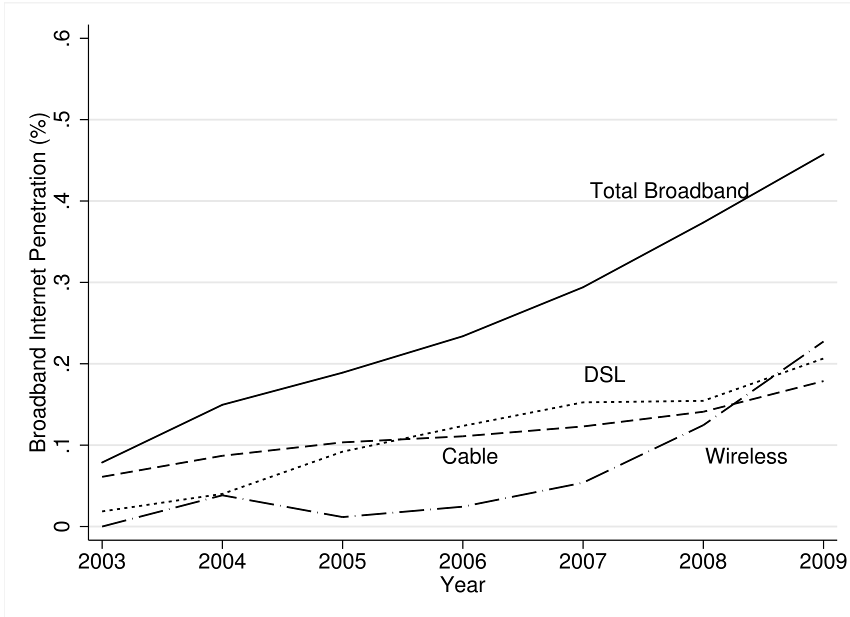


Figure 3.3: Household broadband Internet penetration by technology

Table 3.1: Summary statistics by population density.

	(1)	(2)	(3)
VARIABLES	High Dens.	Interm. Dens.	Low Dens.
Home INet 2009	0.561	0.474	0.368
	(0.496)	(0.499)	(0.482)
Broadband INet	0.538	0.452	0.354
	(0.499)	(0.498)	(0.478)
Wireless INet	0.235	0.197	0.174
	(0.424)	(0.398)	(0.380)
Computer (%)	0.635	0.565	0.447
	(0.482)	(0.496)	(0.497)
Children (%)	0.302	0.322	0.258
	(0.459)	(0.467)	(0.438)
Household Income	4.834	4.290	3.777
	(2.177)	(1.983)	(1.799)
Household Size	2.789	2.892	2.713
	(1.322)	(1.329)	(1.250)
Observations	2,603	2,672	2,407

Table 3.2: Summary statistics by household type.

VARIABLES	(1) No Children	(2) Children
Home INet 2009***	0.377	0.693
	(0.485)	(0.461)
Broadband INet***	0.358	0.670
	(0.480)	(0.470)
Wireless INet***	0.149	0.331
	(0.356)	(0.471)
Computer (%)***	0.437	0.825
	(0.496)	(0.380)
Household Income***	4.091	4.844
	(2.038)	(1.953)
Household Size***	2.293	4.013
	(0.968)	(1.200)
Observations	5,413	2,269

Standard deviations in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2005 and June 2009. School traffic is measured at the school's edge router and consists of all traffic exchanged between the school and the Internet. For our measure of school broadband use, we aggregate the total traffic (upload plus download) of all schools in a municipality over the entire academic period.[14]

Internet use in schools grew significantly since the introduction of ADSL in late 2005 (Figure 3.4). Internet use grew from nearly zero in 2005 to 1.15 GB on average per student per year in 2009. The latter statistic corresponds to watching almost ten hours of YouTube video (at 260 Kbps), browsing 3,500 webpages (at 320 KB per page), or exchanging 8,500 emails (at 130 KB per email).[15] Broadband use per student exhibits high variability across municipalities (see Figure 3.6 for a histogram) and, as depicted in Figure 3.5, it grew more in low density areas, which is consistent with H_2. Overall, broadband use per student in

[14] We use as academic year the period between July 1st and June 30th.

[15] Average webpage size was obtained from *http://code.google.com/speed/articles/web-metrics.html*. We use the average email size of one of the authors as reference, as we found no reliable information on this statistic.

school is considerable.

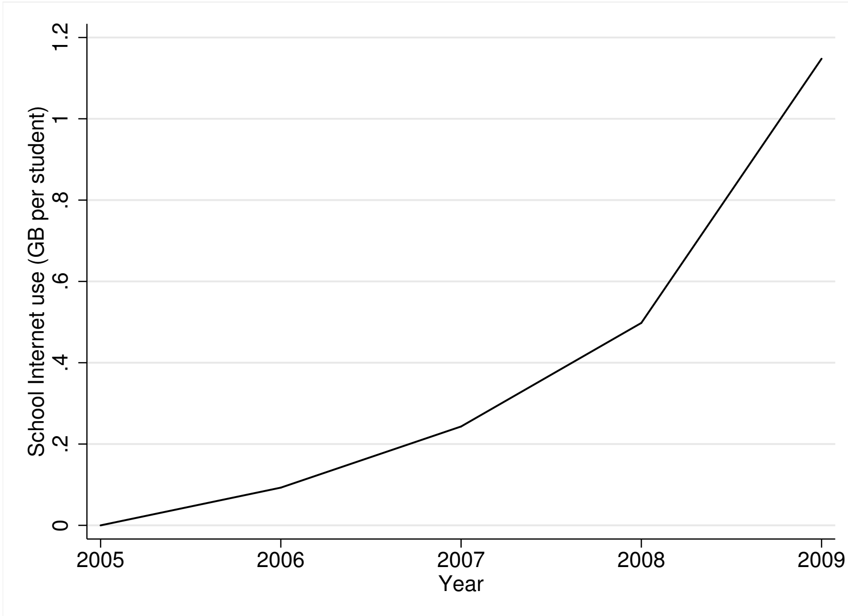


Figure 3.4: School Internet use per student.

Municipality Data

Finally, municipality data were provided by the Portuguese National Statistics Institute. These data include population density, average income, and population by age bracket across municipalities. Table 3.3 presents summary statistics for these variables as well as for household level variables.[16]

Municipalities exhibit also high variability in terms of percentage of households with children (see Figure 3.7).

[16]Portugal has a population of 10.6 million. The country is divided into 308 municipalities. Our sample covers schools in 195 municipalities.

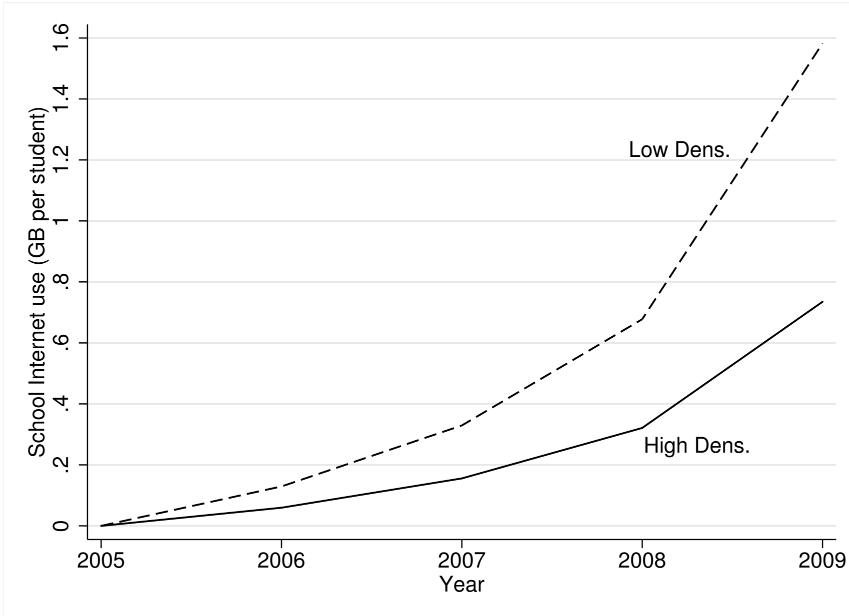


Figure 3.5: School Internet use per student by population density.

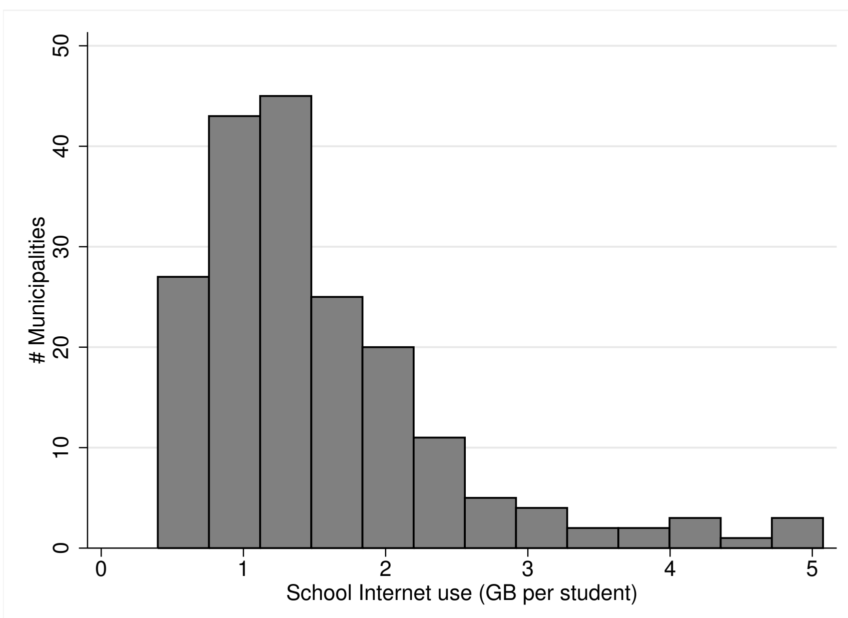


Figure 3.6: School Internet use per student in 2009 (GB).

Table 3.3: Summary statistics for municipality data.

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Home INet 2009	7,866	0.479	0.500	0	1
Broadband INet	7,866	0.457	0.498	0	1
Wireless INet	7,864	0.227	0.419	0	1
Computer (%)	7,868	0.560	0.496	0	1
Children (%)	7,868	0.276	0.447	0	1
Household Income	7,686	4.408	2.097	1	9
Household Size	7,868	2.704	1.204	1	14
School INet / Student (GB)	5,585	1.148	0.668	0.398	5.075
Pop. Dens. (Municipality)	7,868	1,297	1,822	5.600	7,183
Avg. Income (Municipality)	7,868	13.20	3.119	8.660	23.34
Age 0 to 14 (%)	7,868	0.151	0.0213	0.0676	0.239
Age 15 to 24 (%)	7,868	0.113	0.0151	0.0793	0.179
Age 25 to 64 (%)	7,868	0.556	0.0241	0.429	0.592
Distance to CO (Km)	5,585	1.322	0.446	0.351	2.612

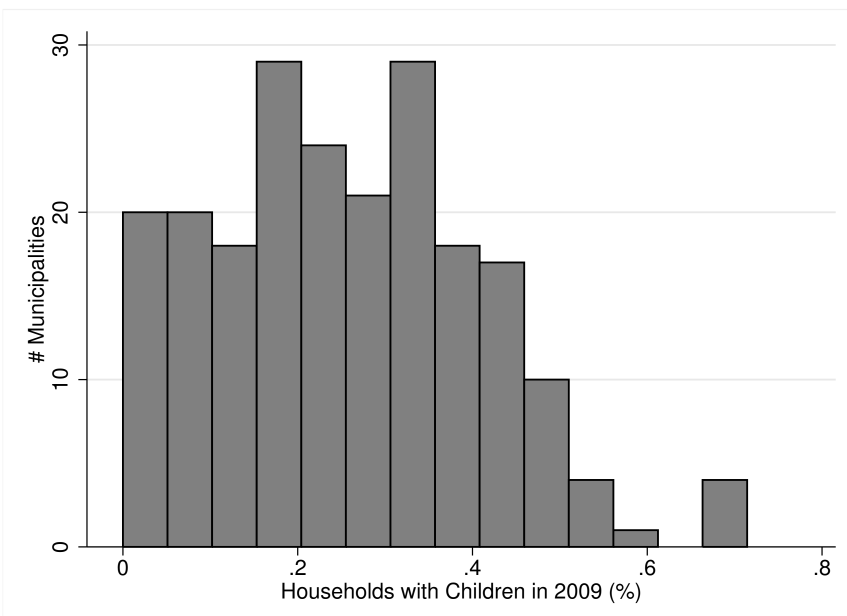


Figure 3.7: Households with Children in 2009 (%).

3.5 Empirical Specification

We use a Linear Probability Model (LPM) to estimate the effect of school Internet use on household Internet adoption. We use LPM rather than a binary response model (e.g., probit) because fewer assumptions are required to obtain consistency with the former and because the coefficients obtained can be directly interpreted as average marginal effects . Also, using LPM avoids several problems when applying instrumental variables further on.

From equations (3.4) and (3.5), the probability of adoption for household i in municipality j at time t given that no adoption has yet ocured is given by

$$\begin{aligned} P(u_a^*|I=1 \geq u_a^*|I=0)_{ijt} &= d_c \cdot \delta_c + d_c \cdot \phi_1 s_{cjt} + \\ &+ \phi_2 \frac{N_c}{N} s_{cjt} + \phi_3 \frac{N_c}{N} + \phi_4 t + \beta \mathbf{x}_{it} + \nu_j + \varepsilon_{it} \qquad (3.5) \end{aligned}$$

where d_c is an indicator variable for whether the household has children; δ_c is a constant that corresponds to the difference in the child's utility between having and not having Internet at home; ϕ_1 and ϕ_2 are the parameters of interest, corresponding to the household-level and municipality-level spillover effects, respectively; s_{jt} is the use of Internet at the municipality level; $\mathbf{x}_{it}$ is a household level vector of observed covariates (we decompose θ_{it} in equation 3.4 into a vector of observables, $\mathbf{x}_{it}$, and an unobserved term, ε_{it}); and ν_j is a municipality level unobserved term.

