

# High-speed broadband and academic achievement in teenagers: Evidence from Sweden

Erik Grenestam \*      Martin Nordin †

April 21, 2018

## Abstract

This study examines the effects of super-fast internet connections on the academic achievement of students in upper secondary school. We link detailed register data on around 250,000 students to local levels of access to optic fiber broadband and estimate the effect of broadband on student GPA. We show that reaching full coverage in the student's parish of residence causes a GPA reduction ranging from 3 to 6 percent of a standard deviation. Estimates are consistently more negative for boys and students with low ability and/or low-educated parents. Using PISA survey data, we provide evidence that students living in areas with faster broadband expansion also saw a greater increase in the number of hours spent online during weekdays, suggesting student time use as a mechanism.

*JEL classifications:* J24, H52, I24, I28, O33

*Keywords:* Education, Broadband, Internet, High-school, GPA

---

\*Department of Economics, Lund University. PO Box 730, SE-220 07, Lund, Sweden. E-mail: erik.grenestam@nek.lu.se

†AgriFood Economics Centre and Department of Economics, Lund University. PO Box 730, SE-220 07, Lund, Sweden. The authors would like to thank Kaveh Majlesi, Petter Lundborg, Alessandro Martinello and Jan Bietenbeck for their feedback as well as seminar participants at Lund University, Agrifood Economics Centre and EALE 2017 in S:t Gallen, Switzerland.

# 1 Introduction

In countries worldwide, governments are committing to large scale investments in high-speed broadband infrastructure (OECD, 2011). Both the European Commission and the U.S. Federal Communications Commission (FCC) have set ambitious targets stating that by 2020, half of all households should have access to at least a 100 Mbit/s connection, a speed which can only be achieved via new information and communications technology (ICT) infrastructure based on optic fiber.

While policymakers often claim large benefits of upgrading networks to give more people access to a fast broadband connection (Kenny and Kenny, 2011), the effects of widespread broadband adoption are poorly understood and require further research. As many countries are currently transitioning to next-generation broadband networks (see e.g., OECD Broadband Portal for cross-country access statistics), understanding the costs and benefits of widespread high-speed broadband adoption is crucial.

This study is part of a growing body of literature on the socio-economic consequences of broadband. Several empirical studies have employed differences in the timing and location of broadband roll-out to estimate causal effects. To our knowledge, ours is the first empirical study of the effect of high-speed broadband access at home on educational outcomes. We examine the effect of high-speed broadband at home on the academic performance of upper secondary school students in Sweden. We use the differing roll-out of high-speed broadband via optic fiber across localities (250 m by 250 m grid squares), and the resulting within-student variation in broadband exposure, in order to identify the causal effect of broadband on upper secondary school grades. Our register data comprise detailed records on school achievements for all upper secondary school students in Sweden up until 2012. To measure local levels of broadband access, we use an annual nationwide survey conducted among Swedish internet service providers (ISPs) by the Swedish Post and Telecom Authority (PTS), providing high-resolution spatial data on locally available access techniques which we aggregate to the parish level. Our results indicate a small but robust negative effect of high-speed broadband on GPA. Our estimates demonstrate an overall effect size of 3-6 percent of a standard deviation, with boys, low-ability students and children to parents with low education suffering the largest negative effects. We also find evidence that students spend more time online following an upgrade to household access speed, which is a likely mechanism.

So far, there have been few studies examining the effect of broadband on educational outcomes. Among the best is the rigorous work by Faber et al. (2015), exploiting the fact that distance to the nearest telephone exchange determines the speed of copper-based DSL (digital subscriber line) broadband. They use a regression discontinuity design (RDD), with residential distance to the nearest exchange as the running variable and the as-good-as-random geographical borders between telephone exchange catchment areas as thresholds. They report a zero effect on standardized test scores, but also show that crossing a boundary seems to have a limited effect on access speed, meaning that the marginal speed increase could be too small to impact behavior. A related area of research is how computer use affects student achievement. Vigdor et al. (2014) analyze the effect of home computer use on American students' standardized test scores in school grades 5-8. Using a student fixed effects model similar to that employed in the present study, they show that a home computer has a small, but significant, negative

effect on math and reading scores. Using the number of ISPs connected to a local node as a proxy for broadband coverage, they also show that broadband access reduces homework effort and seems to widen racial and socio-economic gaps.

Using an RDD to examine the effects of a home computer voucher program directed at low-income households in Romania, Malamud and Pop-Eleches (2011) demonstrate that home computer use decreases children's grades, but increases their computer skills. In that study, children reported using the computers not for educational purposes, but rather to play games. However, the share of households with access to the internet in the study was low and the program did not seem to have any effects on internet use. Broadband expansion has also been associated with decreased teen fertility (Guldi and Herbst, 2016), with decreased sexual activity suggested as a mechanism. Aguiar et al. (2017) show that online activities can crowd out offline leisure and labor supply. Using a structural approach, they show that improved leisure technology (e.g., broadband) causes young men to reduce hours worked in favor of gaming and other recreational computer use. This shift can explain 40-80 per cent of the decline in hours worked compared with older men since 2004. Turning to labor demand, there is evidence that ICT complements the skills acquired by a formal education (Akerman et al., 2015). In this paper we study the effects on grades. However, it is important to recognize that effects of ICT may go beyond conventional measures of human capital formation and that certain cognitive skills may benefit from ICT use.

The remainder of the paper is organized as follows. Section 2 provides background on internet access technologies, the role of access speed, and underlying concepts and potential mechanisms through which access speed could affect educational outcomes. Sections 3 and 4 describe our data and empirical specification, respectively. Section 5 presents the results, section 6 reports effects on time use, and some conclusions are presented in section 7.

## 2 Background

Broadly speaking, there have been three generations of internet access technologies. With the advent of the internet and home computers came the dial-up modem. Using existing telephone lines and dial-up modems, internet access came at a low cost to consumers and ISPs. Applications such as e-mail, online chats, and browsing became common.

The second generation – what became known as 'broadband' – also entered homes over existing infrastructure, specifically phone lines (for DSL) or copper cables used for cable television. With new modems and upgrades to operator node networks, access speeds increased by orders of magnitude compared with dial-up. Access to DSL or cable broadband has now become commonplace in many developed countries (the OECD average is currently around 25 DSL and cable broadband subscriptions per 100 inhabitants). However, the infrastructure underpinning the first- and second-generation access has many limitations, stemming from the fact that the copper cables are not suited to carry high-frequency signals, limiting access speed and reliability. For example, DSL speeds quickly decrease with the distance between the consumer and the operator node. In the latest generation of broadband, copper cable is fully or partly replaced with optic fiber, enabling further increases in access speed. This study focuses

on the rollout of broadband delivered via optic fiber directly to consumers' homes, a system also known as 'fiber to the home' (FTTH).

A distinguishing feature of our study compared with previous studies on the socio-economic consequences of broadband is the margin of access speed. Most other studies have focused on the effects of going from a dial-up modem to a DSL or cable connection. By 2007, the first year for which we have data on local availability of broadband, the share of the Swedish population with DSL access is reported at 97.8 percent (Swedish Post and Telecom Authority, 2008)<sup>1</sup>. Thus, our baseline household has access to a conventional copper-based broadband connection and at the margin, the vast majority of households in our study are transitioning from broadband to faster broadband. It is important to note that, in parallel to the expansion of FTTH, the DSL and cable connections are also upgraded by replacement of copper cable with fiber, as the distance that the signal has to travel over old telephone lines or coaxial cables is reduced. Therefore, our measure of fiber coverage also captures upgrades to older technologies, in addition to FTTH. What are the effects of going from a copper-based connection to fiber? Or, what difference does a super-fast internet connection make when the user already has a fast connection? To quantify the speed increase associated with a fiber upgrade, we use data from a Swedish NGO that provides an online tool for measuring consumer access speeds (Stiftelsen för internetinfrastruktur, 2013). Using 96 million measurements taken between 2008 and 2013, it puts the average speed of DSL at 11 megabits per second (Mbps), compared with the 57 Mbps of a FTTH connection. It also reports a 50 percent reduction in average latency times when going from DSL to fiber. Surprisingly, the observed fivefold speed increase is similar to what consumers experienced when upgrading from a dial-up model to a first-generation DSL connection (the 'extensive' margin), which was typically equivalent to transitioning from 50 to 250 kbps.

