WORKING PAPER 2017:5

Erik Grenestam
Martin Nordin

High-speed broadband and academic achievement in teenagers

- Evidence from Sweden

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

Grenestam, Erik[a], Martin Nordin[b]

[a]Department of Economics, Lund University. PO Box 730, SE-220 07, Lund, Sweden.
[b]AgriFood Economics Centre and Department of Economics, Lund University. PO Box 730, SE-220 07, Lund, Sweden.

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, in order to estimate a causal 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 the greatest high-speed broadband expansion also spend more time online during weekdays, suggesting student time use as a potential mechanism.

JEL classification: J24, H52, I24, I28, O33
Key words: Education, Broadband, Internet, High-school, GPA

1 Introduction

In countries worldwide, governments are committing billions of dollars to rolling out high-speed broadband networks (Friedrich et al, 2009). Both the European Commission and the 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 & 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 risks 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 rollout 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 rollout 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. Those authors employ 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 & 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 way of a formal education (Akerman et al., 2015). In this paper we study the effects on academic 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 (see Vermillion, 2003, for a technical overview), 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 (PTS, 2008)[1]. 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

[1]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 is reported at around 70 percent (Statistics Sweden, 2007).

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)[2]. 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[3] 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

[2]The speed advertised by an ISP can be quite different from the actual access, which is why we use third-party measurements.
[3]Latency measures the time it takes to send a signal from the consumer's device to a server and back again. Low latency times can improve the quality of e.g., online gaming.

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[4]. 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 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

[4]The marginal cost can be high if there is congestion, but there have been no reports of network congestion in Sweden. Furthermore, no Swedish ISP has introduced tiered pricing, which has been used as a response to alleged congestion in countries such as the U.S. and Australia.

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 (Statens medieråd, 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., 2014). 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 (Statens medieråd, 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 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 the negative effect is wide, be it through active supervision, inherited traits and abilities (both cognitive and non-cognitive), or resource endowments. Thus we use parental education as a proxy for the family environment[5] when examining heterogeneity in the negative effect.

3 Data

Since 2007, PTS conducts an annual survey of telecom operators and ISPs regarding broadband access in Sweden (PTS, 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[6], 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

[5]See e.g., Björklund and Salvanes (2011), who review the literature on education and family background.
[6]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.

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^2.

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[7]. 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 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

[7]Since drop-outs are not recorded consistently for all years, we excluded them from our sample.

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 privde 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. The measurement error is arguably white noise and the resulting attenuation bias means that our effects should be interpreted as lower bounds. 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 rollout, 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[8], we assume that the relationship between GPA and fiber coverage for individual i living in parish p in year t can be described as:

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

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:

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

where GPA_{0ipt} denotes the (counterfactual) GPA of individual i in the absence of fiber rollout. 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 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

[8] We choose to describe our model using standard FE notation but, as we only have two observations per individual, we could also write it as a first-difference model.

take-up of fiber. Our estimates of β 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 econometrical results, we make a descriptive investigation of 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 rollout. To eliminate confounding effects of fiber rollout 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 rollout 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. Ideally, no pre-determined variable should have any predictive power. As can be seen from Table 3, fathers' income and education are both statistically significant, even after including parish and cohort fixed effects. However, the coefficient is very small, as a one log point increase in income is associated with a 0.04 percent decrease in coverage. For an additional year of education, the effect is much smaller. Arguably, the statistical significance is the result of having a large sample size, and should not invalidate causal interpretation. One should also bear in mind that our main regressions control better for factors such as parental income and education, since we condition on an individual fixed effect.

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 8 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.

When regressing fiber coverage on a linear trend for each parish separately, the average R-squared across all parishes is just over 70 percent. This suggests that, by controlling for a linear trend at the parish level (column 2 of Table 2), we are eliminating some of the true effect. Put differently, since treatment resembles a linear trend, joint estimation is problematic. In column 3 of Table 2, we introduce a linear trend at the school level instead. A school-specific GPA trend can be better identified separately from fiber rollout, 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 negative 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 affcted 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 would explain just over 10 percent of the increase in the GPA gender gap between 2003 and 2012.