We start by estimating a pooled OLS for 2008 and 2009. We include household income, size and locality type as household level controls, as well as the percentage of households with children as a municipality level control. We use sampling weights and cluster the standard errors at the municipality level. We do not include municipality-level dummies because this would preclude us from estimating the municipality-level spillover effects. We have, however, performed similar regressions with municipality-level dummies and obtained similar results in all specifications. These results are available upon request.

If all right hand side variables in equation (3.5) are exogenous we can consistently estimate ϕ_1, ϕ_2, ϕ_3, ϕ_4 and β. However, our estimates for these coefficients are not consistent if school's Internet traffic, s_c, is endogenous, as our structural model suggests. Thus, we need to instrument school Internet traffic to obtain consistent estimates. We use two sets of instruments. First, we use municipality level variables as instruments for school Internet use . Second, we use an exogenous measure of Internet quality at the school to proxy the use of Internet at school. We detail each of these strategies below.

Instrumenting with municipality-level variables.

We instrument school Internet traffic with municipality-level variables, such as population density and average income levels (see, for example, Goolsbee and Klenow, 2002). These covariates predict well school Internet traffic as shown in our first-stage regressions in Table 3.5 in Appendix. However, they might not be valid instruments if correlated with unobserved household level variables. However, if we include the corresponding household level variables in the regression (e.g., household income), any bias due to unobservables would be captured by the household level variables, leaving the municipality-level coefficients unbiased.[17] For example, if technology savvy people tend to locate in high-income

[17]This is true as long as we assume that $E(\varepsilon_{it}|\mathbf{x}_{it}, \mathbf{w}_{jt}) = E(\varepsilon_{it}|\mathbf{x}_{it})$, where $\mathbf{w}_{jt}$ is a vector with municipality level variables. Potential household level unobservables would bias household variables but not

areas then including household income in our regressions would capture all the correlation between technology savvyness and income.

Instrumenting with broadband quality.

Alternatively, we exploit the variation in the quality of broadband connections across schools as an exogenous source of variation. Schools that benefit from a better connection to the Internet are more likely to use it more and therefore more likely to register more traffic. With ADSL technology, a greater distance between the costumer's premises and the ISP's Central Office (CO) results in a lower maximum transfer bitrate. Therefore, schools further away from the CO are likely to obtain lower throughput on their connection. Such lower throughput leads to degraded performance decreasing the attractiveness of the broadband connection at the school and thus lowering the amount of traffic exchanged with the Internet. Consequently, we use the average — weighted by the number of students in each school — of the line-of-sight distance between each school and its closest CO as a proxy for the quality of the school's broadband connection.[18]

Still, we posit that the distance between schools and COs might not be a valid instrument because it might be correlated with Internet availability in the region and, consequently, with home Internet penetration.[19] This positive correlation would bias upwards our estimates. Therefore, we use a proxy for household Internet penetration that is not correlated with distance to the CO . While wired broadband home Internet is likely to be provided over the same infrastructure put in place to provide Internet access to schools,

---

municipality level variables.

[18]Line-of-sight distance is calculated from information on the GPS coordinates of both schools and the ISP's closest COs.

[19]Although the distance between schools and COs might be unrelated with the distance between households and COs and thus the quality of broadband Internet in schools and households might be uncorrelated

wireless broadband Internet is provided not only by the ISP that wired schools but also, and to a large extent, by other ISPs in the country that use their own backbone infrastructure to connect cell towers. This breaks the correlation between the school's distance to the CO and wireless broadband penetration as table 3.7 in the Appendix shows. Therefore we use wireless Internet penetration as a proxy for home Internet penetration, which now renders the distance between schools and COs as a valid instrument.

Given that school Internet use shows up twice in equation (3.5), we need at least two excluded instruments to ensure identification. Thus, we use the interaction between the predicted Internet use and the *Children* indicator as additional instrument in both IV strategies.

3.6 Results

3.6.1 OLS Results

We start by presenting the cross-section regressions of school Internet traffic on home Internet penetration (see Table 3.4). Our dependent variable is binary and represents whether the household has Internet access or not; the independent variable, school Internet traffic, is continuous and is measured in Gigabytes (GB) per student in a given year. As in the model, it shows up twice in the regressions: interacted with the children indicator (SchInetPS×Children), and multiplied by the percentage of households with children (SchInetPS×ChildrenMuni (%)). As stated above, we use a set of household and municipality-level variables as additional controls. As expected, household income, household size, and population density are positively correlated with having Internet at home (column (1)). Also, having children is correlated with home Internet penetration (columns (3)-(4)). Having children at home is associated with an increase of 5% to 10% in the likelihood of having Internet at home. School Internet traffic, corresponding to the regional spillover effect described by Goolsbee and Klenow (2002), does not seem to be correlated with home Internet penetration in municipalities with higher percentage of households with children (columns (2) and (3)) (H_3). The interesting result is that school Internet use seems to be highly correlated with the home Internet penetration for households with children (column (4)). This suggests that the household level spillover effect is stronger than the municipality level effect. For each additional GB used by a student in a given year, the probability of having Internet for a household with children increases by 6.5%. This corroborates the hypothesis that there are spillover effects from schools to households and that children play a key role in the process (H_4). Additionally, it is not only availability at school that matters for the spillover effect, but it is the level of Internet usage (or quality) as well.

There are, however, several alternative explanations for the observed results. One such alternative is the reverse causality hypothesis. Suppose that households with children are more likely to adopt Internet — either because they are younger and more technology savvy, or because children ask for it — and this makes children get to school and make more use of the Internet when compared to children without Internet at home. School Internet use would increase *because* students have Internet at home, and not the other way around. To overcome the endogeneity problem we need to make use of exogenous sources

Table 3.4: Internet at home as a function of school Internet traffic.

VARIABLES	(1) HomeInet	(2) HomeInet	(3) HomeInet	(4) HomeInet
SchInetPSChildren				0.0651***
				(0.0227)
SchInetPSChildrenMuni		0.144***	0.0807*	0.0144
		(0.0315)	(0.0475)	(0.0504)
Children			0.103***	0.0513*
			(0.0189)	(0.0305)
ChildrenMuni			0.0597	0.113
			(0.0699)	(0.0700)
HouseholdIncome	0.183***	0.183***	0.185***	0.187***
	(7.92e-03)	(8.31e-03)	(8.05e-03)	(8.14e-03)
HouseholdSize	0.0937***	0.0951***	0.0717***	0.0715***
	(5.91e-03)	(6.37e-03)	(6.35e-03)	(6.33e-03)
LocalityType	-0.0439***	-0.0518***	-0.0434***	-0.0435***
	(8.66e-03)	(8.93e-03)	(9.12e-03)	(9.12e-03)
Year2009	0.0546***	0.0243	0.0375*	0.0445**
	(0.0181)	(0.0198)	(0.0205)	(0.0206)
Constant	-0.192***	-0.194***	-0.187***	-0.190***
	(0.0294)	(0.0305)	(0.0352)	(0.0353)
Observations	11,814	8,263	8,263	8,263
R-squared	0.307	0.313	0.319	0.320

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

of variation correlated with school Internet use, but not correlated with the unobserved error term. In any case, we should emphasize that the cross-section results clearly indicate that children mediate the observed positive correlation between school Internet use and home Internet penetration.

3.6.2 IV Results

Instrumenting with municipality-level variables

Table 3.5 shows the instrumental variables regressions using municipality-level variables and broadband quality as instruments.[20] Municipality-level IV results (columns (3) and (4)) are similar to OLS results (columns (1) and (2)), but higher in magnitude.[21] Again, school Internet use seems to affect home Internet penetration only for households with children. The IV coefficient (0.124) indicates that an increase of 1 GB in school Internet traffic per capita increases the probability of having Internet at home by 12.4 percentage points for households with children. In 2008 and 2009 school Internet use was on average 1.6 GB per student, which corresponds to an increase in the probability of Internet adoption of 20% in households with children, or 5% in the total population. As with OLS, we do not find evidence for municipality level effects.

[20] Table 3.8 in the Appendix includes the first-stage regressions.

[21] This difference between OLS and IV results is consistent with the hypothesis that there is a substitution effect between home Internet use and school Internet use: students that have Internet at home are likely to use less Internet at school. By instrumenting school Internet use we overcome this attenuation effect.

Table 3.5: Internet penetration as a function of school Internet traffic.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HomeInet				WirelessInet			
VARIABLES	OLS	OLS	Muni IV	Muni IV	OLS	OLS	Dist IV	Dist IV
SchInetPSChildren		0.0651***		0.124*		0.0868***		0.121***
		(0.0227)		(0.0643)		(0.0209)		(0.0402)
SchInetPSChildrenMuni	0.0807*	0.0144	0.371	0.240	0.126***	0.0379	0.494	0.368
	(0.0475)	(0.0504)	(0.275)	(0.307)	(0.0371)	(0.0373)	(0.331)	(0.334)
Children	0.103***	0.0513*	0.104***	0.00618	0.0611***	-0.00757	0.0626***	-0.0328
	(0.0189)	(0.0305)	(0.0192)	(0.0599)	(0.0121)	(0.0213)	(0.0123)	(0.0339)
ChildrenMuni	0.0597	0.113	-0.223	-0.118	-0.122**	-0.0517	-0.480	-0.379
	(0.0699)	(0.0700)	(0.264)	(0.289)	(0.0598)	(0.0600)	(0.328)	(0.332)
HouseholdIncome	0.185***	0.187***	0.184***	0.187***	0.0657***	0.0674***	0.0643***	0.0667***
	(0.00805)	(0.00814)	(0.00804)	(0.00860)	(0.00738)	(0.00738)	(0.00714)	(0.00724)
HouseholdSize	0.0717***	0.0715***	0.0709***	0.0705***	0.0391***	0.0388***	0.0380***	0.0376***
	(0.00635)	(0.00633)	(0.00651)	(0.00646)	(0.00464)	(0.00462)	(0.00475)	(0.00475)
LocalityType	-0.0434***	-0.0435***	-0.0690**	-0.0686**	-0.00499	-0.00503	-0.0374	-0.0371
	(0.00912)	(0.00912)	(0.0267)	(0.0266)	(0.00642)	(0.00644)	(0.0303)	(0.0305)
Year2009	0.0375*	0.0445**	-0.0138	0.000448	-0.00321	0.00620	-0.0682	-0.0545
	(0.0205)	(0.0206)	(0.0533)	(0.0564)	(0.0166)	(0.0166)	(0.0575)	(0.0579)
Constant	-0.187***	-0.190***	-0.101	-0.108	-0.142***	-0.146***	-0.0326	-0.0390
	(0.0352)	(0.0353)	(0.0940)	(0.0952)	(0.0264)	(0.0264)	(0.103)	(0.104)
Observations	8,263	8,263	8,263	8,263	8,261	8,261	8,261	8,261
R-squared	0.319	0.320	0.314	0.315	0.146	0.149	0.134	0.136

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Instrumenting with broadband quality

Applying distance to CO as instrument yields estimates consistent with OLS as well as with municipality-level variables IV results, despite the use of a different dependent variable (wireless Internet instead of home Internet). Household income and size are still statistically significant, but population density (LocalityType) is not. This might relate to the fact that wireless Internet is more readily available in more areas and not specifically in urban centers. Having children is important to have wireless Internet at home, as well as school Internet use in households with children. In general, the effect is bigger for wireless Internet than for home Internet, both in OLS and IV regressions. In the OLS regressions we see significance of school Internet for the households in general (column (5)), however this effect disappears when we include the household level $Children$ indicator: once again school Internet use is significant only for households with children. The IV coefficient (0.121) indicates that an increase of 1 GB in school Internet traffic per capita increases the probability of having wireless Internet at home by 12.1 percentage points for households with children. In 2008 and 2009 school Internet use increased the probability of adopting wireless Internet by 19% in households with children, which represents 5% in the total population.

3.6.3 Aggregating the data at the municipality-level

Given the nature of our data we cannot follow households over time, but we can follow municipalities. Thus, as an alternative estimation strategy we aggregate the data at the municipality level and for each municipality we calculate the gap in Internet adoption rates between households with and without children. We then regress the change in the gap between 2008 and 2009 on the changes in school Internet use over the same period:

$$\Delta(\bar{I}_{child} - \bar{I}_{nochild})_j \quad = \quad \Delta s_j + \beta \mathbf{x}_j + \varepsilon_j \tag{3.6}$$

where Δ represents difference between 2008 and 2009, $\bar{I}_{child}$ and $\bar{I}_{nochild}$ are the municipality-level home Internet adoption rates for households with and without children respectively, s_j is the school Internet use, $\mathbf{x}_j$ correspond to municipality-level covariates, β is a parameter vector, and ε_j is a municipality-level error term. This is a differences-in-differences setting with the particularity that one of the differences is a differences variable itself: the gap in Internet adoption rates between households with and without children.

Figure 3.8 shows a scatter of the change in the Internet adoption gap between households with children and without children, as a function of school Internet use change. The linear fit shows trend with a small but positive slope.

Table 3.6 shows the results of regressing Equation (3.6) in the aggregated data. The coefficient of interest has a similar magnitude as the ones obtained with the disaggregated data, but with a significance level of only 10%.