How does an increase in access speed affect consumer usage? A simple comparison of internet use between consumers on slow and fast connections will be biased, due to self-selection into different plans and technologies. Grover et al. (2016) report a field experiment to estimate the causal effect of access speed on internet use. Working with a large American ISP, those authors randomly upgraded households currently on a 100 Mbps plan to 250 Mbps, without informing the households. Despite already having access to a fast connection, they found that data volumes of these households increased relative to the control group. Interestingly, households that were not fully utilizing the available capacity prior to the upgrade experienced the largest relative increase in demand, suggesting that consumers either started to use more bandwidth-intensive services and/or increased the time spent on internet use.

From a technical perspective, the marginal cost of additional bandwidth can be assumed to be insignificant. Due to higher margins on high-bandwidth plans, ISPs have an interest in convincing the consumer to upgrade their connection. With this in mind, it is interesting to examine their main selling

---

<sup>1</sup>This number reflects households living at an address where at least one ISP reported able to supply a DSL connection "without incurring a significant cost". Actual take-up was reported at around 70 percent in a government survey (Statistics Sweden, 2007).

points. Google (2016) states that fiber means “less time buffering videos [and more] online gaming”. Swedish ISPs present similar arguments. Fiber is described as improving online video streaming (often presenting scenarios where multiple family members are watching different video streams at once). The online video-streaming service Netflix recommends a 25 Mbit/s connection to stream high-definition content, well above the average Swedish DSL connection. Online gaming and downloading large files (the reference to online piracy is never explicit, for understandable reasons) are also big selling points. The aforementioned activities are all bandwidth-intensive in the sense that they benefit from increases in bandwidth and/or decreases in latency times. As an example, downloading a 3-gigabyte movie takes about 40 minutes using an average DSL connection, while with a 100 Mbps fiber connection the time is reduced to 4 minutes. To some, this reduction may be of little importance. However, a decrease in the time between deciding to watch a movie and pressing play may influence the consumption decision. Another factor relevant for bandwidth constraints is household size. With multiple people sharing available bandwidth, the marginal effect of fiber should increase, assuming diminishing returns to bandwidth.

An annual survey of online habits in Sweden puts the share of daily internet users among 15- to 19-year-olds at above 90 percent (IIS, 2015). Sixty percent of this age group report daily use of the internet for games and/or movies, but there seem to be significant differences between boys and girls. The share of boys and girls aged 16-25 who report playing games online on a daily basis is 41 and 11 percent, respectively. Another survey puts the share of boys and girls aged 15-18 who play games for more than three hours per day at 40 and 5 percent, respectively (Swedish Media Council, 2015). In 2011, researchers working with the infamous file-sharing site ‘The Pirate Bay’ conducted a survey among visitors to the site (Svensson et al., 2013). Of the 2000 respondents based in Northern Europe aged 17 or younger, about 4 percent were girls. Among daily file sharers, girls made up just over 2 percent of the sample. However, adolescent girls are avid users of other online services, e.g., they are over-represented when it comes to music streaming and social media use (Swedish Media Council, 2015). However, these two activities do not require a lot of bandwidth and consequently do not benefit from faster access speed to the same extent as the services where boys make up the majority of users. Consequently, there is reason to believe that any negative effect of broadband on academic achievement should be greater in magnitude for boys than for girls.

To clarify our hypotheses, we draw on the conceptual framework presented by Vigdor et al. (2014). The adolescent faces a resource allocation problem, where time and money can be used in activities that promote future wellbeing, i.e., investment, or activities that provide instantaneous entertainment (consumption). Technology alters the cost of both types of activities, but it is unlikely that technology will have the same effect on the marginal cost of all activities. Assuming diminishing marginal benefits, a new technology will change resource allocation in a way that favors activities whose marginal costs have decreased most. If going from fast to very fast broadband primarily lowers the cost of leisure activities (Aguilar et al., 2017) for adolescents, fiber thus contributes to lower academic effort such as time devoted to homework (referred to as “The Facebook Effect” by Faber et al. (2015)). Conversely, if fiber promotes academic productivity, the net effect of fiber could be positive. Our reduced form analysis is intended to provide insights into the net effect of fiber.

Within the family, the scope for mitigation of negative effect and reinforcement of positive effects is wide, be it through active supervision, inherited traits and abilities (both cognitive and non-cognitive), or resource endowments. To separate households along these dimensions, we use parental education as a proxy for the family environment<sup>2</sup> when examining heterogeneity in the effect of fiber.

### 3 Data

Since 2007, PTS conducts an annual survey of telecom operators and ISPs regarding broadband access in Sweden (Swedish Post and Telecom Authority, 2008). This information is matched to register data on location, transcripts from upper secondary school, and other background information. The respondents in the PTS survey are asked to produce a list of all the addresses where they supply a connection to consumers. As responding to the survey is considered mandatory, more than 90 percent of telecom operators complete the survey every year. A single grid square is considered to be covered if at least one building within the square has access to fiber, i.e., each 250 m by 250 m grid is either covered or not. Note that a consumer located in a covered square could face a significant cost of actual take-up should they live far away from the ISP node.

Using data on the working-age (16-64) population within each 250 m by 250 m square<sup>3</sup>, we calculate a measure of coverage at parish level by weighting the coverage in each square by its population (see Figures 1 and 2 for a visual representation). Our measure of fiber coverage can thus be interpreted as the share of the parish population covered or, equivalently, as the probability of being covered conditional on the parish of residence. A small share of students (about 1 percent of the sample) live in parishes where the recorded change in coverage between 2007 and 2011 is negative. In one case, we confirmed that this is due to a reporting error by the ISP. Consequently, we excluded these students from our sample. However, all results are robust to including parishes where coverage reportedly decreased during the period. The 1376 parishes in our data are on average home to about 7000 people and cover an area of about 300 km<sup>2</sup>.

Figure 3 shows the variation in our sample in terms of treatment intensity (change in fiber coverage between 2007 and 2011) and the initial level of coverage (in 2007). To reduce visual clutter, we bin our data into equal-size bins and plot within-bin means. As can be seen, almost everyone in our sample is treated to some extent, but there is much variation in treatment intensity even for a given level of initial coverage, which gives credibility to our identification strategy.

From register data, we have detailed transcripts on all students graduating upper secondary education up until 2012. Our data include GPA at graduation and grades for all individual courses. An important consideration for our study is that students in Swedish upper secondary school are graded

---

<sup>2</sup>Education is known to be correlated with several socio-economic characteristics of the family. See Björklund and Salvanes (2011) for a review of the literature on education and family background.

<sup>3</sup>We use 2013 data on grid population, as this is the earliest year available to us. Our results are robust to using total population instead.

continuously throughout their three years of study. Subjects are typically taught over multiple courses, and courses taken during the first year carry the same weight as courses taken during the last year when calculating GPA at graduation. GPA is calculated as a course credit-weighted average of grades and presented as a number from 0 to 20, where 10 is the equivalent of obtaining a pass grade in all courses taken. While we do not know when a course was completed, course codes (e.g., MA1201 for introductory math and MA1204 for calculus) provide some information regarding when a course was taken. All students in upper secondary school take a number of mandatory introductory courses ('core courses') equivalent to about a full year of studies. During the years in our study, these courses consisted of introductory Swedish, Mathematics, English, Arts & Crafts, Physical Education, Religion, Science, and Social Studies.