We also explore heterogeneity by student ability. As a proxy for cognitive and non-cognitive skills, we use GPA in the 9th grade (the final year of elementary school). To address some of the endogeneity concerns, we split the sample into 9th grade GPA quartiles within each parish, to have a prior distribution that is conditional on initial local broadband coverage and other local characteristics. Figure 6 presents the results. While estimates are imprecise, the results suggest that low-ability students experience a more negative effect. When we perform the same analysis by maternal education[9] (see Figure 7), we find that low-ability students born to mothers with low education suffer the greatest negative effects[10]. Maternal education has an offsetting effect, greatly reducing the negative impact of fiber access for low-ability students, suggesting that parental intervention has an important role to play in mitigating the effect. As shown in Table 4, the negative effect among children of

[9]We define a high education level as having a degree from a three-year academic track in upper secondary school or more (i.e., more than 12 years of education). We drop the 10 percent of mothers for whom information on education is missing.

[10] See columns 3 and 4 of table 5. We have also examined the role of fathers' education, with similar, although smaller, differences.

mothers with low education is almost 60 percent larger than for the sample as a whole. Finally, we look at rural municipalities separately. The small sample size means low precision, but the results do not indicate that the effect in rural areas much different from that in the full sample. This suggests that differences in local amenities does not have a big impact on the effect.

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. As expected, the counterfactual rollout does not yield any significant effects. However, this does not constitute a clean placebo test, since the rollout we observe 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. Examining broadband subsidies granted in the early 2000s, we find that the Swedish government sought to identify areas where commercial investment would not be feasible to be targeted with broadband subsidies (SOU 2000:111). The earmarked funds were distributed among municipalities in proportion to the estimated cost of extending coverage in sparsely populated areas. A larger absolute gap between projected cost and awarded funds should, ceteris paribus, be detrimental for early broadband rollout. Interpreting the cost-subsidy gap as a proxy for the rollout speed of first-generation broadband, we regress the increase in municipal fiber coverage during 2007-2011 on the cost-subsidy gap and find that an additional 10 million SEK in subsidies back in 2000 is associated with an additional expansion during 2007-2011 of 2 percentage points[11], suggesting that the effects of these early subsidies were still at play during 2007-2011 and rollout speeds are positively autocorrelated. In addition, using a 2003 PTS survey detailing the number of ISPs that owned

[11]Significant at the 5 percent level.

local internet infrastructure in Swedish towns and cities, we find that additional operators in 2003 are positively correlated with broadband expansion during 2007-2011[12].

Table 6 documents further robustness checks. First, we split the sample by initial treatment level. Splitting at the median of the initial *parish* level fiber coverage reveals, unsurprisingly, that a majority of students live in parishes that had initial coverage above the median. A negative effect is found only for the sample with high initial coverage, which suggests a non-linear effect, possibly driven by peer or neighborhood effects. 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, as our baseline estimate barely changes when we exclude foreign-born students (column 3 in Table 6). We also exclude second-generation immigrants with similar results.

5.3 Effects on dropping out and tertiary eligibility

Next, we examine the effect of fiber on the probability of dropping out of upper secondary school. For students who drop out, we do not know course grades and we do not know at what time they dropped out. With only a single data point per student, we have to resort to OLS on repeated cross-sections for students who started between 2007 and 2009 (and would have graduated in 2010 and 2012). As shown in Table 7, we do not find any significant effects on either the probability of dropping out or on being eligible for tertiary studies[13].

6 Effects on student time use

Perhaps the most important channel through which broadband could affect academic achievement is student time use. As explained in section 3, the negative effect we find in our

[12]The number of ISPs operating in an area has been used as a proxy for local coverage in previous research (see Vigdor, 2014; Kolko, 2012).
[13]To be eligible for tertiary education, students must obtain pass grades in 90 percent of course credits, including introductory Swedish, English, and Math.