3.7 Conclusion

In this paper we look at the effects of providing broadband Internet to schools in terms of home Internet penetration. We find that besides its main purpose, this project generated

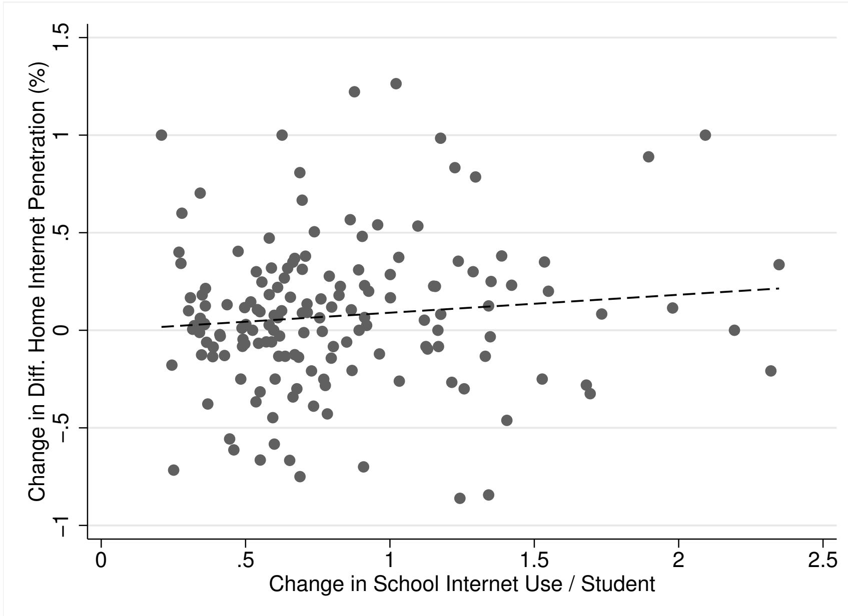


Figure 3.8: Change in the Internet adoption gap between households with children and without children as a function of school Internet use change.

Table 3.6: Summary statistics by Diff Regs

	(1)
VARIABLES	$\Delta(\bar{I}_{child} - \bar{I}_{nochild})$
Δ SchInetPS	0.160*
	(0.0865)
HouseholdIncome	-0.0367
	(0.0376)
LocalityType	-0.0790
	(0.0598)
PTFixedLine	-0.0203
	(0.0937)
Constant	0.222
	(0.202)
Observations	131

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

spillover effects in terms of Internet adoption in the surrounding communities.

We develop a structural model that provides insight on how Internet use at school affects home Internet penetration and how Internet penetration affects school Internet use. We use household level survey data to assess the net magnitude of each of these effects, and provide empirical evidence on the existence and magnitude of these spillovers.

We find evidence that children play a key role in home Internet adoption. We find that school broadband use contributes directly for a higher adoption rate in households with children. In 2008 and 2009 school Internet use increased the probability of adopting Internet by 20% in households with children, which represents an increase of 5% in the total population. For wireless Internet adoption, school Internet use increased the probability of adoption by 19%, representing 5% in the total population. We have found no evidence of a statistically significant effect at the municipality level.

Our study is not without limitations. The nature of the data does not allow building a traditional panel, leaving identification mainly dependent on the validity of our instruments. Even though we have tried to show evidence of their validity, we can never completely rule out the possibility of unobserved factors being correlated both with the instruments and with the dependent variable. We have run instrument validity tests obtaining encouraging results. We have also aggregated the data at the municipality level and run a diff-in-diff-in-diff setting yielding similar results in terms of magnitude, which gives us some more confidence on the results.

Finally, an increase in home Internet penetration does not necessarily mean that adults are now using more Internet. It might be that children are the ones using Internet at home.

In this case one could question the real value of these spillovers, a task that is out of the scope of this work.

3.A Appendix

Table 3.7: Internet at home as a function of distance to CO and other controls.

	(1)	(2)
VARIABLES	HomeInet	WirelessInet
Dist2COkm	-0.0432***	-2.13e-03
	(0.0137)	(0.0110)
Children	0.101***	0.0605***
	(0.0189)	(0.0120)
ChildrenMuni	0.157***	1.60e-03
	(0.0494)	(0.0407)
HouseholdIncome	0.184***	0.0661***
	(8.02e-03)	(7.48e-03)
HouseholdSize	0.0732***	0.0395***
	(6.32e-03)	(4.69e-03)
LocalityType	-0.0275***	6.55e-03
	(8.54e-03)	(6.67e-03)
Year2009	0.0508***	0.0190
	(0.0184)	(0.0156)
Constant	-0.174***	-0.178***
	(0.0330)	(0.0249)
Observations	8,263	8,261
R-squared	0.320	0.145

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.8: Internet penetration as a function of school Internet traffic.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		HomeInternet						Wireless			
		Muni IV	Muni IV	Muni IV		Dist IV	Dist IV	Dist IV	Dist & Muni IV	Dist & Muni IV	Dist & Muni IV
VARIABLES	OLS	1st Stg	1st Stg	2nd Stg	OLS	1st Stg	1st Stg	2nd Stg	1st Stg	1st Stg	2nd Stg
SchInetPSChildren	0.0653***			0.120***	0.0880***			0.147***			0.118***
	(0.0200)			(0.0407)	(0.0189)			(0.0551)			(0.0347)
SchInetPSChildrenMuni	0.0420			-0.887	0.0578			-0.0269			-0.0333
	(0.0845)			(0.583)	(0.0673)			(0.0748)			(0.451)
Children	0.0460*	-0.785***	-0.0700***	2.64e-03	-8.36e-03	1.110***	3.83e-03	-0.0551	-0.570***	-0.0659***	-0.0320
	(0.0243)	(0.0288)	(8.29e-03)	(0.0377)	(0.0215)	(0.0206)	(4.21e-03)	(0.0471)	(0.0321)	(0.0100)	(0.0323)
ChildrenMuni	-0.0630	0.0677	0.994***	0.852	-0.0717	0.0606	0.996***	0	0.0417	0.994***	0.0128
	(0.120)	(0.0448)	(0.0220)	(0.579)	(0.0991)	(0.0541)	(0.0226)	(0)	(0.0447)	(0.0220)	(0.449)
HouseholdIncome	0.178***	-8.38e-03***	2.33e-03**	0.181***	0.0665***	-0.0186***	1.94e-03*	0.0678***	-7.91e-03***	2.34e-03**	0.0672***
	(6.49e-03)	(2.24e-03)	(1.10e-03)	(6.80e-03)	(7.01e-03)	(3.43e-03)	(1.10e-03)	(7.05e-03)	(2.25e-03)	(1.10e-03)	(7.09e-03)
HouseholdSize	0.0770***	1.39e-04	-4.87e-05	0.0770***	0.0390***	3.23e-03	2.49e-04	0.0387***	4.77e-05	-5.04e-05	0.0389***
	(6.25e-03)	(2.50e-03)	(8.33e-04)	(6.26e-03)	(5.28e-03)	(3.17e-03)	(8.36e-04)	(5.27e-03)	(2.49e-03)	(8.32e-04)	(5.27e-03)
LocalityType	-0.0576***	8.43e-04	5.85e-05	-0.0621***	-0.0286	-7.52e-03	-5.15e-03**	-0.0287	3.53e-04	4.92e-05	-0.0289*
	(0.0222)	(6.68e-03)	(2.42e-03)	(0.0223)	(0.0175)	(9.12e-03)	(2.59e-03)	(0.0175)	(6.67e-03)	(2.42e-03)	(0.0176)
HouseIncomeMuni		-0.105***	-0.0189***						-0.105***	-0.0189***	
		(9.00e-03)	(3.93e-03)						(9.00e-03)	(3.93e-03)	
LocalityTypeMuni		-0.373***	-0.235***						-0.359***	-0.235***	
		(0.0605)	(0.0287)						(0.0598)	(0.0287)	
HouseIncomeMuniChild		0.220***	0.0146***						0.224***	0.0147***	
		(5.67e-03)	(1.61e-03)						(5.72e-03)	(1.62e-03)	
LocalityTypeMuniChild		0.463***	0.0103***						0.400***	9.14e-03**	
		(0.0131)	(3.18e-03)						(0.0134)	(3.62e-03)	
DistanceIV						0.305	0.0511		-0.606**	-0.516***	
						(0.258)	(0.0794)		(0.301)	(0.105)	
DistanceIVChildren						-0.774***	-9.36e-03		-0.297***	-5.60e-03	
						(0.0327)	(6.88e-03)		(0.0252)	(7.58e-03)	
Year2009	0.0433*	0.189***	0.187***	0.191**	5.81e-03	0.0592***	0.163***	0.0164	0.189***	0.187***	0.0190
	(0.0252)	(0.0174)	(7.29e-03)	(0.0962)	(0.0210)	(0.0111)	(3.66e-03)	(0.0232)	(0.0173)	(7.29e-03)	(0.0747)
Constant	-0.0583	0.602***	7.08e-03	-0.203	-0.146*	-0.119	-0.148***	-0.123	1.254***	0.605***	-0.0894
	(0.118)	(0.0842)	(0.0413)	(0.169)	(0.0784)	(0.0931)	(8.76e-03)	(0.146)	(0.210)	(0.0888)	(0.150)
Observations	8,263	8,265	8,265	8,263	8,261	8,265	8,265	8,261	8,265	8,265	8,261
R-squared	0.345	0.840	0.903	0.333	0.177	0.758	0.900	0.176	0.843	0.903	0.177
Muni. Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.9: Internet penetration as a function of school Internet traffic, by Income.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	Muni IV	Muni IV	OLS	OLS	Dist IV	Dist IV
VARIABLES	Low Income	High Income	Low Income	High Income	Low Income	High Income	Low Income	High Income
SchInetPSChildren	0.0344	0.0831***	0.115*	0.0949	0.129***	0.0563*	0.181***	-2.15e-03
	(0.0308)	(0.0283)	(0.0628)	(0.0724)	(0.0234)	(0.0325)	(0.0448)	(0.0792)
SchInetPSChildrenMuni	0.0525	-0.0348	-1.058	-0.866	0.0591	0.0169	-0.500	-0.191
	(0.105)	(0.138)	(0.801)	(1.382)	(0.0686)	(0.142)	(0.548)	(1.206)
Children	0.132***	-0.0139	0.0626	-0.0211	-0.0475*	0.0168	-0.0916**	0.0603
	(0.0410)	(0.0308)	(0.0628)	(0.0564)	(0.0260)	(0.0347)	(0.0419)	(0.0651)
ChildrenMuni	-0.136	0.0291	0.996	0.803	-0.0579	-0.0453	0.509	0.153
	(0.151)	(0.191)	(0.828)	(1.279)	(0.101)	(0.191)	(0.566)	(1.118)
HouseIncome	0.139***	0.144***	0.139***	0.145***	0.0352***	0.103***	0.0348***	0.103***
	(0.0146)	(0.0102)	(0.0148)	(0.0104)	(9.52e-03)	(0.0113)	(9.62e-03)	(0.0115)
HouseholdSize	0.0862***	0.0584***	0.0870***	0.0581***	0.0403***	0.0381***	0.0407***	0.0380***
	(9.82e-03)	(8.54e-03)	(9.84e-03)	(8.60e-03)	(6.52e-03)	(8.71e-03)	(6.50e-03)	(8.74e-03)
LocalityType	-0.0831***	-0.0167	-0.0903***	-0.0198	-0.0207	-0.0360	-0.0243	-0.0374
	(0.0285)	(0.0364)	(0.0291)	(0.0365)	(0.0194)	(0.0329)	(0.0195)	(0.0332)
Year2009	0.0303	0.0270	0.191	0.148	0.0463***	0.145***	0.126	0.188
	(0.0211)	(0.0257)	(0.122)	(0.200)	(0.0131)	(0.0263)	(0.0832)	(0.174)
Constant	0.0725	-0.494***	0.0718	-0.608***	0.232	-0.400***	0.233	-0.438**
	(0.237)	(0.121)	(0.230)	(0.224)	(0.291)	(0.108)	(0.284)	(0.199)
Observations	4,745	3,518	4,745	3,518	4,744	3,517	4,744	3,517
R-squared	0.221	0.187	0.198	0.176	0.137	0.156	0.123	0.154
Muni. Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Chapter 4

The Role of Young People on Adults' Computer and Internet Use Patterns and Skills

Abstract: Previous studies have shown that the presence of children in the household contributes to higher levels of Internet adoption. In this paper we go a step further and try to assess the effects of living with children or young adults on the likelihood of computer and Internet use (as opposed to adoption) and on skills acquisition. We make use of a detailed individual level and household level survey on computer and Internet use and skills. In order to alleviate potential selection issues in our data, namely the fact that skill acquisition is contingent on use and use is contingent on adoption, we apply a three-step selection model. We provide empirical evidence that the presence of children or young adults in the household does contribute to an increase in the likelihood of having a

computer or Internet at home, but does not contribute to an increase in use patterns and skills. Moreover, we find that the presence of children and young adults is associated with lower levels of computer and Internet use and skills.

4.1 Introduction

The increasing availability of computers and Internet in the household pushed the digital divide debate to a new level, where the effective use of technology and skills become more important than access. It has been shown that there are significant discrepancies in the ways people use ICT technologies, and this is true even within the same socioeconomic and educational backgrounds (Hargittai, 2010). This heterogenity is often referred to as digital inequality (DiMaggio et al., 2001) or second level digital divide (Hargittai, 2002). Therefore it important to understand which factors contribute to digital inequality and what actions can be taken to reduce it. Answering this question has important policy implications for the use of Internet and for the digital divide debate.

Previous studies have shown that providing Internet to schools has a positive impact on Internet adoption at the household level, specifically in households with children (Belo et al., 2012b; Ward, 2010), the main argument being that children would get acquainted with the technology at schools and ask their parents to adopt it at home. Similarly, the presence of young people in the household can have an impact in how older adults use technologies such as computers and Internet. In this study we explore the possibility that there is knowledge transmission among members of a household, and whether this knowledge transmission occurs from children or young adults to older adults.