While the exact timing of the courses is not regulated, we rely on the fact that completing the introductory course is formally recommended before taking more advanced courses in the same subject, meaning that the curricula of more academic tracks will typically have students completing the core courses early on. Moreover, if the student passes the course, the grade cannot be revised. Vocational programs typically give practical courses within the chosen field as well as provide provide apprenticeships for their students during the latter two years of secondary school, meaning that they are likely to complete the academic part of the program early on. Consequently, we use GPA on core courses as a proxy for first-year GPA, and the GPA based on non-core courses as a proxy for second- and third-year GPA. Within a student, the measurement error caused by proxying the timing of grades has must have mean zero (a positive measurement error in first year GPA will be canceled out by a negative error in latter year GPA). We could be systematically mis-measuring GPA within a period, but this would only be a problem if student curricula is somehow correlated with local fiber roll out, which seems unlikely.

To identify a causal effect, we exploit within-student variation in fiber coverage between the first and second/third year. We associate first-year GPA with coverage in fall of the first year of upper secondary school, and later-year GPA with the average coverage in fall of the second and third year. For our analysis, we standardize GPA to have zero mean and unit standard deviation.

To exploit the large cross-sectional differences in fiber roll out, we require data on location. We use data from the Swedish pharmaceutical register and the national patient register, which records the parish of residence each time a prescription is processed by a pharmacy and during outpatient visits (not including primary care visits). An interesting feature of the pharmaceutical register is that in many cases it reports a location multiple times during each year, allowing us to minimize measurement error due to students moving e.g., to take up university studies in fall in graduation year. While this register only provides information on students who have been in contact with healthcare providers, through a sequence of matches based on the prescriptions and healthcare visits by the individual, younger siblings, and parents (see appendix), we can trace 97 percent of students in our regression sample to a parish during their upper secondary school years (and about 93 percent of all students).

However, attrition is unlikely to be independent of student characteristics and we note that the sample without location differs significantly from the main sample with regard to several observables (see Table 1). Boys are over-represented in the group without location, probably because Swedish girls are issued medical prescriptions more frequently, e.g. contraceptives. As shown in Table 1, parish coverage

increases by about 50 percent (15 percentage points) between the first and last two years. GPA also increases between year 1 and the later years, which could be due to several reasons. It could reflect the effect of initially having to take ‘core courses’ that the student might not find very interesting and only taking courses within their chosen field later, but might also be the result of students maturing or the formation of student-teacher bonds.

Our main regression sample consists of students graduating upper secondary school between 2010 and 2012, excluding students who move between their first and third year of upper secondary school (about 10 percent of students). While dropping these students could bias our estimate if the decision to move is related to broadband expansion, our main concern is that we would risk capturing other effects of the move, such as the effects of changing schools, peers, and neighborhood. Another concern is measurement error in our parish variable, e.g., if a student’s parents are separated, we might match their location based on the father’s residence in one year and the mother’s residence in the next. It is only when we consistently match a student to the same parish during all three years of upper secondary school that we can be reasonably sure that we are observing the true parish of residence. However, all our results are robust to including movers. Since we require data on fiber coverage when starting upper secondary school, our earliest cohort consists of students starting in fall 2007 (i.e., graduating in summer 2010).

## 4 Empirical specification

For our individual fixed effects specification, we assume that the relationship between GPA and fiber coverage for individual  $i$  living in parish  $p$  in year  $t$  can be described as

$$GPA_{ipt} = \beta Fiber_{pt} + \alpha_i + \gamma_t + \varepsilon_{ipt} \quad (1)$$

For  $\hat{\beta}$  to represent an estimate of the causal effect of fiber on GPA, we require that, conditional on time and individual fixed effects, students in less treated parishes represent an unbiased estimate of the counterfactual GPA difference for more treated students (the parallel trends assumption). Formally:

$$E[GPA_{0ipt} | \alpha_i, t, Fiber_{pt}] = E[GPA_{0ipt} | \alpha_i, t] \quad (2)$$

where  $GPA_{0ipt}$  denotes the (counter-factual) GPA of individual  $i$  in the absence of fiber roll out. By only exploiting variation over a short period of time (two years), we reduce the risk of biased estimates due to transitory GPA shocks. However, we also attempt to control for violations of the parallel trends assumption by estimating specifications with school-specific and parish-specific linear trends.

It can be useful to think of our specification as the reduced form of an instrumental variables specification, where coverage at parish level is used as an instrument for actual household take-up of fiber (which we do not observe). While the decision to connect one’s home to the local fiber network is endogenous, short-term changes in parish coverage are arguably exogenous from the student’s perspective, and a valid instrument for household take-up of fiber. Our estimates of  $\beta$  represent intention-to-treat (ITT)



estimates, meaning we estimate the effect of gaining access to fiber. Since actual take-up of fiber is less than 100 percent, our estimates represent a lower bound of the average effect of the treatment.

## 5 Results

Before presenting the estimation results, we provide some descriptives on the relationship between fiber coverage and GPA. In order to visualize the pre- and post-treatment GPA, we sort students into ‘control’ and ‘treatment’ groups by splitting the distribution of the change in fiber coverage between 2007 and 2011 at the median. Figure 4 presents the trends based on the distribution of fiber roll out. To eliminate confounding effects of fiber roll out occurring before 2007, the diagram only includes parishes with zero fiber coverage in 2007 (around 60 percent of parishes). Since 2005, the gap between treatment and control has increased by about 5 percent of a standard deviation. Turning to the GPA gender difference, the gap between girls and boys has increased by about 4 percent, as Figure 5 shows. Note that in order to better illustrate the change in the gender gap, mean GPA for both boys and girls is normalized to zero in 2003 in Figure 5. The absolute gender gap during the period is on average 40 percent of a standard deviation.

Fiber roll-out is not randomly assigned. In Table 2, we document differences between the ‘treatment’ and ‘control’ groups for a number of observables. The differences in parental education and income suggest that students who receive more intensive treatment are positively selected. Our identifying assumption is that treatment, conditional on our covariates, is as good as random. As a balancing test, we predict treatment status using pre-determined variables with and without fixed effects (table 3). Here, we include all student-year observations with non-missing data on the parents. As expected, once we condition on parish and cohort fixed effects, the explanatory power of the socio-economic characteristics of the parents goes to zero.

Table 4 presents our baseline estimates. In our OLS regression, we define fiber coverage as the average coverage during all three years of upper secondary school. The size of our OLS estimate (column 1) is about twice that of our preferred difference-in-differences estimate (column 2), putting the fiber effect at a negative 9 and 4 percent of a standard deviation, respectively. The interpretation of the marginal effect differs slightly between the two, as the within-student estimate should be interpreted as the ‘contemporaneous’ effect of going from 0 to 100 percent parish fiber coverage, whereas the marginal effect of the OLS estimate is the effect of going from 0 to 100 percent parish coverage during all three years of high school. Only the fixed-effect estimate is statistically significant.

A linear trend explains on average 70 percent of the variation in fiber expansion. This suggests that, by controlling for a linear trend at the parish level, we are eliminating some of the true effect since we are only exploiting “jumps” in coverage. Put differently, since treatment resembles a linear trend, joint estimation is problematic. In column 3 of Table 4, we introduce a linear trend at the school level instead. A school-specific GPA trend can be better identified separately from fiber roll out, since students in a school come from different parishes and vice versa. While we prefer the conservative estimate in column 2, this is likely a lower bound of the true effect size. We also estimate a specification with both a school-specific and parish-specific trend. The effect is still statistically significant, but small (column 4).