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 rollout 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 the last two waves (2012 and 2015), 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 in terms of negative effects on grades, as it directly competes with time devoted to e.g., homework. Unfortunately, the waves do not overlap with the period in our main study, but as fiber rollout was still very much ongoing in 2015, this should not be a major issue. Each wave includes a representative sample of about 5000 Swedish 15-year-olds, most attending the final year of lower secondary school (9^{th} 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[14]. 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:

$$High\,use_i = \alpha + \beta Metro_i + \gamma D_{i,2015} + \delta Metro_i * D_{i,2015} \qquad (3)$$

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

[14]The population-weighted change in coverage between 2011 and 2014 was 11 percentage points for metropolitan areas (Stockholm, Gothenburg, and Malmö) and 15 percentage points for the rest of Sweden.

interest is δ, 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 $High\ use$ baseline of 56 percent in 2012. This suggests that fiber rollout 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 largely unchanged if we define high use as spending more than four hours per day online, but less precise.

7 Conclusions

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. The effect is larger for boys and for children born to parents with low education. 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 probably 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. Third, our proxies for first and later-year GPA introduce a (arguably classical) measurement error, which would bias estimates towards zero. 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.

A back-of-envelope calculation indicates that the gender difference in the fiber effect can explain part of the increase in the GPA gender gap during the past decade and a half. Seeking to explain a widening GPA gender gap, Fortin et al. (2015) use detailed Canadian survey data (mostly pre-dating the internet era) to disentangle possible explanations. They identify lowered post-secondary academic ambitions among boys as one of the main reasons, e.g., an increased share of boys aiming for only a two-year college degree. 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. At the extensive margin, we do not find any evidence suggesting an increased probability of dropping out of upper secondary school. 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, M., Bils, M., Charles, K.K., & Hurst, E. (2017). "Leisure Luxuries and the Labor Supply of Young Men." NBER Working Papers 23552.

Akerman, A., Gaarder, I. & Mogstad, M. (2015). "The Skill Complementarity of Broadband Internet." *The Quarterly Journal of Economics*, 130(4), 1781-1824.

Björklund A, Fredriksson P, Gustafsson J-E, Öckert B (2010) "Den svenska utbildningspolitikens arbetsmarknadseffekter: vad säger forskningen?" IFAU Rapport 2010:13.

Björklund A, Salvanes KG: Chapter 3 - Education and Family Background: Mechanisms and Policies. In *Handbook of the Economics of Education*, Volume 3. 2011 edition. Edited by: Eric A, Hanushek SM, Ludger W. Elsevier; 2011:201–247.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *Quarterly Journal of Economics, 120*(2), 669-700.

Faber, B., Sanchis-Guarner, R., and Weinhardt, F. (2015). "ICT and Education: Evidence from Student Home Addresses." NBER Working Papers 21306.

Fortin, N. M. & Oreopoulos, P. & Phipps, S. (2015). "Leaving Boys Behind: Gender Disparities in High Academic Achievement." *Journal of Human Resources* 50(3), 549-579.

Friedrich, R., Sabbagh, K., El-Darwiche, B. and Singh, M. (2009), "Digital Highways: The Role of Government in 21st-Century Infrastructure", Booz & Co., New York, NY, available at: www.booz.com/ media/uploads/Digital_Highways_Role_of_Government.pdf.

Google. (2016). *About Google Fiber*. Available online: https://fiber.google.com/about/

Grover, S., Ensafi, R., and Feamster, N. (2016). "A Case Study of Traffic Demand Response to Broadband Service-Plan Upgrades." *Passive and Active Measurement.* Heraklion, Greece, March 31 – April 1, 2016, 124-135. Switzerland: Springer International Publishing.

Guldi, M., & Herbst, C. M. (2016). "Offline effects of online connecting: the impact of broadband diffusion on teen fertility decisions." *Journal of Population Economics*, 1-23.

Kenny, R., Kenny, C. (2011), "Superfast broadband: is it really worth a subsidy?" *info*, 13(4), 3-29.

Malamud, O., & Pop-Eleches, C. (2011). "Home Computer Use and the Development of Human Capital." *The Quarterly Journal of Economics, 126*(2), 987-1027.