Living with young people can affect computer and Internet skills in ways that are not related with knowledge transmission. For example, children and young adults can spend more time using the computer and Internet, making it more difficult for older adults toe use these resources and to learn from their use. Another alternative to the knowledge transmission hypothesis is that adults may rely on young people to get the services they need from computers and Internet without needing to learn how to use these technologies.

We use survey data at the individual and household level to test the outlined hypotheses, and provide empirical evidence that the presence of children or young adults in the household does contribute to an increase in the likelihood of having a computer and Internet at home, but does not contribute to an increase in use patterns and skills. Moreover, we find that the presence of children and young adults is associated with lower levels of computer and Internet use and skills. We apply a Heckman selection model to control for potential selection on technology adoption and use, and the results remain the same in qualitative terms. To further explore the potential effects of living with children and young adults we partition the data into age, education and income bracket dummies, and perform t-tests on each partition. We find that living with a young adult is associated with lower levels of computer and Internet use and skills but only for individuals with low levels of education and in the age bracket of 25 to 44.

This paper is structured as follows. Section 4.2 revises and extends the related literature on household technology adoption and skill acquisition, section 4.3 presents used data and delineates the empirical strategy, section 4.4 presents the results and 4.5 concludes.

4.2 Household technology adoption and skill acquisition

Technology adoption has been a topic of research in economics, psychology and sociology for a long time. Its origins can be traced to the works of Rogers (2003) — who classifies adopters into categories depending on their time adoption — and Bass (1969) — that develops an adoption model in which new adopters are influenced by the proportion of previous adopters. In the psychology and sociology literatures, the technology adoption model (TAM) (Davis, 1985) has been a reference to explain the factors that lead people to adopt a given technology. This model and others alike assume perceived easiness of use, perceived usefulness, attitudes towards using the technology, and behavioral intention of using the technology, among others, as the main explanatory variables for adtopion. Venkatesh and Brown (2001) apply this model specifically to household PC adoption. They find that utilitarian and hedonic factors are the major drivers for adoption, while factors for non-adoption are related with rapid change of technology and cost. In their study the role of children as drivers for PC purchase is very limited. Only under 2% of adopter respondents indicate children to be an important factor in their decision. These models explain the potential factors that lead people to adopt a given technology or product, but they do not explicitly focus on the relations that exist in the household and on the potential for knowledge transfer between members of the household.

Selwyn et al. (2005) surveyed 1,001 adults and performed 100 semi-structured interviews on a follow-up study on their Internet use habits. From their interviews they find that it is frequent that adults acquire a computer and Internet connection for their children, but end up using them themselves. In their study there is evidence that in some situations adults and children use the computer and Internet together, but there is no other evidence that knowledge transmission occurs.

One of the most significant theoretical contributions concerning knowledge transmission and skill acquisition, is introduced by DiMaggio et al. (2001) where they argue that, as Internet access becomes more ubiquitous, the term “digital divide” gives place to “digital inequality” as the most important variable to measure. This refinement is referred to as the second level digital divide (Hargittai, 2002). They identify five dimensions of digital inequality — equipment, autonomy of use, skill, social support and purpose of use — and set an agenda for future research that is in part followed by studies by Eszter Hargittai and co-authors (e.g., Hargittai, 2002, 2007; Hargittai and Hinnant, 2008; Hargittai, 2010). In these studies they find a large variation among young people in terms of use patterns and skills, and find that not all young people are savvy in what respects the use of ICTs. Even controlling for Internet access, their socioeconomic background has a significant influence on how they make use of the technology (Hargittai, 2010). These studies do not focus, however, on knowledge transmission among members of a household. Given the focus of our study on the role of children and young adults in computer and Internet use and skills, we concentrate on the role of social support, more specifically in the form of technical assistance and emotional reinforcement by family members as defined by DiMaggio et al. (2001). Namely, we test their hypothesis that “social support of all kinds increases users’ motivation to use the technology and the extent to which they develop their own digital competence".

Besides the above-stated hypothesis, there is little theory about the existence and nature of knowledge transmission between children and older adults. We refine the knowledge transmission hypothesis presented by DiMaggio et al. (2001), and develop additional alternative hypotheses on the effects of living with children and young adults on adults' computer and Internet use and skills.

As mentioned by DiMaggio et al. (2001), social support may be an important catalyst for computer and Internet use and consequent skill acquisition. Given that children and young adults are likely to exhibit higher levels of computer and Internet skills, when compared with their parents and other adults in the same household, the presence of a child or young adult in the household might foster the transmission of knowledge learned elsewhere, for instance at school (Belo et al., 2012a), to the adult. These effects can be different for children and young adults, due to the different nature of activities they perform using the technology and the degree of knowledge they have. Children are more likely to require the presence of an adult to aid them in performing some specific task, and are more likely to be monitored by their parents regarding their activity in the Internet. This may lead adults to try learning how to use a computer and Internet in order to help their children. Young adults, on the other hand, are more likely to be independent and to use computer and Internet by themselves without an older adult overseeing their activities. This may lead to a lower level of adult's use and skills, as adults are not required to acquire skills in order to assist younger adults. However, the presence of a young adult can be valuable in situations where older adults need help to complete a specific task, given young adults' knowledge.

Thus, all of these mechanisms are associated with higher computer and Internet use and skills for older adults living with children or young adults.

Alternatively, children and young adults may monopolize computer and Internet access, preventing adults from using the technology and learning from it. This resource sharing effect would be associated with a negative correlation between living with a child or young adult and computer and Internet use and skills. This effect would be more prevalent in households with lower income levels, given that in these households the likelihood of computer sharing would be higher. Another mechanism that is possibly at play is that people with children have less free leisure time to dedicate to computer and Internet use. Yet another alternative scenario is that adults rely on children's or young adult's computer and Internet skills and feel they do not need to learn how to use these technologies given they will be able to ask their children to perform the necessary tasks for them (e.g., performing online orders of printing information about a specific topic). All of these mechanisms are associated with lower computer and Internet use and skills for adults that live with children or young adults.

In summary, there are several competing alternative explanations for the effects of living with children and young adults. Additionally, identification becomes harder in the presence of selection. For instance, if households that adopt Internet are intrinsically different from households that do not adopt, estimating the effects of living with children on Internet use in households with Internet may lead to biased estimates. The problem gets worse if the independent variables of interest are endogenous. For instance, if households with young adults are per se different than households without young adults even after controlling for all observables, all the estimates can be biased as well. These scenarios are discussed in the next section.

4.3 Data and Empirical Framework

To test the alternative explanations enumerated above, we make use of very detailed survey data administered to households in Portugal by the Portuguese National Statistics Institute (PNSI) in 2008. PNSI administers this survey on a yearly basis to track the use of Information and Communication Technologies (ICT) in the country. This survey is administered since 2003, and from year to year 3/4 of the sample are kept from the previous year while 1/4 are new households. Household identifiers are available only for 2008 and thus we are not able to construct true a panel or know household composition for other years, so we restrict our analysis to 2008. Surveyed computer skills include being able to create files, perform a copy&paste operation, use a spreadsheet software, compress and uncompress files, and use a programming language. Internet skills include being able to use a search engine, to send an email with attachment, to use instant messaging services, to perform a phone call, to use peer-to-peer tools, and to build a web-page For 2008 all individuals in a household are associated to a household identifier, so it is possible to know computer and Internet use patterns and skills of all members of a household. In 2008 4,116 households were surveyed, corresponding to 7,680 older adults (> 24 years old) and 1,102 young adults (16 to 24 years old). Along with individual and household level data, we use municipality level data publicly available from the Portuguese National Statistics Institute. These data include population density, average income, and population by age bracket across municipalities.

Table 4.1 presents some descriptive statistics. *InetHouse*, *InetUse* and *CompUse* are dicotomous variables representing, respectively, whether an individual has Internet at home or not, whether she uses it, and whether she uses a computer. *SkillsInet* and *SkillsComp* are count variables representing the number of skills individuals reported having regardin Internet and computer use, respectively. *Age* and *EducLevel* represent age and education brackets. Age bracket 2 corresponds to [25, 34]; 3 to [35, 44]; 4 to [45, 54]; 5 to [55, 64]; and 6 corresponds to 65 and older. Education bracket 1 corresponds to basic education, 2 corresponds to high school education, and 3 corresponds to college education or higher. Adults living in households with children are more likely to have Internet and use it. They also exhibit higher levels of Internet and computer skills. In contrast, adults that live with young adults are more likely to have Internet at home, but are less likely to use it, which suggests that Internet adoption in these households is driven mainly by the young adults, and might point to the lack of knowledge transmission, or skill acquisition by adults.

We model computer and Internet adoption, use, and skills acquisition as a three-stage process. In the first stage the household decides whether or not to adopt the technology (i.e., computer or Internet), in the second stage each member of the household decides whether to use the technology or not, and finally in the third stage the individual decides how much time to spend using the technology and acquiring new skills. Due to the nature of the process, information on skills is contingent on technology use, and information on use is contingent on technology adoption. In the third stage the individual acquires skills based on how much time she spends using the technology and on how easy it is to acquire

Table 4.1: Summary statistics for adults (> 24): Internet use and skills by household composition.

VARIABLES	(1) All Total (s.d.)	(2) Children (s.d.)	(3) Young A. (s.d.)	(4) InetHouse Total (s.d.)	(5) Children (s.d.)	(6) Young A. (s.d.)
InetHouse	0.471	0.623	0.721			
	(0.499)	(0.485)	(0.449)			
InetUse	0.319	0.450	0.330	0.558	0.617	0.424
	(0.466)	(0.498)	(0.470)	(0.497)	(0.486)	(0.494)
SkillsInet	1.095	1.502	1.033	2.027	2.152	1.325
	(1.952)	(2.096)	(1.822)	(2.339)	(2.262)	(1.961)
CompUse	0.379	0.523	0.396	0.613	0.672	0.496
	(0.485)	(0.500)	(0.489)	(0.487)	(0.470)	(0.500)
SkillsComp	1.396	1.892	1.360	2.440	2.595	1.715
	(2.261)	(2.436)	(2.158)	(2.616)	(2.585)	(2.297)
Age	4.155	3.251	3.787	3.728	3.268	3.787
	(1.342)	(1.003)	(0.997)	(1.194)	(0.941)	(0.960)
EducLevel	1.291	1.335	1.235	1.518	1.482	1.306
	(0.632)	(0.659)	(0.571)	(0.781)	(0.753)	(0.638)
Observations	7,680	2,299	1,757	3,616	1,432	1,266

new skills. Skills acquisition can be easier or harder depending on age, education level and other factors, but also on whether there is someone knowledgeable nearby to help with the difficulties. Children and young adults can fulfill this role. In the second stage, the individual will decide to use the technology if her expected utility is increased by the use of the technology: $Pr_i(use_i) = \Phi(u_i|\text{Use} = 1 - u_i|\text{Use} = 0 > 0)$, where u_i is the utility of individual i. On the first stage the decision to adopt the technology is taken by the household as a whole which is influenced by expected usage by its members, price and other availability constraints: $Pr_h(adopt_h) = \Phi(u_h|\text{adopt} = 1 - u_h|\text{adopt} = 0 > 0)$, where u_h is the utility of household h.

Therefore, there are two levels of selection in this process. The first selection occurs at the household level, where households decide whether to buy the technology or not, and the second selection occurs in a household that has the technology, in which individuals decide whether to use it or not. Finally, for those individuals that decide to use the technology, living with a child or young adult might have an impact in skills acquisition, the effect set out to explore in this study. Equations (4.1)-(4.3) represent the described selection model.

$$\begin{aligned} skills_i &= \mathbf{x}_{1i}\beta_1 + \varepsilon_{1i} & (4.1) \\ use_i &= 1[\mathbf{x}_{2i}\delta_2 + \varepsilon_{2i} > 0] & (4.2) \\ adopt_h &= 1[\mathbf{x}_{3i}\delta_3 + \varepsilon_{3i} > 0] & (4.3) \end{aligned}$$

where $\mathbf{x}_{1i}, \mathbf{x}_{2i}, \mathbf{x}_{3i}$ are covariate vectors, $\beta_1, \delta_2, \delta_3$ are parameter vectors and $\varepsilon_{2i}, \varepsilon_{2i}, \varepsilon_{3i}$ are error terms.

To estimate the effect of living with children and young adults in the likelihood of using a given technology, we estimate the binary response model with sample selection represented by equations (4.2) and (4.3). Households select into adopting or not the technology; individuals in adopting households decide whether to use the technology or not. We can estimate this model by assuming that the errors are bivariate normal and independent of the explanatory variables. Since we only observe usage upon adoption, we estimate the following second stage equation:

$$use_i = 1[\mathbf{x}_{2i}\delta_2 + \lambda_2 \frac{\hat{\phi_{2i}}}{\Phi_{2i}} + \varepsilon_{2i} > 0] \tag{4.4}$$

where $\frac{\hat{\phi_{2i}}}{\Phi_{2i}}$ are the estimates of the inverse Mills ratios over the population, and λ_2 is the corresponding coefficient.

Due to the non-linearities of the probit models, identification is contingent on $\mathbf{x}_3$ containing at least one variable not present in $\mathbf{x_2}$. This means that the selection equation must contain a variable not in the *use* equation (Wooldridge, 2002, pp. 570-571). Given that selection occurs at the household level, we restrict $\mathbf{x}_3$ to household and municipality level variables. In particular, we include average household age and education level, household income, locality type, population density and municipality level average Internet quality, proxied by the average distance between schools and the ISP's central offices.