## 5.1 Heterogeneity

Table 5 documents heterogeneity in the causal effect of fiber. We predict the GPA of boys to be affected to a larger degree more than that of girls, and the results confirm this prediction. As the estimates in columns 1 and 2 show, the point estimate for boys is about 25 percent larger than for girls, with the effect size estimated at 4.3 and 3.5 percent, respectively. As mentioned previously, we observe a 4 percent increase in the GPA gender gap between 2003 and 2012. In 2012, the average parish coverage across all students in our sample was 55 percent. Assuming that fiber had no impact on GPA back in 2003, as Sweden was then still in the very early stages of extending coverage, we can do a back-of-envelope calculation of the impact of fiber on the gender gap. Taking the difference in point estimates and multiplying by 0.55, the differential effect of fiber explains just over 10 percent of the increase in the GPA gender gap between 2003 and 2012.

We also explore heterogeneity by parental education (columns 3 and 4 of table 5). Splitting the sample between mothers with more or less than 12 years of schooling<sup>4</sup>, we find that the negative effect is highly concentrated among mothers with low education. This is consistent with parental investments playing a role, but could also reflect a correlation with student ability. We also split the sample by the type of program chosen by the student (columns 5 and 6). If students sort into programs by ability, the fact that the effect seems mainly driven by students in vocational programs suggests that low-ability students are hit the hardest. We find similar discrepancies when conditioning on the father's years of school, but smaller in magnitude.

To disentangle the roles of parental investment and inherited student characteristics, we use GPA in the 9th grade (the final year of elementary school) as a proxy for student ability. To address some of the concerns associated with conditioning on a potential outcome, we split the sample into 9th grade GPA quartiles within each parish to have a prior distribution that is conditional on local broadband coverage and other local characteristics. When we run regressions on GPA quartiles by maternal education (see Figure 7), we find that maternal education is important even when conditioning on prior GPA, but only for students in the lowest quartile.

Finally, we hypothesize that urban amenities may attenuate the fiber effect. If students already have access to a large number of leisure goods and activities, this may reduce the impact of high-speed broadband. We find that the fiber effect is reduced when we only include students living in major cities but not by much (column 7 of table 5).

## 5.2 Robustness checks

To test the robustness of the results, we first counterfactually introduce fiber 1 to 3 years earlier. The results of this exercise are presented in Figure 8. The counterfactual roll out does not yield any significant effects. However, this does not constitute a clean placebo test, since the roll out we exploit in this paper

---

<sup>4</sup>12 years of schooling is equivalent to at least a degree from a three-year academic track in upper secondary school. We drop the 10 percent of students missing data on maternal education is missing.

is likely to be related to earlier broadband investments. Due to lack of data on broadband expansion pre-2007, we can only provide limited evidence of this. In 2000, prior to the first wave of subsidies to broadband expansion, the Swedish government sought to identify rural areas where commercial investment was unlikely to happen (The Broadband Committee, SOU 2000:111, 2000). The earmarked funds were distributed among municipalities in proportion to the estimated cost of extending coverage. Given this subsidy schedule, a larger absolute gap between projected cost and awarded funds should, *ceteris paribus*, be negatively correlated with broadband coverage. Interpreting the cost-subsidy gap as a proxy for the roll out speed of first-generation broadband, we regress the increase in municipal fiber coverage during 2007-2011 on the cost-subsidy gap (not reported) and find that an additional 10 million SEK in subsidies back in 2000 is associated with an increase in the expansion of coverage during 2007-2011 of 2 percentage points<sup>5</sup>, suggesting that the effects of these early subsidies still affected expansion rates and that roll out rates are positively autocorrelated. In addition, using a 2003 PTS survey detailing the number of ISPs that owned local internet infrastructure in Swedish towns and cities, we find that additional operators in 2003 are positively correlated with broadband expansion during 2007-2011<sup>6</sup>.

Table 7 documents further robustness checks. First, we split the sample by initial treatment level. A majority of students reside in urban parishes with coverage above the median. Nonetheless, it is interesting to note that the negative effect is found only for the sample with high initial coverage. This suggests a non-linear effect, which is consistent with the fact that many online activities can be described as network goods with positive spillovers among peers, such as online gaming and social media. We also estimate our baseline specification without foreign-born students in the sample. If foreign-born students improve their GPA less than native-born students due to e.g., discrimination, this could confound our results, as immigrants are over-represented in the more intensely treated parishes (see Table 2). However, it does not seem as though this discrepancy is driving our results (column 3 in Table 7).

### 5.3 Effects on dropping out and tertiary eligibility

Next, we examine the effect of fiber on two extensive margins, dropping out and being eligible for tertiary education<sup>7</sup>. For students who drop out, our records only indicate that they were enrolled but did not receive a diploma at the time of graduate. We do not know any course grades. With only a single data point per student, we resort to a family fixed-effects specification where we compare dropout rates between siblings:

$$D_{hpt} = \beta_1 Fiber_{pt} + \beta_2 \mathbf{X} + \alpha_h + \gamma_t + \varepsilon_{ipt} \quad (3)$$

---

<sup>5</sup>Significant at the 5 percent level

<sup>6</sup>The number of ISPs operating in an area has been used as a proxy for local coverage in previous research (Vigdor et al., 2014; Kolko, 2012).

<sup>7</sup>Students are eligible to apply to universities and university-colleges if they obtain passing grades on at least 90 percent of their course credits, including introductory Swedish, English, and Math.

where  $D$  is an indicator for dropout or eligibility status for a student in family  $h$ , residing in parish  $p$  and (potentially) graduating in year  $t$ . Our sample consists of students graduating between 2007 and 2012 with at least one sibling graduating during the same period. As a fiber measure, we use fiber coverage in the year of graduation to increase sample size. We exclude families who move between the graduation of the eldest and youngest sibling. We control for a family fixed effect as well as gender, type of program and ninth grade GPA. As shown in table 6, we do not find any significant effects on either margin. Running the family fixed-effects regression for GPA reveals a point estimate similar to our baseline regression (5 percent of a standard deviation), but we cannot statistically distinguish the effect from zero.

## 5.4 Effects on student time use

Perhaps the most obvious mechanism by which broadband could affect academic achievement is student time use. As explained in section 2, the negative effect we find in our reduced form analysis could be caused by lowered marginal costs of activities that compete with schoolwork for the student's time. Using survey data, we provide evidence that the time spent online increases more in areas where the broadband roll out is more extensive. The Program for International Student Assessment (PISA) is a large international triennial survey of the skills and knowledge of 15-year-old students. In 2012 and 2015 waves, the student questionnaire included several questions on the intensity of ICT use. One of these is particularly well suited to our study: "*During a typical weekday, for how long do you use the internet outside of school?*". The intensity of internet use during a weekday is a plausible culprit, as it directly competes with time devoted to homework.

Unfortunately, the waves do not overlap with the period in our main study, but fiber roll out was still very much ongoing in 2015. Each wave includes a representative sample of about 5000 Swedish 15-year-olds, most attending the final year of lower secondary school (9th grade). While student residence is not reported in detail, we do know if the student is located in a metropolitan area, a large city, or a smaller town/rural area. As the metropolitan areas were almost fully covered by fiber in 2015, the rest of the country was catching up. Thus, if fiber access causes more internet use, internet use is expected to increase less in metropolitan areas than in the rest of Sweden between the two survey waves (2012 and 2015). We test this hypothesis using a simple diff-in-diff specification:

$$Highuse_i = \alpha + \beta Metro_i + \gamma D_{i,2015} + \delta Metro_i \times D_{i,2015} \quad (4)$$

where we define high internet use as a dummy indicating whether student  $i$  reported spending more than two hours per weekday using the internet.  $Metro$  is a dummy for attending a school in a metropolitan area and  $D_{2015}$  is a dummy for the 2015 wave. The coefficient of interest is  $\delta$ , where  $\delta < 0$  indicates that time spent online increased more in non-metropolitan areas between 2012 and 2015. As reported in Table 8,  $\hat{\delta}$  has the expected negative sign and is statistically significant at -4.3 percentage points or about -7 percent of the "High use" sample mean of 56 percent in 2012. This suggests that fiber roll out actually causes students to spend more time online, and provides a plausible mechanism for the negative effect

we find in our main analysis. The point estimate is unaffected if we define high use as spending more than four hours per day online, but less precise.