Swedish Post and Telecom Authority (2008). *Bredbandskartläggning 2007.* Available online: https://www.pts.se/upload/Rapporter/Internet/2008/Bredbandskartlaggning_2007_2008_5.pdf

Statens medieråd (2015). *Ungar och medier 2015.* Available online: http://statensmedierad.se/publikationer/ungarochmedier/ungarmedier2015.381.html

SOU 2000:111. *Bredbandsinfrastruktur för stad och land.* Slutbetänkande.

Stiftelsen för internetinfrastruktur (2013). *Bredbandskollen – Surfhastighet i Sverige 2008-2013.* Available online: https://www.iis.se/docs/Bredbandskollen_surfhastighet_i_Sverige_2008-2013.pdf

Svensson, M., Larsson, S., and de Kaminski, M. (2014). ”The Research Bay - Studying the Global File Sharing Community”, in Gallagher, W. T. and Halbert, D. (eds.) *Law and Society Perspectives on Intellectual Property Law.*

Vermillion, W. C. (2003). “End-to-end DSL architectures.” Indianapolis: Cisco.

Vigdor, J., H. Ladd, and E. Martinez (2014). "Scaling The Digital Divide: Home Computer Technology And Student Achievement." *Economic Inquiry* 52(3): 1103-1119.

Wikström C., Wikström M., (2005) “Grade inflation and school competition. An empirical analysis based on the Swedish upper secondary schools.” *Econ Educ Rev* 24(3), 309–322

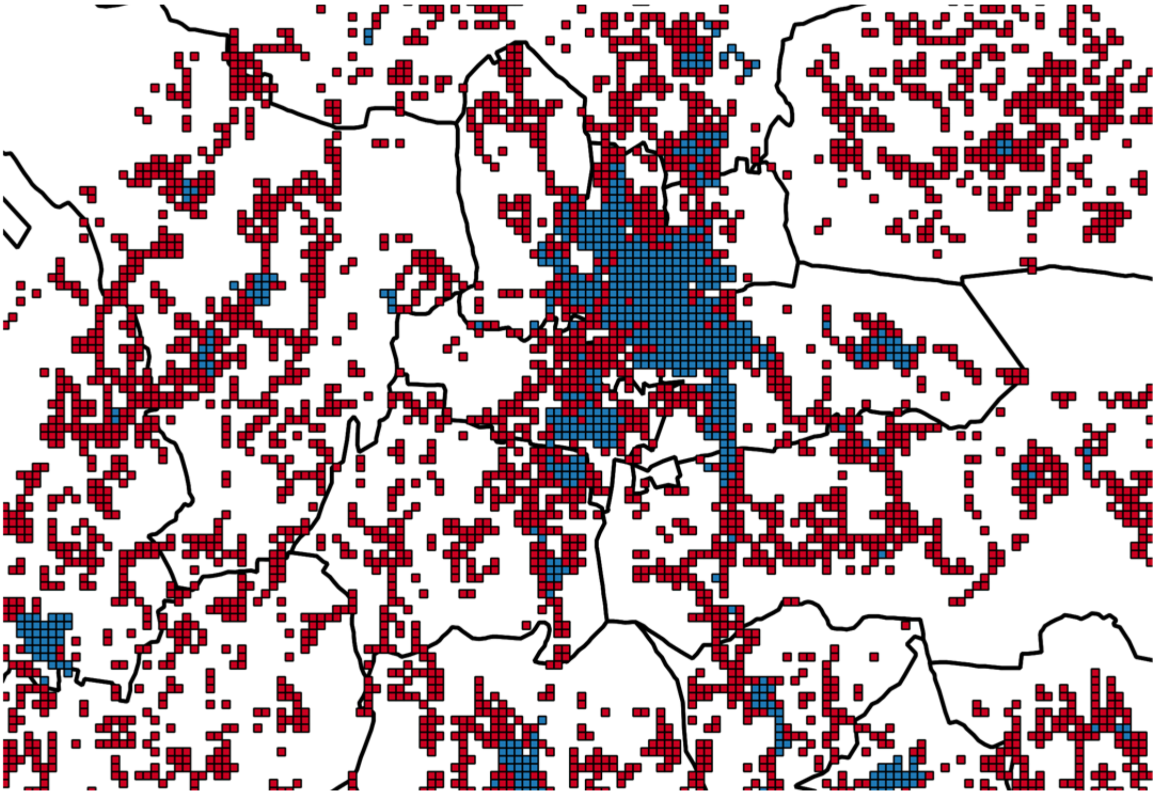

Figure 1: Raw data on fiber coverage in the study area. Blue squares indicate 250 m x 250 m areas that have access to fiber. Red squares are not covered.