Equations (4.1) and (4.2) represent the typical Heckman selection model, or a type II Tobit model, where skills are observed only for those that use the technology. Assuming no other selection issues, we could consistently estimate β_1 by including the estimates of the inverse Mills ratios over individuals in households that adopted the technology, $\frac{\hat{\phi_{1i}}}{\Phi_{1i}}$, in equation (4.1) and estimating it. However, individuals with in a household that adopted the technology are already a selection from the whole population. Therefore, to correctly estimate β_1, we need to include the estimates for the inverse Mills ratios in the first selection equation as well:

$$skills_i = \mathbf{x}_{1i}\beta_1 + \lambda_1 \frac{\hat{\phi_{1i}}}{\Phi_{1i}} + \lambda_2 \frac{\hat{\phi_{2i}}}{\Phi_{2i}} + \varepsilon_{1i} \tag{4.5}$$

where $\frac{\hat{\phi_{1i}}}{\Phi_{1i}}$ are the estimates of the inverse Mills ratios over individuals in households that adopted the technology, and λ_1 is the corresponding coefficient.

The Heckman correction solves the selection problem in the case the instruments are correlated with the selection but not correlated with the use equation, or in the case the errors are normal. Given that we can not rule out that the instruments are correlated with the use equation or that the errors are not normal, we interpret the corresponding results with caution.

4.4 Results

4.4.1 OLS estimates

Table 4.2 shows multivariate regressions for the likelihood of having Internet at home as a function of household composition, education level, age, household income, and locality type. Higher levels of income, number of people in the household, and education are associated with a higher likelihood of having Internet at home. Living with children and young adults is positively associated with having Internet at home. This indicates that children and young adults are a strong driver for having Internet at home. The coefficients for children and young adults in column (5), 0.05 and 0.246, respectively, mean that living with a child increases the probability of having Internet by about 5%. This effect is 25% for young adults.

Focusing only in households with Internet, we look at Internet use as a function of living with children and young adults (Table 4.3). After controlling for education and age we find no association between Internet use and living with children, but we see a negative association for young adults. This suggests that once a household has Internet, children do not influence older adults' likelihood of using Internet, but for young adults this is not the case. The negative coefficient means that either (i) there is a negative effect of living with young adults on Internet use, or that (ii) household with young adults select into adopting Internet even if older adults have no intention of using it. We have explored

Table 4.2: Internet at home as a function of household composition.

VARIABLES	(1) InetHouse	(2) InetHouse	(3) InetHouse	(4) InetHouse	(5) InetHouse
Children		0.101***	0.0326	0.106***	0.0482**
		(0.0215)	(0.0235)	(0.0210)	(0.0230)
YoungAdults				0.260***	0.246***
				(0.0222)	(0.0228)
HouseIncome2	0.192***	0.218***	0.189***	0.182***	0.158***
	(0.0270)	(0.0280)	(0.0270)	(0.0266)	(0.0257)
HouseIncome3	0.393***	0.466***	0.390***	0.402***	0.336***
	(0.0304)	(0.0307)	(0.0305)	(0.0304)	(0.0300)
HouseIncome4	0.566***	0.675***	0.562***	0.603***	0.495***
	(0.0354)	(0.0342)	(0.0355)	(0.0346)	(0.0354)
HouseIncome5	0.601***	0.757***	0.595***	0.698***	0.524***
	(0.0336)	(0.0295)	(0.0339)	(0.0297)	(0.0342)
PopDensMed	-0.0252	-0.0390*	-0.0261	-0.0550**	-0.0369*
	(0.0221)	(0.0230)	(0.0221)	(0.0224)	(0.0216)
PopDensLow	-0.0522**	-0.0671***	-0.0519**	-0.0840***	-0.0644***
	(0.0232)	(0.0241)	(0.0233)	(0.0231)	(0.0223)
Age35to44	0.0558**		0.0501*		0.0155
	(0.0253)		(0.0258)		(0.0259)
Age45to54	0.0576**		0.0648***		0.00527
	(0.0251)		(0.0249)		(0.0251)
Age55to64	-0.104***		-0.0909***		-0.0854***
	(0.0240)		(0.0244)		(0.0241)
Age65to74	-0.246***		-0.233***		-0.211***
	(0.0257)		(0.0263)		(0.0259)
EducHigh	0.139***		0.141***		0.166***
	(0.0238)		(0.0237)		(0.0232)
EducCollege	0.168***		0.171***		0.213***
	(0.0233)		(0.0233)		(0.0238)
Constant	0.231***	0.153***	0.219***	0.144***	0.216***
	(0.0340)	(0.0270)	(0.0343)	(0.0258)	(0.0334)
Observations	7,489	7,489	7,489	7,489	7,489
R-squared	0.296	0.243	0.297	0.293	0.337

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

these alternatives in the previous section. As with the likelihood of having Internet at home, income and education levels are also positively associated Internet use, however the number of people in the household is negatively associated with Internet use.

Tables 4.4 and 4.5 show the likelihood of having specific computer and Internet skill as a function of household composition for households with Internet. Surveyed computer skills include being able to create files, perform a copy&paste operation, use a spreadsheet software, compress and uncompress files, and use a programming language. Internet skills include being able to use a search engine, to send an email with attachment, to use instant messaging services, to perform a phone call, to use peer-to-peer tools, and to build a web-page. In general, living with children or young adults is associated with lower levels of computer and Internet skills. Again, this can be justified by selection, meaning that households with children and young adults are more likely to adopt a technology even though in general use and skills are lower.

Tables 4.2, 4.3, 4.4 and 4.5 are summarized in Table 4.6. Living with young adults is associated with a lower levels of computer and Internet use as well as corresponding skills. Living with children is associated with lower computer and Internet skills but not associated with lower use patterns.

In Table 4.6 we look at the association between living with children and young adults and Internet and computer use and skills only in the households that have adopted the technology and among the individuals that use it. Our interest lies, however, in the effect of children and young adults in households with the technology as compared with the effect in households without the technology. In Table 4.7 we assume that individuals without

Table 4.3: Internet use in households with Internet.

VARIABLES	(1) InetUseHome	(2) InetUseHome	(3) InetUseHome	(4) InetUseHome	(5) InetUse
Children		0.118***	0.0235	0.0932***	0.00828
		(0.0228)	(0.0222)	(0.0221)	(0.0220)
YoungAdults				-0.192***	-0.106***
				(0.0228)	(0.0219)
HouseIncome2	0.0935*	0.0986*	0.0925*	0.116**	0.102*
	(0.0521)	(0.0570)	(0.0521)	(0.0544)	(0.0529)
HouseIncome3	0.132**	0.183***	0.132**	0.204***	0.144***
	(0.0517)	(0.0564)	(0.0516)	(0.0538)	(0.0525)
HouseIncome4	0.197***	0.316***	0.197***	0.332***	0.210***
	(0.0541)	(0.0598)	(0.0540)	(0.0573)	(0.0551)
HouseIncome5	0.203***	0.449***	0.201***	0.454***	0.216***
	(0.0549)	(0.0582)	(0.0550)	(0.0555)	(0.0556)
PopDensMed	-0.0571**	-0.111***	-0.0582**	-0.0922***	-0.0523**
	(0.0227)	(0.0262)	(0.0227)	(0.0256)	(0.0226)
PopDensLow	-0.0813***	-0.124***	-0.0819***	-0.0955***	-0.0697***
	(0.0244)	(0.0281)	(0.0245)	(0.0270)	(0.0241)
Age35to44	-0.110***		-0.117***		-0.0971***
	(0.0270)		(0.0288)		(0.0284)
Age45to54	-0.276***		-0.273***		-0.239***
	(0.0283)		(0.0282)		(0.0288)
Age55to64	-0.352***		-0.345***		-0.342***
	(0.0294)		(0.0295)		(0.0296)
Age65to74	-0.471***		-0.465***		-0.464***
	(0.0328)		(0.0328)		(0.0327)
EducHigh	0.364***		0.364***		0.351***
	(0.0275)		(0.0274)		(0.0273)
EducCollege	0.429***		0.430***		0.409***
	(0.0270)		(0.0272)		(0.0277)
Constant	0.441***	0.269***	0.432***	0.322***	0.448***
	(0.0562)	(0.0561)	(0.0565)	(0.0536)	(0.0572)
Observations	3,528	3,528	3,528	3,528	3,528
R-squared	0.334	0.108	0.334	0.141	0.343

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.4: Computer Skills.

VARIABLES	(1) SkillsComp	(2) File	(3) CopyPaste	(4) Excel	(5) ZipFiles	(6) Progmng
Children	-0.459***	-0.0210	-0.0458**	-0.0472*	-0.100***	-0.0483**
	(0.114)	(0.0171)	(0.0225)	(0.0267)	(0.0270)	(0.0200)
YoungAdults	-0.404***	-0.00283	-0.0359	-0.0753**	-0.0500*	-0.0236
	(0.122)	(0.0205)	(0.0246)	(0.0297)	(0.0289)	(0.0188)
Constant	4.219***	0.940***	0.809***	0.526***	0.491***	0.141**
	(0.330)	(0.0523)	(0.0787)	(0.0754)	(0.0765)	(0.0551)
Observations	2,366	2,366	2,366	2,366	2,366	2,366
R-squared	0.319	0.093	0.155	0.182	0.255	0.085
Age Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Education Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Household Income Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Density Dummies	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.5: Internet Skills.

VARIABLES	(1) SkillsInet	(2) Google	(3) Attachmnt	(4) Chat	(5) Phone	(6) P2P	(7) Webpage
Children	-0.461***	0.00454	-0.0609**	-0.0763**	-0.0367	-0.110***	-0.0601**
	(0.126)	(0.0119)	(0.0247)	(0.0309)	(0.0339)	(0.0280)	(0.0247)
YoungAdults	-0.483***	0.0113	-0.0901***	-0.0995***	-0.0929***	-0.00896	-0.0384*
	(0.124)	(0.0148)	(0.0290)	(0.0291)	(0.0334)	(0.0260)	(0.0208)
Constant	4.029***	0.958***	0.640***	0.427***	0.335***	0.381***	0.168***
	(0.449)	(0.0326)	(0.130)	(0.0820)	(0.0855)	(0.0832)	(0.0614)
Observations	1,971	1,971	1,971	1,971	1,971	1,971	1,971
R-squared	0.308	0.037	0.158	0.167	0.105	0.171	0.083
Age Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
House Inc. Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dens. Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.6: Computer and Internet use and skills as a function of household composition.

VARIABLES	(1) CompUse	(2) SkillsComp	(3) InetUseHome	(4) SkillsInet
Children	0.0243	-0.459***	8.28e-03	-0.461***
	(0.0218)	(0.114)	(0.0220)	(0.126)
YoungAdults	-0.142***	-0.404***	-0.106***	-0.483***
	(0.0232)	(0.122)	(0.0219)	(0.124)
Constant	0.601***	4.219***	0.448***	4.029***
	(0.0710)	(0.330)	(0.0572)	(0.449)
Observations	3,539	2,366	3,528	1,971
R-squared	0.322	0.319	0.343	0.308
Age Dummies	Yes	Yes	Yes	Yes
Education Dummies	Yes	Yes	Yes	Yes
Household Income Dummies	Yes	Yes	Yes	Yes
Pop. Density Dummies	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

a computer or Internet at home do not use the technology, and that non-users' skills are zero, and show the association between living with children and young adults and computer and Internet use and skills for the whole population. We distinguish between households with and without Internet by adding interaction variables between children and young adults and having Internet. Living with children is associated with lower computer and Internet skills, but not for young adults in general. For households with Internet, living with children is associated with higher levels of computer and Internet skills, as well as higher levels of Internet use. On the other hand, living with young adults is associated with lower levels of computer and Internet skills. It seems that children and young adults have different effects in adults' Internet and computer skills, even if selection plays a role. In the case selection occurs because Internet adoption is driven by the young adults, it is likely that the interaction variable between young adults and home Internet is negative in order to reflect the lower use and skills of older adults living with young adults and that

Table 4.7: Computer and Internet use and skills as a function of household composition.

VARIABLES	(1) CompUse	(2) SkillsComp	(3) InetUseHome	(4) SkillsInet
Children	-0.116***	-0.572***	-0.118***	-0.577***
	(0.0145)	(0.0946)	(0.0138)	(0.0784)
ChildrenInetHouse			0.162***	0.469***
			(0.0231)	(0.109)
ChildrenCompHouse	0.154***	0.460***		
	(0.0226)	(0.123)		
YoungAdults	-7.99e-03	-0.188*	-0.0192	-0.0900
	(0.0133)	(0.0989)	(0.0130)	(0.0932)
YoungAdultsInetHouse			-0.109***	-0.471***
			(0.0241)	(0.120)
YoungAdultsCompHouse	-0.128***	-0.316**		
	(0.0241)	(0.130)		
CompHouse	0.290***	0.584***	0.119***	0.453***
	(0.0212)	(0.0986)	(0.0172)	(0.0700)
InetHouse	0.119***	0.484***	0.252***	0.633***
	(0.0181)	(0.0855)	(0.0221)	(0.0892)
Constant	0.212***	2.004***	0.172***	1.876***
	(0.0210)	(0.113)	(0.0210)	(0.108)
Observations	7,489	7,489	7,489	7,489
R-squared	0.514	0.575	0.496	0.564
Age Dummies	Yes	Yes	Yes	Yes
Education Dummies	Yes	Yes	Yes	Yes
Household Income Dummies	Yes	Yes	Yes	Yes
Pop. Density Dummies	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

would not acquire Internet otherwise. A possible explanation for the observed positive sign in households with Internet and children, is that adults that live with children and those that have higher levels of computer and Internet skills are more likely to acquire Internet.

In summary, it is difficult to reach a conclusion about the existence of an effect of living with children or young adults on computer and Internet skills, due to potential selection both on technology adoption and its use. Therefore, we follow the strategy outlined in section 4.3, and apply a Heckman selection model.