## 6 Conclusion

In this study, we provide evidence of a negative causal effect of high-speed broadband on the GPA of students in upper secondary school. Increasing local coverage from zero to 100 percent reduces expected GPA by 4 percent of a standard deviation in our preferred specification. The effect is larger for boys and for children born to parents with low education. We find that parental resources can mitigate the negative effect even when accounting for differences in student ability. While our estimates are quite small, there are several reasons to consider our preferred estimate a lower bound of the average treatment effect. First, we can only estimate the ITT effect. Scaling our estimates by the first stage would yield a significantly larger LATE estimate. Second, our estimates control for a linear trend at the parish level, which is to some extent correlated with treatment. With the current pace of ICT development, it should only be a matter of time before high-speed internet access is practically universal. With the booming number of services provided online, the size and scope of the effect of high-speed internet on student performance is likely to increase.

Our evidence suggests that part of an increased GPA gender gap can be explain by the fiber effect being more harmful to boys. Fortin et al. (2015) identify lowered post-secondary academic ambitions among boys as one of the main causes of a widening GPA gender gap. Our findings suggest that the GPA gap has increased due to boys reducing academic effort in favor of leisure, which is consistent with decreasing post-secondary ambitions and the labor market effects of leisure focused technology documented by Aguiar et al. (2017). At the extensive margin, we do not find any evidence suggesting an increased probability of dropping out of upper secondary school although our methodology here is lacking. We also provide suggestive evidence that parental ability plays a role in mitigating the negative effect of high-speed internet, with the estimated effect on students born to a highly educated mother being just a quarter of the size of the effect on students born to a mother with low education.

Finally, we provide a plausible mechanism using PISA survey data. Students living in areas with more rapid broadband expansion report a larger increase in hours spent online. More time spent online could crowd out time spent on e.g., homework, ultimately causing the drop in GPA.

## References

- Aguiar, Mark, Mark Bilal, Kerwin Kofi Charles, and Erik Hurst**, "Leisure Luxuries and the Labor Supply of Young Men," Working Paper 23552, National Bureau of Economic Research June 2017. DOI: 10.3386/w23552.
- Akerman, Anders, Ingvil Gaarder, and Magne Mogstad**, "The Skill Complementarity of Broadband Internet," *The Quarterly Journal of Economics*, November 2015, 130 (4), 1781–1824.
- Björklund, Anders and Kjell G. Salvanes**, "Chapter 3 - Education and Family Background: Mechanisms and Policies," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 3, Elsevier, January 2011, pp. 201–247. DOI: 10.1016/B978-0-444-53429-3.00003-X.
- Faber, Benjamin, Rosa Sanchis-Guarner, and Felix Weinhardt**, "ICT and Education: Evidence from Student Home Addresses," Technical Report w21306, National Bureau of Economic Research, Cambridge, MA June 2015. DOI: 10.3386/w21306.
- Fortin, Nicole M., Philip Oreopoulos, and Shelley Phipps**, "Leaving Boys Behind: Gender Disparities in High Academic Achievement," *Journal of Human Resources*, August 2015, 50 (3), 549–579.
- Grover, Sarthak, Roya Ensafi, and Nick Feamster**, "A Case Study of Traffic Demand Response to Broadband Service-Plan Upgrades," in "Passive and Active Measurement" Lecture Notes in Computer Science 2016, pp. 124–135.
- Guldi, Melanie and Chris M. Herbst**, "Offline effects of online connecting: the impact of broadband diffusion on teen fertility decisions," *Journal of Population Economics*, July 2016, pp. 1–23.
- Kenny, Robert and Charles Kenny**, "Superfast broadband: is it really worth a subsidy?," *info*, 2011, 13 (4), 3–29.
- Kolko, Jed**, "Broadband and local growth," *Journal of Urban Economics*, 2012, 71 (1), 100 – 113.
- Malamud, Ofer and Cristian Pop-Eleches**, "Home Computer Use and the Development of Human Capital," *The Quarterly Journal of Economics*, May 2011, 126 (2), 987–1027.
- OECD**, "National Broadband Plans," *OECD Digital Economy Papers*, 2011, 181.
- Statistics Sweden**, "Use of computers and the Internet by private persons in 2007," 2007.
- Svensson, Måns, Stefan Larsson, and Marcin De Kaminski**, "Professionalization, Gender and Anonymity in the Global File Sharing Community," *Piracy Effect*, 2013, pp. 1–8.
- Swedish Media Council**, "Ungar och medier 2015," <http://statensmedierad.se/publikationer/ungarochmedier/ungarmedier2015.381.html> 2015.

**Swedish Post and Telecom Authority**, "Bredbandskartläggning 2007," [https://www.pts.se/upload/Rapporter/Internet/2008/Bredbandskartlaggning\\_2007\\_2008\\_5.pdf](https://www.pts.se/upload/Rapporter/Internet/2008/Bredbandskartlaggning_2007_2008_5.pdf) 2008.

**The Broadband Committee**, SOU 2000:111, "IT-infrastruktur för stad och land," 2000.

**Vigdor, Jacob L., Helen F. Ladd, and Erika Martinez**, "Scaling the Digital Divide: Home Computer Technology and Student Achievement," *Economic Inquiry*, 2014, 52 (3), 1103–1119.

Table 1: Summary statistics by location information

	With location	Without location	p-value
Year of birth	1991.9	1991.9	0.00
Boy	0.49	0.69	0.00
GPA at graduation	14.1	14.1	0.44
GPA, year 1	13.8	13.7	0.48
GPA, years 2-3	14.2	14.1	0.18
GPA, 9 <sup>th</sup> grade	223.4	222.8	0.39
Fiber coverage, year 1	0.31	-	
Fiber coverage, years 2-3	0.42	-	
Academic program	0.47	0.49	0.00
Immigrant	0.049	0.096	0.00
Mother's years of schooling	11.9	11.9	0.073
Father's years of schooling	11.8	11.9	0.0068
Mother's income (log SEK)	11.5	11.4	0.00
Father's income (log SEK)	11.9	11.8	0.00
Number of students	246,059	8,050	



Table 2: Summary statistics by change in fiber coverage

	Bottom 50 percent of fiber rollout 2007-2011	Top 50 percent of fiber rollout 2007-2011	p-value
Year of birth	1991.9	1991.9	0.027
Boy	0.49	0.49	0.94
GPA at graduation	14.0	14.1	0.00
GPA, year 1	13.5	13.8	0.00
GPA, years 2-3	14.2	14.1	0.074
GPA, 9 <sup>th</sup> grade	221.0	224.1	0.00
Fiber coverage, year 1	0.22	0.33	0
Fiber coverage, years 2-3	0.23	0.47	0
Academic program	0.44	0.48	0.00
Immigrant	0.026	0.055	0.00
Mother's years of schooling	11.7	11.9	0.00
Father's years of schooling	11.5	11.9	0.00
Mother's income (log SEK)	11.4	11.5	0.00
Father's income (log SEK)	11.8	11.9	0.00
Number of students	55,914	190,145	

Table 3: Balancing test, predicting fiber by parental schooling and income

Father's income (log SEK)	-0.00664** (0.00293)	-0.00733*** (0.00177)	-0.000321 (0.000202)
Mother's income (log SEK)	0.00346 (0.00278)	0.000875 (0.00144)	-0.000128 (0.000203)
Observations	444,808	444,808	444,808
Father's years of schooling	0.0115*** (0.00142)	0.00785*** (0.000760)	-6.37e-05 (5.13e-05)
Mother's years of schooling	0.00459*** (0.00119)	0.00294*** (0.000706)	-2.89e-06 (6.86e-05)
Observations	424,068	424,068	424,068
Cohort FE	YES	YES	YES
Municipal FE	NO	YES	NO
Parish FE	NO	NO	YES

Standard errors clustered at the parish level.