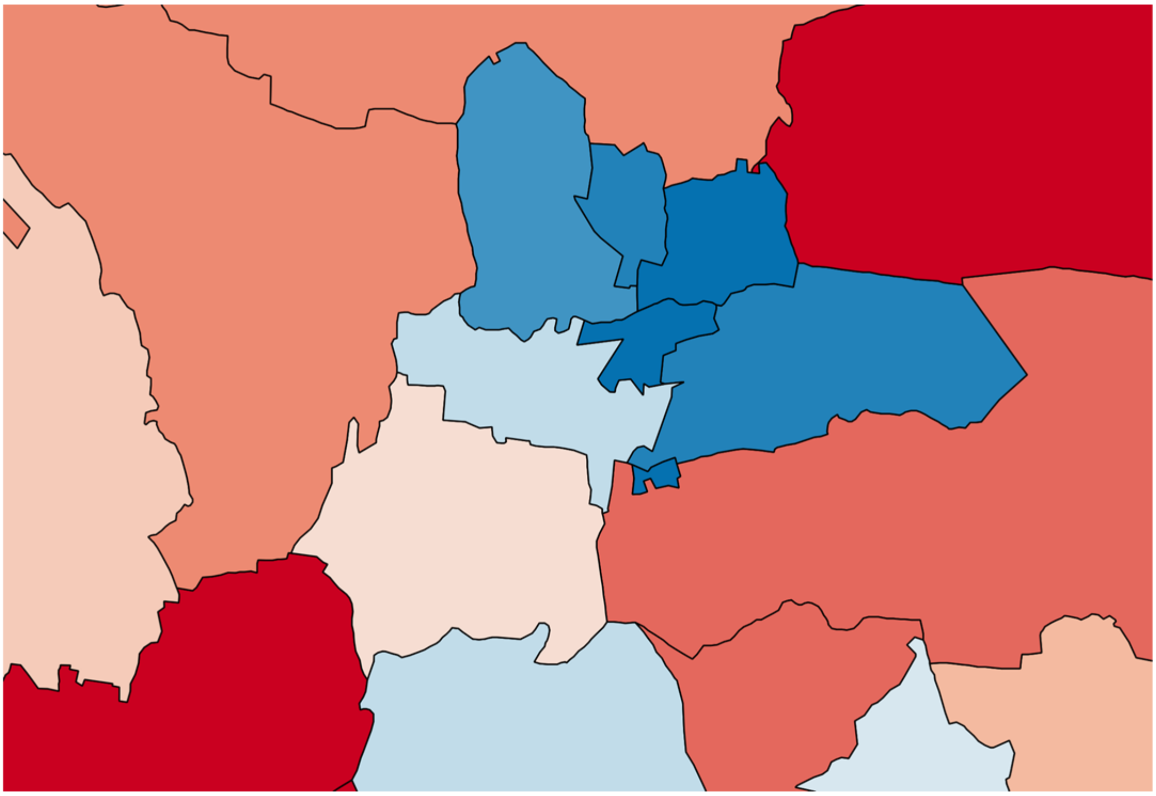

Figure 2: Fiber coverage aggregated to parish level, based on data on the population in each square. Increasingly red areas have close to zero coverage. Increasingly blue areas have closer to 100 percent coverage.