4.4.2 Heckman Selection Model

We estimate the three-step process described in section 4.3 to account for potential selection on observed patterns in technology use and corresponding skills. Table 4.8 shows results for computer use and skills. In the first selection equation, in column (4), we include household and municipality level covariates, leaving out individual level covariates, given that this step corresponds to a household level decision. Besides the variables of interest — living with a child or young adult — we include average household education level, household income bracket dummies, and regional level variables such as average earnings in the municipality and a proxy for the quality of Internet available in the municipality: the average distance between schools in the municipality and the corresponding ISP central offices.[1] Living with children and young adults is positively associated with the likelihood of having a computer at home. Average education level in the household is also positively associated with computer ownership, and the lower the average Internet quality, the lower the likelihood of having a computer at home. This is somewhat expected considering that many computer acquisitions may be motivated by the intention of using the Internet. Column (3) represents probit estimates of the likelihood of using a computer at home as a function of living with children and young adults, and including the estimated inverse Mills ratio from the selection equation in column (4). Column (2) is identical to column (3) except that the estimate is applied only over the sub-sample that exhibits computer skills. Living with a child does not seem to affect the likelihood of using a computer, but living

[1]In 2008 the vast majority of broadband Internet services were provided using ADSL through the copper lines of the incumbent ISP, Portugal Telecom. The average distance between schools and the central offices of the ISP are a good proxy for the relative quality and attractiveness of Internet in a given region. See Belo et al. (2012c) for more details.

with a young adult seems to negatively affect computer use. Column (3) represents OLS estimates of computer skills as a function of living with children and young adults, and including the estimated inverse Mills ratio from the selection equation in column (2). Both children and young adults are associated with lower computer skills. This can be justified by either a monopolization effect, in which children and young adults use the computer more time and adults to not get to learn how to perform some tasks with the technology, or by adults relying on younger people to perform some tasks, they need, which prevents them from learning how to perform the tasks by themselves. An alternative explanation is that the selection model is not capturing the selection correctly, given that the inverse Mills ratios parameters are not statistically significant, which is an indication that there is no selection, or that the model is not well specified.

Table 4.9 has the same structure as Table 4.8 but applied to Internet adoption, use and skills, and with an extra column corresponding to the regression of hours spent on Internet as a function of living with children and young adults. Living with children or young adults increases the likelihood of having Internet at home (column (5)), but negatively impacts Internet skills (column (1)). Living with children is also associated with a lower amount of time spent online, although there is no relation with young adults. In line with computer use, Internet use is also negatively affected by young adults but not by children (columns (3) and (4)). As with computer adoption, there is no evidence of selection in Internet adoption, but there is evidence of selection in Internet use, i.e., individuals more likely to use Internet are the ones that exhibit more skills.

To test the resource sharing hypothesis we correlate the amount of time older adults

Table 4.8: Heckman Selection Models of Computer Skills and Use.

	(1)	(2)	(3)	(4)
	Type II Tobit		Binary Resp. Model	
VARIABLES	SkillsComp	CompUseHome	CompUseHome	CompHouse
Children	-0.401***	0.0687	0.0766	0.617***
	(0.139)	(0.0905)	(0.102)	(0.0835)
YoungAdults	-0.327*	-0.357***	-0.413***	1.081***
	(0.171)	(0.110)	(0.133)	(0.0937)
HouseEduc				0.998***
				(0.0947)
Dist2COkm				-0.183*
				(0.0943)
PopDensMuni				-1.95e-05
				(3.40e-05)
AvEarnk				-0.00176
				(0.0195)
InvMillsCompHouse	0.277	-0.182		
	(0.299)	(0.175)		
Constant	3.997***	0.264	0.384	-2.141***
	(0.619)	(0.363)	(0.469)	(0.329)
λ	0.052		0.108	
	(0.066)		(0.227)	
Age Dummies	Yes	Yes	Yes	No
Education Dummies	Yes	Yes	Yes	No
Household Inc. Dummies	Yes	Yes	Yes	Yes
Pop. Density Dummies	Yes	Yes	Yes	Yes
Observations	2,701	2,701	4,782	4,782

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.9: Heckman Selection Models of Internet Skills and Use.

	(1)	(2)	(3)	(4)	(5)
		Type II Tobit		Binary Resp. Model	
VARIABLES	SkillsInet	InetHours	InetUseHome	InetUseHome	InetHouse
Children	-0.515***	-0.374***	-0.0410	-0.0399	0.337***
	(0.166)	(0.0940)	(0.0870)	(0.0862)	(0.0752)
YoungAdults	-1.075***	-0.101	-0.480***	-0.453***	0.918***
	(0.220)	(0.131)	(0.118)	(0.105)	(0.0861)
HouseEduc					0.825***
					(0.0834)
Dist2COkm					-0.152*
					(0.0871)
PopDensMuni					-2.09e-05
					(3.11e-05)
AvEarnk					-8.93e-03
					(0.0185)
InvMillsInetHouse	-0.388	0.0877	-0.293		
	(0.446)	(0.241)	(0.217)		
Constant	2.384***	3.210***	0.260	0.295	-1.733***
	(0.520)	(0.432)	(0.315)	(0.395)	(0.303)
λ	1.946***	-0.237		-0.255	
	(0.191)	(0.228)		(0.198)	
Age Dummies	Yes	Yes	Yes	Yes	No
Education Dummies	Yes	Yes	Yes	Yes	No
Household Inc. Dummies	Yes	Yes	Yes	Yes	Yes
Pop. Density Dummies	Yes	Yes	Yes	Yes	Yes
Observations	2,406	2,406	2,406	5,051	5,051

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

spend online with the amount of time young adults spend online. For the 561 older adults that use Internet and live with young adults, there is a positive correlation between the amount they report spending online and the number of online time reported by the young adults (correlation of 0.15; statistical significance at the 1% level when regressing older adults' time online on young adults' time online, and 10% significance level including controls for age, education, household income and population density). This result suggests that the resource sharing hypothesis does not apply for young adults.

Table 4.10: Hours of Internet use as a function of Young Adults' Hours of Internet Use.

	(1)	(2)
VARIABLES	InetHours	InetHours
YoungInetHours	0.124**	0.0761
	(0.0600)	(0.0530)
HouseIncome2		0.0287
		(0.357)
HouseIncome3		0.290
		(0.366)
HouseIncome4		0.264
		(0.356)
HouseIncome5		0.390
		(0.361)
PopDensMed		-0.114
		(0.135)
PopDensLow		0.0847
		(0.158)
Age35to44		-0.908***
		(0.174)
Age45to54		-0.783***
		(0.160)
Age55to64		-1.024***
		(0.240)
Age65to74		-0.706***
		(0.193)
EducHigh		0.716***
		(0.162)
EducCollege		1.087***
		(0.176)
Constant	1.888***	2.105***
	(0.214)	(0.372)
Observations	509	494
R-squared	0.013	0.288

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.4.3 Partitioning the data

In order to further understand whether or not selection plays an important role in our estimates, we partition the data by age, education and income brackets and perform t-tests for living with children and young adults on computer and Internet use and skills. We also partition the population by household Internet ownership. Table 4.11 shows for computer and Internet use and skills conditional on living with children, while Table 4.12 shows t-tests conditional on living with young adults. In general we see that living with a child or young adult is associated with significantly different levels of computer and Internet use and skills, but only for people with low education. Moreover, this is significant mainly for the age brackets 25 to 34 and 35 to 44.

In Table 4.11, living with a child is generally associated with higher levels of computer and Internet use and skills for the age bracket 35 to 44, but associated lower levels of computer and Internet use and skills for the age bracket 25 to 34. This is especially significant for households with Internet. This effect is not apparent in previous estimates probably due to the mixed signs in different age brackets. In Table 4.12 the existence of a young adult at home is associated with lower levels of computer and Internet use and skills both among households with Internet and household without Internet.

In summary, both children and young adults are associated with lower levels of computer and Internet use and skills, but only for lower levels of education and for the age bracket of 25 to 34. For the age bracket 35 to 44 children seem to have a positive effect in computer and Internet use and skills while young adults keep associated with a negative effect. Thus, it is unlikely that the negative effect observed on living with young adults derives from

Table 4.11: T-tests for Computer and Internet Use and Skills by Presence of Children in the Household.

			All					InetHouse					NoInetHouse				
Age	Ed.	Inc.	N	CompUse	SkillsComp	InetUse	SkillsInet	N	CompUse	SkillsComp	InetUse	SkillsInet	N	CompUse	SkillsComp	InetUse	SkillsInet
2	1	1	34	-.0069	-.201	.0139	-.285	6	-.333	-3.33	-.333	-2	28	.0615	.487	.0923	.0974
2	1	2	296	-.0925	-.851***	-.0909	-.641***	99	-.172**	-1.79***	-.179**	-1.6***	197	-.0136	-.175	.0038	.0825
2	1	3	195	-.014	-.206	-.061	-.376	95	.111	-.257	.0273	-.634	100	-.159	-.305	-.18*	-.302
2	1	4	66	-.356***	-1.57***	-.278**	-.872*	47	-.33***	-1.28*	-.243*	-.611	19	-.262	-1.49*	-.143	-.857
2	1	5	22	-.15	-.667	-.233	-1.32	18	0	.333	0	-.556	4	-.667	-4.33	-1	-3
2	2	1	5	0	-5	0	-4.5	3	–	–	–	–	2	0	-5	0	-4
2	2	2	63	.0541	-.5	.0156	-.786	41	0	-.593	0	-.571	22	.2*	.517	.117	0
2	2	3	86	-.0102	-.925**	-.0498	-1.15**	62	-.04	-1.23**	-.08*	-1.39**	24	.0667	-.156	.0222	-.689
2	2	4	53	.0571	.243	.0857	.325	43	0	-.18	.04	-.549	10	–	–	–	–
2	2	5	39	.0455	-1.25*	.0455	-1.78***	35	0	-1.4**	0	-2.03***	4	.5	1	.5	1
2	3	1	2	0	2	0	-2	2	0	2	0	-2	0	–	–	–	–
2	3	2	33	0	1.34*	0	1.13	23	0	1.3*	0	1	10	0	1.22	0	1.22
2	3	3	51	0	-.778	0	-1.16*	45	0	-.244	0	-.487	6	0	-1	0	-1.33**
2	3	4	44	0	-.741	0	-.829	39	0	-.6	0	-.6	5	0	-2	0	-3
2	3	5	92	0	-1.04***	0	-1.28***	90	0	-1.04***	0	-1.27***	2	–	–	–	–
3	1	1	82	.169*	.162	.128	.127	21	.5**	.214	.143	.357	61	-.0183	.0075	.0409	-.109
3	1	2	487	.124***	.109	.115***	.128	208	.0118	-.477	.0651	-.152	279	.0969*	.0731	.048	-.0356
3	1	3	401	.125**	.0778	.0928*	.197	248	.133*	-.13	.103	.105	153	.0552	.18	.0128	.143
3	1	4	93	.132	.911*	.217**	.962**	75	.262**	1.47**	.31***	1.36***	18	-.481**	-1.58*	-.286**	-1*
3	1	5	43	.519***	2.3***	.479***	1.89***	36	.408**	2.09**	.431**	1.86***	7	.5*	1*	0	0
3	2	1	6	-.5	-3.75**	-1	-3.25**	2	0	-3	-1	-2	4	-1	-4.33	-1	-3.67
3	2	2	31	-.121	-.358	.075	.671	21	-.0288	.115	.173	.385	10	-.333*	-1.67	-.19	-.429
3	2	3	59	.0212	1.2**	-.0015	1.25**	44	.0706	1.49**	.0412	1.29**	15	-.1	.7	-.1	1
3	2	4	53	.0625	1.47***	.098	1.55**	44	0	.771*	0	.959	9	.167	.833	0	0
3	2	5	39	0	.0363	.0605	1.16	35	0	-.023	0	.948	4	0	-1.5	-.5	-1.5
3	3	1	0	–	–	–	–	0	–	–	–	–	0	–	–	–	–
3	3	2	14	0	-2.15**	-.25*	-2	7	0	-2.25	-.333	-2.58	7	0	-1.5	0	-.833
3	3	3	31	0	-.0792	0	-.487	25	0	-.66	0	-1.1	6	0	2.33***	0	2
3	3	4	32	0	-.208	0	-.5	30	0	-.216	0	-.33	2	–	–	–	–
3	3	5	108	-.0115	.181	-.023	.0115	104	-.0116	-.209	-.0233	-.362	4	0	1.33	0	1
4	1	1	105	.22***	.732***	.155***	.255***	17	.318	1.48**	.242	.485	88	.127*	.307***	.0833***	.0833***
4	1	2	580	-.0171	-.0625	.0101	.0616	226	-.0267	-.0719	-.0161	.07	354	-.0576*	-.156*	-.0136	-.0391
4	1	3	492	-.0343	-.119	-.0076	-.111	310	-.0116	-.0196	.0458	-.0319	182	-.0038	-.114	-.0293	-.0863
4	1	4	203	.0862	.0586	-.0172	-.0862	173	.0496	.0109	-.0195	-.111	30	.231	-.173	-.269	-.577
4	1	5	80	-.0036	.218	-.0727	-.305	71	.0027	.243	-.0652	-.286	9	-.0714	-.0714	-.286	-.857
4	2	1	5	–	–	–	–	4	–	–	–	–	1	–	–	–	–
4	2	2	18	-.123	.846	.0308	.354	13	0	1.42	0	.306	5	-.75	-1.75	-.25	-.5
4	2	3	49	.129	.403	.105	.303	39	.0583	-.0333	.075	.133	10	.375	1.38	0	.125
4	2	4	33	.25*	1.29	.306*	1.43*	30	.19	1.1	.27	1.37*	3	–	–	–	–
4	2	5	48	.0588	-1.16*	-.084	.134	46	.0606	-.883	-.0163	.45	2	0	-5	-1	-5
4	3	1	0	–	–	–	–	0	–	–	–	–	0	–	–	–	–
4	3	2	2	–	–	–	–	2	–	–	–	–	0	–	–	–	–
4	3	3	25	-.0486	-.431	-.16	-.375	21	-.143	-1.29*	-.286**	-1.14	4	.5	3.5*	.5	3.5***
4	3	4	22	0	-.271	.118	-1.73	21	0	-.3	.0625	-1.95*	1	–	–	–	–
4	3	5	88	-.0085	.31	-.0064	.541	86	-.0078	.274	-.0233	.444	2	–	–	–	–
5	1	1	205	.0601	.0692	.0856	.171	19	.146	-.271	.208	.167	186	-.0389	-.0833	-.0167	-.0111
5	1	2	654	.0598	.131	.0356	.0841	137	-.0119	-.158	-.0382	-.188	517	.0215	.147	.0158	.0937
5	1	3	358	-.097	-.319	-.0915*	-.289**	147	-.112	-.327	-.125	-.467*	211	-.103	-.36	-.0847	-.18
5	1	4	123	-.15	-.114	-.135	-.142	85	-.146	-.0426	-.15	-.22	38	-.257	-.686	-.171	-.286
5	1	5	73	-.242	-1.16*	-.229	-.719	68	-.206	-1.1	-.234	-.747	5	-.5	-1.75	0	0
5	2	1	1	–	–	–	–	0	–	–	–	–	1	–	–	–	–
5	2	2	15	.143	-.571	.5	2.36	5	.25	-.5	.25	.25	10	–	–	–	–
5	2	3	17	.188	2.75	.188	.625	11	–	–	–	–	6	.2	2.8	0	.6
5	2	4	27	-.42*	-1.02	-.34	-.46	24	-.409*	-1.18	-.364	-.727	3	–	–	–	–
5	2	5	39	.108	.162	.135	1.16	36	.0882	.0588	.118	1.21	3	–	–	–	–
5	3	1	2	–	–	–	–	1	–	–	–	–	1	–	–	–	–
5	3	2	6	-.8	-3	-.6	-2	2	–	–	–	–	4	-.667	-2	-.333	-.333
5	3	3	7	–	–	–	–	6	–	–	–	–	1	–	–	–	–
5	3	4	23	.0455	1.23	.136	-.136	20	.0526	1.16	.105	-.263	3	–	–	–	–
5	3	5	75	-.194	-1.5	-.111	-.375	72	-.188	-1.51	-.101	-.348	3	–	–	–	–
6	1	1	298	.173**	.349*	.19***	.19***	12	-.0909	-.182	-.0909	0	286	.225***	.454**	.243***	.239***
6	1	2	796	-.038	-.055	-.0183	-.0314	74	-.113	-.113	-.0968	-.161	722	-.0313	-.0499	-.0114	-.0199
6	1	3	245	-.0868	-.206	-.0734	-.17	84	-.12	-.269	-.141	-.324	161	-.0884	-.211	-.0408	-.0952
6	1	4	62	-.0536	-.324	-.0833	-.188	36	.0509	-.0291	-.04	-.12	26	-.217	-.696	-.13	-.261
6	1	5	39	-.107	.19	.031	.0276	20	-.143	-.429	-.143	-.357	19	-.0167	1.05	.25**	.5**
6	2	1	2	–	–	–	–	0	–	–	–	–	2	–	–	–	–
6	2	2	12	-.545	-1	-.182	-.364	2	–	–	–	–	10	-.444	-.667	-.111	-.333
6	2	3	11	–	–	–	–	5	–	–	–	–	6	–	–	–	–
6	2	4	6	.4	2.8	.4	1.2	3	0	1.5	0	.5	3	–	–	–	–
6	2	5	11	–	–	–	–	8	–	–	–	–	3	–	–	–	–
6	3	1	0	–	–	–	–	0	–	–	–	–	0	–	–	–	–
6	3	2	8	–	–	–	–	2	–	–	–	–	6	–	–	–	–
6	3	3	19	–	–	–	–	10	–	–	–	–	9	–	–	–	–
6	3	4	10	.125	1	-.25	-.25	3	.5	2	-.5	-.5	7	-.333	-.333	-.167	-.167
6	3	5	35	.364	.227	.515	1.11	29	.259	-.278	.407	.796	6	–	–	–	–