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

Table 4: Main results

	OLS	FE	FE	FE
Fiber	-0.0937 (0.0666)	-0.0393*** (0.0138)	-0.132*** (0.0147)	-0.0290** (0.0136)
Observations	229,078	494,064	494,064	494,064
Students	229,078	247,032	247,032	247,032
R-squared	0.457	0.901	0.909	0.910
Linear trends	Parish	Parish	School	Parish and School

Column 1 includes controls for parents' age at birth, education, and income, and the student's GPA in 9th grade, as well as a set of dummies for academic program, sex, sibling order, and parish. Columns 2-5 present the individual FE estimates of equation 1. All FE regressions include a full set of year dummies. Standard errors clustered at the parish level. Our sample excludes movers, dropouts, and students in parishes with a negative reported change in fiber coverage during 2007-2011. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 5: Heterogeneity

Fiber	-0.0448** (0.0202)	-0.0349* (0.0184)	-0.0623*** (0.0208)	-0.0158 (0.0175)	-0.0269 (0.0169)	-0.0439** (0.0199)	-0.0321** (0.0156)
Observations	244,306	255,092	260,292	239,106	235,176	264,222	303,528
Students	122,153	127,546	130,146	119,553	117,588	132,111	151,764
R-squared	0.890	0.906	0.884	0.912	0.914	0.893	0.905
Sample	Boys	Girls	Mother low edu	Mother high edu	Academic program	Vocational program	Major cities

This table presents the individual FE estimates of equation 1. Columns 1-2 splits the sample by the sex of the student. Columns 3-4 splits the sample by mother's education (< and > 12 years of schooling, respectively). Columns 5-6 splits the sample by the type of program attended by the student. Column 7 including only students living in major cities. All regressions include a full set of year dummies and parish specific linear trends. Standard errors clustered at the parish level. Our sample excludes movers, dropouts, and students in parishes with a negative reported change in fiber coverage between 2007 and 2011. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 6: Extensive margins

	Dropout	Eligibility	GPA
Fiber	-0.000591 (0.0116)	0.0109 (0.0119)	-0.0489 (0.0328)
Observations	241,470	202,078	202,078
R-squared	0.614	0.960	0.814
Sample mean	0.099	0.894	-

This table presents OLS estimates for the effect of fiber on drop out an eligibility rates. Controls include parents age, schooling and income, program type, sex and 9th grade GPA as well as a full set of parish, sibling order and year fixed effect as well as parish specific linear trends. Standard errors clustered at the parish level. Our sample excludes movers, dropouts, and students in parishes with a negative reported change in fiber coverage between 2007 and 2011. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 7: Robustness checks

Fiber	-0.0404*** (0.0144)	-0.00717 (0.0296)	-0.0437*** (0.0160)
Observations	470,074	106,514	387,550
Students	235,037	53,257	193,775
R-squared	0.901	0.892	0.903
Sample	Excluding foreign born	Low initial coverage (mean = 0,2 %)	High initial coverage (mean = 33,0 %)

This table presents the individual FE estimates of equation 1. Column 1 excludes foreign born students. Column 2 (3) conditions on being in the bottom (top) 50 percent of the fiber coverage distribution in 2008. All regressions include a full set of year dummies and parish specific linear trends. Standard errors clustered at the parish level. Our sample excludes movers, dropouts, and students in parishes with a negative reported change in fiber coverage between 2007 and 2011. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 8: Student time use

	High internet use
Metro	0.00917 (0.0143)
$D_{2015}$	0.157*** (0.0119)
Metro $\times D_{2015}$	-0.0432** (0.0198)
Observations	10,074
R-squared	0.022

The table presents the estimation of equation (3) using individual responses from two PISA waves (2012 and 2015). High use is a dummy indicating more than two hours of internet use on a typical weekday. Standard errors in brackets. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

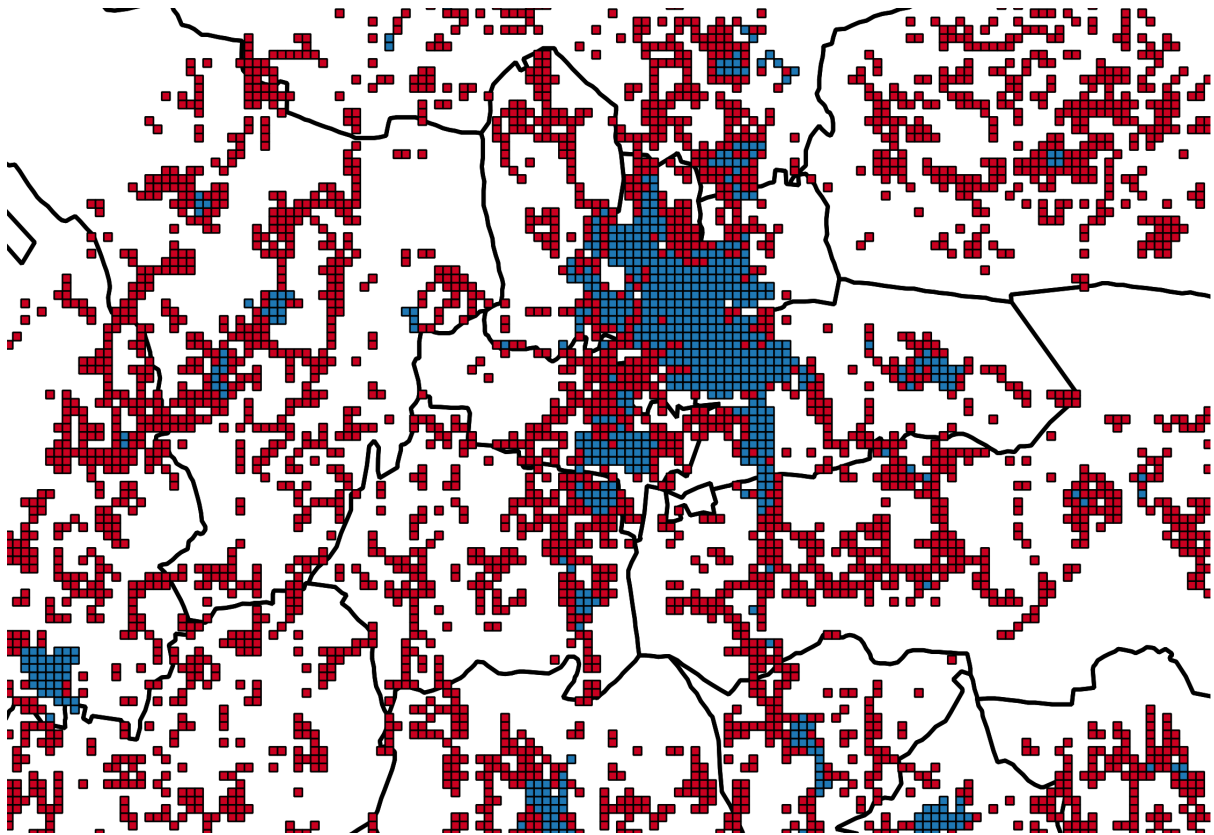


Figure 1: Raw 250\*250 meter grid data. Blue areas are covered, white areas are unpopulated.