Table 4.12: T-tests for Computer and Internet Use and Skills by Presence of Young Adults in the Household.

			All					InetHouse					NoInetHouse				
Age	Ed.	Inc.	N	CompUse	SkillsComp	InetUse	SkillsInet	N	CompUse	SkillsComp	InetUse	SkillsInet	N	CompUse	SkillsComp	InetUse	SkillsInet
2	1	1	34	-.576	-1.24	-.394	-.939	6	-.6	-2	-.6	-1.6	28	–	–	–	–
2	1	2	296	-.252***	-.349	-.167*	-.267	99	-.394***	-1.1	-.358***	-1.12*	197	-.247**	-.244	-.151	-.181
2	1	3	195	-.218**	-.364	-.106	-.0683	95	-.22**	-.553	-.129	-.14	100	-.215*	-.182	-.0825	.0026
2	1	4	66	-.396***	-1.62***	-.314**	-.898*	47	-.466***	-2.09***	-.401***	-1.27**	19	-.0778	.278	.111	.667
2	1	5	22	.15	.667	.233	.95	18	0	-.333	0	.111	4	.667	4.33	1	3
2	2	1	5	–	–	–	–	3	–	–	–	–	2	–	–	–	–
2	2	2	63	.0377	.721	.0566	.951	41	0	-.471	0	-.626	22	.105	3.07***	.158	4.05***
2	2	3	86	-.0575	-.698	-.101	-.103	62	-.0769**	-.965	-.154***	-1.01	24	.0435	.826	.13	3.26
2	2	4	53	.0444	-.694	.0667	-.467	43	0	-1.14	.0263	.0789	10	.286	1.14	.286	.333
2	2	5	39	-.0909	1.33*	-.0909	1.08	35	0	1.68***	0	1.5**	4	-1	-2.67	-1	-3.33
2	3	1	2	–	–	–	–	2	–	–	–	–	0	–	–	–	–
2	3	2	33	0	-.907	0	-.722	23	0	.533	0	-.533	10	0	-2.24**	0	-.333
2	3	3	51	0	-.208	0	1.58	45	0	-.405	0	1.33	6	–	–	–	–
2	3	4	44	0	1.61***	0	2.32***	39	0	1.61***	0	2.22***	5	–	–	–	–
2	3	5	92	0	.227	0	.634	90	0	.23	0	.661	2	–	–	–	–
3	1	1	82	-.0446	-.344	-.14	-.445	21	-.4*	-1.37	-.5**	-1.3	61	-.0053	.0212	-.0926	-.259
3	1	2	487	-.15***	-.401**	-.135***	-.411***	208	-.23***	-.98***	-.301***	-1.15***	279	-.201***	-.316**	-.108**	-.126
3	1	3	401	-.13***	-.574***	-.127***	-.505***	248	-.266***	-1.27***	-.298***	-1.06***	153	-.0918	.0116	-.0484	-.151
3	1	4	93	-.192*	-1.64***	-.201*	-1.62***	75	-.244**	-1.81***	-.211**	-1.87***	18	.013	-.987	-.182	-.636
3	1	5	43	-.417***	-1.69**	-.376**	-1.37**	36	-.476***	-1.8**	-.356**	-1.36*	7	.25	.5	0	0
3	2	1	6	-.5	-1.5	-.25	-1.75	2	0	3	1	2	4	-1	-4.33	-1	-3.67
3	2	2	31	-.0846	-.892	-.0462	1.18	21	-.075	-1.01	-.075	.475	10	–	–	–	–
3	2	3	59	.0714	-.34	.0952	-.471	44	.0714	-.295	.107	-.688	15	.0714	-.0714	.0714	-.5
3	2	4	53	.0208	-1.04	.0625	-1.14	44	0	-1.57***	0	-1.74**	9	–	–	–	–
3	2	5	39	0	-.922	-.0444	-1.34*	35	0	-1.11	0	-1.25*	4	0	1.5	.5	1.5
3	3	1	0	–	–	–	–	0	–	–	–	–	0	–	–	–	–
3	3	2	14	–	–	–	–	7	–	–	–	–	7	–	–	–	–
3	3	3	31	0	-1.31	0	-.586	25	0	-1.35	0	-.739	6	–	–	–	–
3	3	4	32	0	.0333	0	.533	30	0	.0357	0	.393	2	–	–	–	–
3	3	5	108	.0102	-.265	.0204	-.929	104	.0106	-.362	.0213	-1.02	4	–	–	–	–
4	1	1	105	-.0333	-.122	.0333	.0778	17	-.319	-1.1	-.0972	-.194	88	-.0617	-.0741	-.0123	-.0123
4	1	2	580	.0288	.0119	.0048	-.0162	226	-.0378	-.136	-.089	-.245*	354	-.0325	-.116	-.0308	-.0727
4	1	3	492	.0127	.161	.0176	.0836	310	-.0896	-.0211	-.0705	-.136	182	-.0026	.0219	-.0054	.0401
4	1	4	203	.0135	.211	.0436	.038	173	.0528	.316	.066	.121	30	-.295*	-.616	-.17	-.67*
4	1	5	80	.0757	.693	.155	.513	71	.161	.802	.149	.511	9	-.55*	-.1	.05	.15
4	2	1	5	0	-.75	.25	0	4	0	-1.33	.333	-.333	1	–	–	–	–
4	2	2	18	.154	.569	.308	.631	13	0	.05	0	-.225	5	–	–	–	–
4	2	3	49	-.0966	-.416	-.178	-.505	39	-.119	-.905	-.286**	-1.02*	10	-.25	.125	0	.75
4	2	4	33	.156	-.0111	0	-.767	30	.172	.0045	.0136	-.751	3	-.5	-2	-.5	-2
4	2	5	48	-.1*	-.429	-.114	-.486	46	-.1*	-.577	-.15**	-.565	2	–	–	–	–
4	3	1	0	–	–	–	–	0	–	–	–	–	0	–	–	–	–
4	3	2	2	–	–	–	–	2	–	–	–	–	0	–	–	–	–
4	3	3	25	-.0873	-.405	-.23	-1.13	21	-.167*	-1	-.333***	-1.63*	4	.333	3	.333	1.67
4	3	4	22	0	-.533	.0167	.0667	21	0	-.611	-.0833	-.278	1	–	–	–	–
4	3	5	88	-.05	-.25	-.0375	-.183	86	-.05	-.299	-.0565	-.301	2	–	–	–	–
5	1	1	205	.189***	.983***	.215***	.512***	19	.257	1.5	.329*	.914	186	.0913	.438**	.114**	.119***
5	1	2	654	-.0419	-.082	-.0371	-.105	137	-.174**	-.356	-.155**	-.475**	517	-.05	-.134	-.037	-.0638
5	1	3	358	.0135	.131	.0265	.122	147	.0132	.135	.0041	.027	211	-.0652	-.048	-.0406	.0441
5	1	4	123	-.0576	-.611*	-.134	-.438	85	-.0683	-.699*	-.145	-.598*	38	-.0833	-.552	-.188	-.313
5	1	5	73	-.0362	.274	-.181	-.66	68	-.0433	.279	-.216	-.76*	5	–	–	–	–
5	2	1	1	–	–	–	–	0	–	–	–	–	1	–	–	–	–
5	2	2	15	.143	-1.64	.5	1.29	5	.25	-1.75	.25	-1	10	–	–	–	–
5	2	3	17	-.19	.31	-.19	.714	11	-.208	.542	-.0833	.875	6	–	–	–	–
5	2	4	27	-.25	-1.63	-.167	-1.79	24	-.238	-1.81	-.19	-2.1*	3	–	–	–	–
5	2	5	39	-.419*	-1.95	-.392	-.419	36	-.441**	-2.06	-.412*	-.382	3	–	–	–	–
5	3	1	2	–	–	–	–	1	–	–	–	–	1	–	–	–	–
5	3	2	6	–	–	–	–	2	–	–	–	–	4	–	–	–	–
5	3	3	7	.167	.5	.333	-.5	6	.2	.4	.4	-.6	1	–	–	–	–
5	3	4	23	.0556	.222	.167	.856	20	.0625	-.5	.125	0	3	0	4	.5	5.5
5	3	5	75	.0275	.871	.129	.257	72	.0357	.875	.143	.295	3	–	–	–	–
6	1	1	298	-.0308	-.0582	-.0137	-.0137	12	-.143	-.286	-.143	0	286	-.0281	-.0526	-.0105	-.014
6	1	2	796	-.0383	-.0554	-.0185	-.0317	74	-.13*	-.13	-.111	-.185	722	-.0313	-.0497	-.0114	-.0199
6	1	3	245	-.0236	-.116	-.0392	-.0659	84	-.0375	-.172	-.0906	-.163	161	-.0872	-.208	-.0403	-.094
6	1	4	62	-.14	-.4	-.08	-.18	36	-.0769	-.154	-.0385	-.115	26	-.208	-.667	-.125	-.25
6	1	5	39	-.206	-.412	-.0882	-.206	20	-.111	-.333	-.111	-.278	19	-.313	-.5	-.0625	-.125
6	2	1	2	–	–	–	–	0	–	–	–	–	2	–	–	–	–
6	2	2	12	–	–	–	–	2	–	–	–	–	10	–	–	–	–
6	2	3	11	–	–	–	–	5	–	–	–	–	6	–	–	–	–
6	2	4	6	.4	-.8	.4	0	3	0	-3	0	-1	3	–	–	–	–
6	2	5	11	–	–	–	–	8	–	–	–	–	3	–	–	–	–
6	3	1	0	–	–	–	–	0	–	–	–	–	0	–	–	–	–
6	3	2	8	–	–	–	–	2	–	–	–	–	6	–	–	–	–
6	3	3	19	–	–	–	–	10	–	–	–	–	9	–	–	–	–
6	3	4	10	–	–	–	–	3	–	–	–	–	7	–	–	–	–
6	3	5	35	.364	2.35	-.0152	.576	29	.259	1.87	-.13	.259	6	–	–	–	–

selection, which strengthens the non-selection hypotheses.