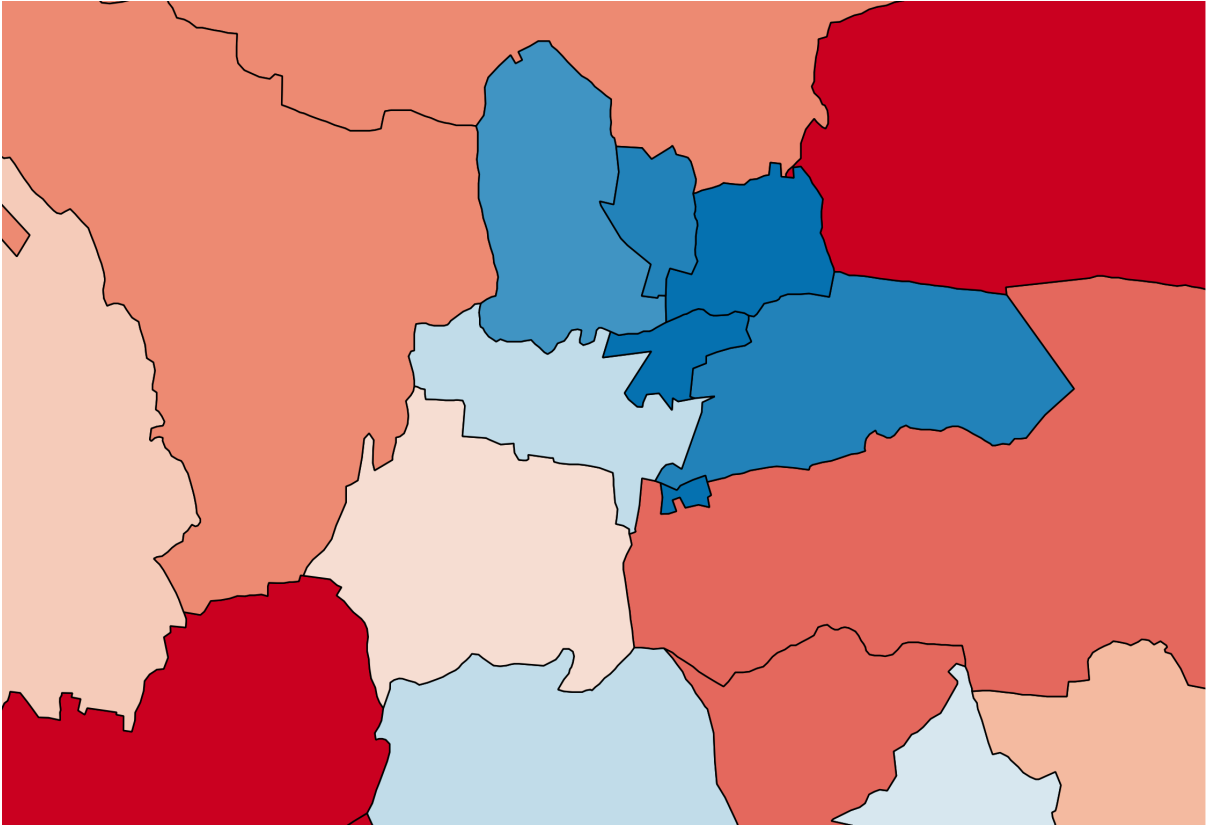


Figure 2: Parish measure of fiber access. Areas in blue have higher coverage.

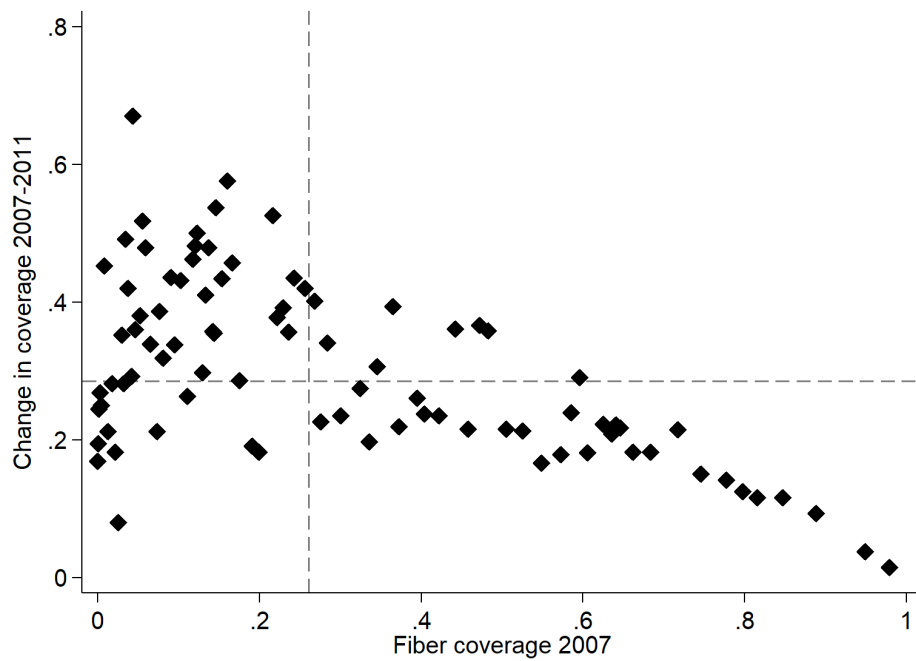


Figure 3: Change in fiber coverage plotted against the initial level of coverage, averages within 87 equal-size bins. Dashed lines represent sample means.

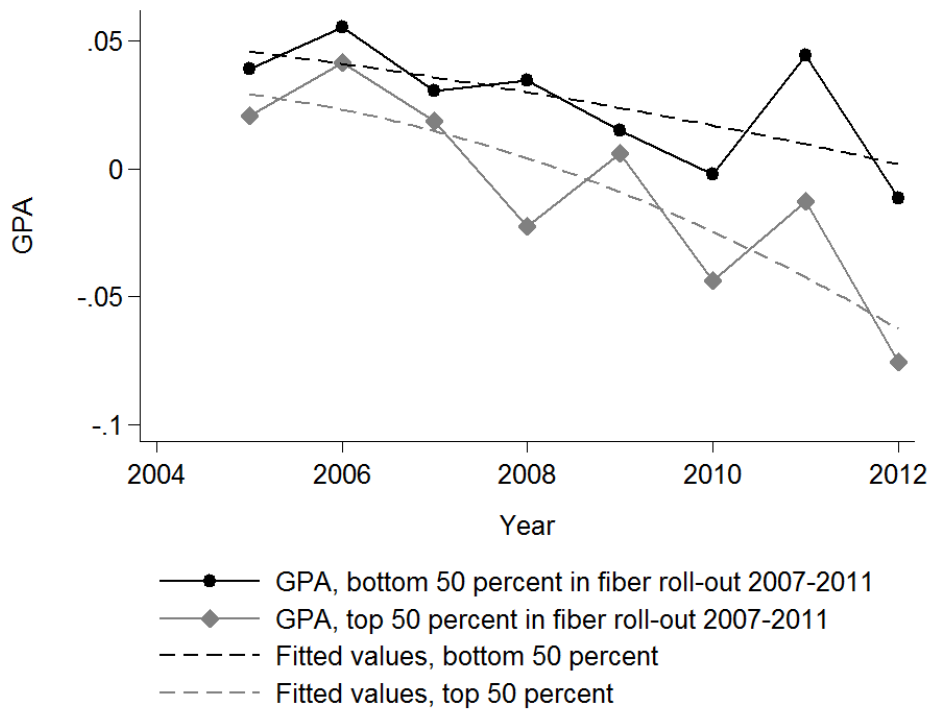


Figure 4: GPA trends for parishes with zero coverage in 2007.