4.5 Conclusion

We look at the effects of living with young people on computer and Internet use patterns and skill acquisition. Living with children and young adults is associated with higher levels of computer and Internet adoption at the household level, but not associated with higher computer or Internet use and skills. Older adults that live with young people exhibit lower levels of computer and Internet skills. We have explored several alternative explanations for the negative association between young people and use patterns and skills, and were able to rule out the resource sharing hypothesis as a potential explanation.

Due to the cross-section nature of our data, we can not completely rule out selection as a potential explanation for the negative association between computer and Internet skills and living with young people, even when using the selection model to estimate the parameters of interest.

Chapter 5

Conclusion

ICTs are shaping the ways people interact with each other and are already an essential part of our lives. The introduction of ICTs in areas such as education creates opportunities for improvement but can also cause disruption and generate unforeseen or unintended effects. While there are merits in providing broadband to schools — enabling students and teachers to access information — the studies presented here show that unintended consequences can occur.

Using a comprehensive dataset on school broadband use, I find evidence that broadband does not improve student performance, and that in some cases grades decrease. The analysis shows that on average broadband is responsible for a decline of 0.97 of a standard deviation in grades in the 2005-2009 window. Additionally I have looked at the impacts the deployment of broadband in schools may have at the community level. I find that school broadband use contributes directly to a higher adoption rate in households with children. In 2008 and 2009 school Internet use increased the probability of adopting Internet by 20%

in households with children, which represents an increase of 5% in the total population. I have found no evidence of a statistically significant effect at the neighborhood level. Finally I have looked at how household composition shapes computer and Internet use patterns and skills acquisition. Living with children and young adults is associated with higher levels of computer and Internet adoption at the household level, but not associated with higher computer or Internet use and skills. Additionally, older adults that live with young people exhibit lower levels of computer and Internet skills.

The main question remaining to be answered is how to minimize unintended consequences and how to take full advantage of the available technology. I identify some potential directions that might help to improve the outcomes of deploying ICT technology in large-scale. First, mere broadband provisioning is not enough, and schools must be aware of potential pitfalls and must be prepared to absorb the technology so that they can benefit from it without compromising students' performance. In line with this, such large-scale programs could benefit from being implemented in phases, in order to detect flaws and correct errors before all schools get the technology. Although Internet at school increases household Internet adoption for households with children, adoption does not necessarily improve adults' Internet use patterns and skills. Young people seem to be the main beneficiaries of Internet at home, despite the possibility that older people may also benefit indirectly from the technology.

Bibliography

ANACOM. Histórico dos elementos do serviço de acesso à internet. Technical report, ANACOM, 2010. 2.4

M. Angelucci and V. Di Maro. Program evaluation and spillover effects. *SPD Working Papers*, 2010. 2

J. Angrist and V. Lavy. New evidence on classroom computers and pupil learning. *Economic Journal*, pages 735–765, 2002. 1, 2.2, 2.7.4

J.D. Angrist and V. Lavy. Using maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 114(2):533–575, 1999. 2.2, 2

Abhijit V. Banerjee, Shawn Cole, Esther Duflo, and Leigh Linden. Remedying education: Evidence from two randomized experiments in india. *Quarterly Journal of Economics*, 122(3):1235–1264, 2007. doi: 10.1162/qjec.122.3.1235. 2.2, 2.7.4

Felipe Barrera-Osorio and Leigh L. Linden. The Use and Misuse of Computers in Education : Evidence from a Randomized Experiment in Colombia. *SSRN eLibrary*, 2009. 2.2

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

52–74, 2009. 2.2

F.M. Bass. A New Product Growth for Model Consumer Durables. *Management Science*, 15(5):215, 1969. 3.1, 4.2

R. Belo, P. Ferreira, and R. Telang. The Effects of Broadband in Schools: Evidence from Portugal. Working Paper, 2010. 3.1, 3.3

Rodrigo Belo, Pedro Ferreira, and Rahul Telang. Spillover Effects of Broadband in Schools and the Critical Role of Children. Working Paper, 2012a. 4.2

Rodrigo Belo, Pedro Ferreira, and Rahul Telang. Spillover Effects of Broadband in Schools and the Critical Role of Children. In *Academy of Management Annual Meeting (forthcoming)*. Boston, Massachusetts, August 3-7 2012b. 4.1

Rodrigo Belo, Pedro Ferreira, and Rahul Telang. The Effects of Broadband in Schools On Students' Performance. Submitted to Management Science (Second Round Review), 2012c. 1, 1

Erik Brynjolfsson and Lorin Hitt. Paradox lost? firm-level evidence on the returns to information systems spending. *Management Science*, 42(4):541–558, 1996. ISSN 00251909. URL `http://www.jstor.org/stable/2634387`. 2.1

Erik Brynjolfsson and Lorin M. Hitt. Computing productivity: Firm-level evidence. *The Review of Economics and Statistics*, 85(4):pp. 793–808, 2003. ISSN 00346535. 3.1, 4

P. Carrillo and J. Ponce. Information Technology and Student's Achievement: Evidence from a Randomized Experiment in Ecuador. *RES Working Papers*, 2011. 2.7.4

Menzie D. Chinn and Robert W. Fairlie. Ict use in the developing world: An analysis of differences in computer and internet penetration. Working Paper 12382, National Bureau of Economic Research, July 2006. 3.1, 5

L. Cuban and H. Kirkpatrick. Computers Make Kids Smarter–Right?. *Technos*, 7(2): 26–31, 1998. 2.2

F.D. Davis. *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. PhD thesis, Massachusetts Institute of Technology, Sloan School of Management, 1985. 4.2

P. DiMaggio, E. Hargittai, et al. From the 'digital divide'to 'digital inequality': Studying internet use as penetration increases. *Princeton: Center for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University*, 2001. 4.1, 4.2

FCT. Fct - internet na escola, 2001. URL `http://www.fct.mctes.pt/programas/interescola.htm`. (in Portuguese). 2.3.1

C. Forman, A. Goldfarb, and S. Greenstein. The internet and local wages: A puzzle. *American Economic Review (forthcoming)*, 2012. 4

Chris Forman, Avi Goldfarb, and Shane Greenstein. How did location affect the adoption of the commercial internet? global village vs. urban density. *Journal of Urban Economics*, 58(3):389–420, 2005. 2.1

V. Glass and S.K. Stefanova. An empirical study of broadband diffusion in rural america. *Journal of Regulatory Economics*, pages 1–16, 2010. 3.1

A. Goldfarb. The (teaching) role of universities in the diffusion of the internet. *International Journal of Industrial Organization*, 24(2):203–225, 2006. 3.2

A. Goolsbee and J. Guryan. The impact of Internet subsidies in public schools. *The Review of Economics and Statistics*, 88(2):336–347, 2006. 1, 2.2, 3.1, 3.2

A. Goolsbee and P.J. Klenow. Evidence on Learning and Network Externalities in the Diffusion of Home Computers*. *The Journal of Law and Economics*, 45(2):317–343, 2002. 3.1, 3.2, 3.5, 3.6.1

Austan Goolsbee. The value of broadband and the deadweight loss of taxing new technology. *NBER Working Paper*, Jan 2006. 3.1

S. Greenstein and R.C. McDevitt. The broadband bonus: Accounting for broadband internet's impact on US GDP, 2009. 3.1, 4

Z. Griliches. Hybrid corn: An exploration in the economics of technological change. *Econometrica, Journal of the Econometric Society*, pages 501–522, 1957. 3.1

E. Hargittai. Second-level digital divide. *First monday*, 7(4-1), 2002. 4.1, 4.2

E. Hargittai. Whose space? differences among users and non-users of social network sites. *Journal of Computer-Mediated Communication*, 13(1):276–297, 2007. 4.2

E. Hargittai. Digital Na (t) ives? Variation in Internet Skills and Uses among Members of the Net Generation. *Sociological Inquiry*, 80(1):92–113, 2010. 4.1, 4.2

Eszter Hargittai and Amanda Hinnant. Digital inequality. *Communication Research*, 35 (5):602–621, 2008. doi: 10.1177/0093650208321782. URL `http://crx.sagepub.com/content/35/5/602.abstract`. 4.2

J.A. Hauge and J.E. Prieger. Demand-Side Programs to Stimulate Adoption of Broadband: What Works? *Review of Network Economics*, 9(3):4, 2010. ISSN 1446-9022. 1, 3.1, 3.2

H.E. Hudson. Universal access: What have we learned from the e-rate? *Telecommunications Policy*, 28(3-4):309–321, 2004. 1

H.E. Hudson. The future of the e-rate: Us universal service fund support for public access and social services. *and Communications for All: An Agenda for a New Administration. Lanham, MD: Lexington Books*, 2009. 1

Adam B. Jaffe, Manuel Trajtenberg, and Rebecca Henderson. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598, 1993. ISSN 00335533. URL `http://www.jstor.org/stable/2118401`. 3.2

K.P. Jayakar. Reforming the e-rate. *info*, 6(1):37–51, 2004. 1

A.B. Krueger. Experimental Estimates of Education Production Functions*. *Quarterly Journal of Economics*, 114(2):497–532, 1999. 2.2, 2

V. Lavy. Using dual natural quasi-experimental designs to evaluate the effect of school hours and class size on student achievement. *Department of Economics, Hebrew University, Jerusalem*, 1999. 2.2, 2

S. Lee and J.S. Brown. Examining broadband adoption factors: an empirical analysis between countries. *info*, 10(1):25–39, 2008. 3.1

E. Leuven, M. Lindahl, H. Oosterbeek, and D. Webbink. The effect of extra funding for disadvantaged pupils on achievement. *The Review of Economics and Statistics*, 89(4): 721–736, 2007. 1, 2.2

S. Machin, S. McNally, and O. Silva. New Technology in Schools: Is There a Payoff?*.

The Economic Journal, 117(522):1145–1167, 2007. 2.2

Ofer Malamud and Cristian Pop-Eleches. Home computer use and the development of human capital. *The Quarterly Journal of Economics*, 126(2):987–1027, 2011. doi: 10.1093/qje/qjr008. 1, 2.1, 2.2, 2.7.4

S. Monjon and P. Waelbroeck. Assessing spillovers from universities to firms: evidence from french firm-level data. *International Journal of Industrial Organization*, 21(9):1255–1270, 2003. 3.2

J.E. Prieger and W.M. Hu. The broadband digital divide and the nexus of race, competition, and quality. *Information economics and Policy*, 20(2):150–167, 2008. ISSN 0167-6245. 5

Everett M. Rogers. *Diffusion of Innovations*. The Free Press, New York, 1962. 3.1

Everett M. Rogers. *Diffusion of Innovations, 5th Edition*. Free Press, August 2003. ISBN 0743222091. 4.2

G.L. Rosston, S.J. Savage, and D.M. Waldman. Household demand for broadband internet in 2010. *The BE Journal of Economic Analysis & Policy*, 10(1):79, 2010. 3.1

C.E. Rouse and A.B. Krueger. Putting computerized instruction to the test: a randomized evaluation of a "scientifically based" reading program. *Economics of Education Review*, 23(4):323–338, 2004. 2.2, 2.7.4

B. Sacerdote. Peer Effects with Random Assignment: Results for Dartmouth Roommates*. *Quarterly Journal of Economics*, 116(2):681–704, 2001. 2.2, 2

N. Selwyn, S. Gorard, and J. Furlong. Whose internet is it anyway? exploring adults'(non) use of the internet in everyday life. *European Journal of Communication*, 20(1):5–26, 2005. 4.2

Thomas D. Snyder and Sally A. Dillow. Digest of education statistics, 2010. Technical report, National Center for Education Statistics, 2011. 1

Douglas Staiger and James H. Stock. Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586, 1997. ISSN 00129682. URL `http://www.jstor.org/stable/2171753`. 2.B

J.H. Stock, M. Yogo, and L. Center. Testing for weak instruments in linear IV regression. *NBER Working Paper*, 2002. 2.B

A.S. Tanenbaum. *Computer networks*. Prentice Hall, 2002. 2.A

UMIC. Todas as escolas públicas de portugal acedem à internet em banda larga, August 2007. URL `http://www.umic.pt/index.php?option=com_content&task=view&id=2595&Itemid=86`. (in Portuguese). 3.4

J. Underwood, A. Ault, P. Banyard, K. Bird, G. Dillon, M. Hayes, I. Selwood, B. Somekh, and P. Twining. The impact of broadband in schools. *Nottingham Trent University/Becta*, 2005. 2.1

Viswanath Venkatesh and Susan A. Brown. A longitudinal investigation of personal computers in homes: Adoption determinants and emerging challenges. *MIS Quarterly*, 25 (1):pp. 71–102, 2001. ISSN 02767783. 4.2

Jacob L. Vigdor and Helen F. Ladd. Scaling the Digital Divide: Home Computer Technology and Student Achievement. *SSRN eLibrary*, 2010. 2.2

Michael R. Ward. Learning to Surf: Spillovers in the Adoption of the Internet. *SSRN eLibrary*, 2010. 3.2, 4.1

H.D. Webbink. Causal effects in education. *Journal of Economic Surveys*, 19(4):535–560, 2005. 1, 2.2

B.E. Whitacre. Factors influencing the temporal diffusion of broadband adoption: evidence from Oklahoma. *The Annals of Regional Science*, 42(3):661–679, 2008. ISSN 0570-1864. 3.1, 5

Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, 2002. 4.3