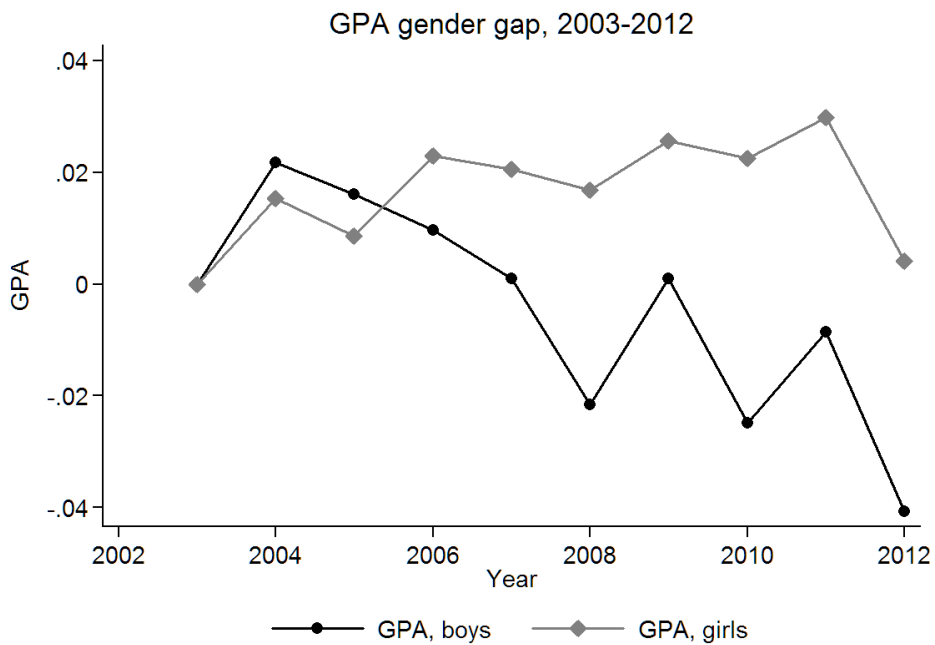


Figure 5: GPA sex differential. GPA is normalized to zero in 2003 for both boys and girls.

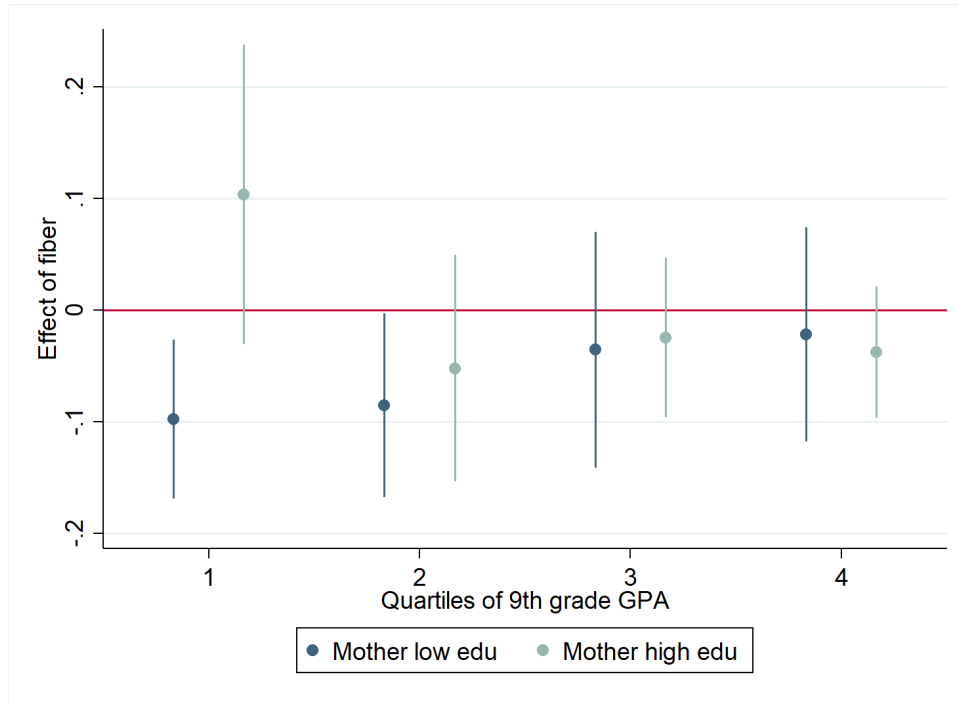


Figure 6: Heterogeneity in the effect of fiber by student ability by mothers' education.

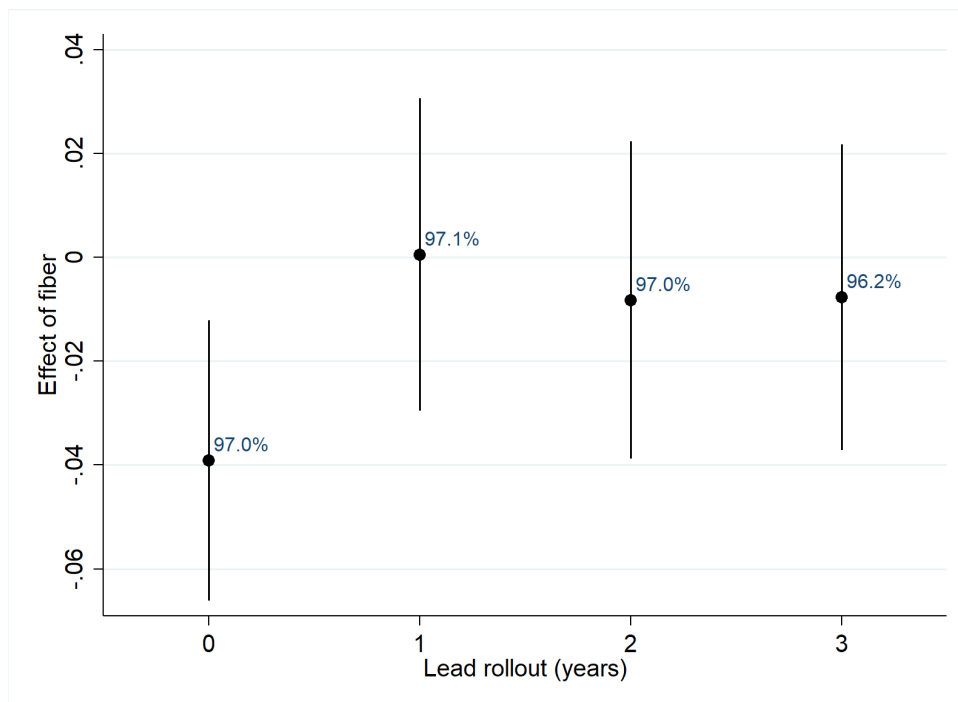


Figure 7: Effect of fiber when the roll out is pushed forward by up to 3 years (e.g., 2004-2008 instead of 2007-2011). The labels denote the share of students that we can match to a parish for all three years of upper secondary school. The "0" estimate corresponds to column 2 of Table 4.



## 7 Appendix

Table A1: Parish matching sequence

Match	Cumulative share of students with location data
Students at time t	37.3%
Parents at time t*	52.8%
Younger siblings at time t <sup>†</sup>	55.4%
Mothers at time t	77.5%
Fathers at time t	87.1%
Students at time t-1	87.6%
Mothers at time t-1	90.9%
Fathers at time t-1	92.3%
Mothers at time t+1	92.8%
Fathers at time t+1	93.1%

\* Conditional on observing the same parish for both mother and father

<sup>†</sup> As students graduating in 2012 is the last cohort on record in our data, we are missing a lot of younger siblings.