Table 1: Descriptive statistics on students graduating high school 2010-2012

	With location	Without location
Year of birth	1991.85	1991.91**
Sex	0.49	0.69**
GPA at graduation	14.10	14.07
GPA, year 1	13.78	13.75
GPA, year 2-3	14.15	14.11
GPA, 9th grade	223.36	222.81
Fiber coverage, year 1	0.31	-
Fiber coverage, year 2-3 (average)	0.42	-
Academic program	0.47	0.49**
Foreign-born	0.05	0.10**
Mother's years of schooling in 1990	11.87	11.92
Father's years of schooling in 1990	11.80	11.87**
Mother's income in 1990, log SEK	11.46	11.41**
Father's income in 1990, log SEK	11.86	11.81**
Number of students	246,034	8,050

**p<0.05. SEK = Swedish crowns.

Table 2: Descriptive statistics on students graduating high school 2010-2012, by fiber rollout

	Bottom 50 percent of fiber rollout 2007-2011	Top 50 percent of fiber rollout 2007-2011
Year of birth	1991.86	1991.85**
Sex	0.49	0.49
GPA at graduation	14.04	14.12**
GPA, year 1	13.53	13.85**
GPA, year 2-3	14.17	14.15
GPA, 9th grade	220.98	224.06**
Fiber coverage, year 1	0.22	0.33**
Fiber coverage, year 2-3 (average)	0.23	0.47**
Academic program	0.44	0.48**
Foreign-born	0.03	0.06**
Mother's years of schooling in 1990	11.72	11.92**
Father's years of schooling in 1990	11.50	11.89**
Mother's income in 1990, log SEK	11.44	11.47**
Father's income in 1990, log SEK	11.84	11.86**
Number of students	55,914	190,120

**p<0.05.

Table 3: Balancing test. Predicting fiber by (pre-determined) parental schooling and income

	(1)	(2)	(3)
		Panel A: Parental schooling	
Mother's years of schooling in 1990	0.00558***	0.00374***	8.17e-05
	(0.00122)	(0.000702)	(6.70e-05)
Father's years of schooling in 1990	0.0126***	0.00842***	-9.14e-05*
	(0.00143)	(0.000772)	(5.16e-05)
Observations	664,807	664,806	664,807
		Panel B: Parental income	
Mother's income in 1990 (log SEK)	-0.000592	-0.00289*	-0.000405**
	(0.00291)	(0.00166)	(0.000175)
Father's income in 1990 (log SEK)	0.00733***	0.00306**	9.00e-05
	(0.00279)	(0.00142)	(0.000176)
Observations	697,792	697,791	697,792
Cohort FE	YES	YES	YES
Municipal FE	NO	YES	NO
Parish FE	NO	NO	YES

Robust standard error in brackets. SEK = Swedish crowns

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

Table 4: Effect of fiber on GPA, baseline OLS and FE estimates

	OLS	FE	FE	FE
Fiber	-0.0785 (0 .0617)	-0.0391*** (0.0137)	-0.132*** (0.0146)	-0.0290** (0.0136)
Observations	210,264	494,014	494,014	494,014
Students	210,264	247,007	247,007	247,007
R-squared	0.503	0.901	0.909	0.910
Linear trends	PARISH	PARISH	SCHOOL	SCHOOL & PARISH
Sample	FULL	FULL	FULL	FULL

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: Heterogeneous effects: Sex, parental education, type of program, and location

	(1)	(2)	(3)	(4)	(5)
Fiber	-0.0435**	-0.0346*	-0.0600***	-0.0164	-0.0551
	(0.0207)	(0.0192)	(0.0194)	(0.0175)	(0.0364)
Observations	241,636	252,378	257,440	236,574	135,044
Students	120,818	126,189	128,720	118,287	67,522
R-squared	0.900	0.914	0.893	0.919	0.899
Sample	Boys	Girls	Mother low edu	Mother high edu	Rural

All regressions include year and individual fixed effects along with linear parish-specific trends. Standard errors clustered at the parish level. ***p<0.01, **p<0.05, *p<0.1

Table 6: Results of robustness checks

	(1)	(2)	(3)
Fiber	-0.00714	-0.0435***	-0.0402***
	(0.0294)	(0.0159)	(0.0143)
Observations	106,514	387,550	470,024
Students	53,257	193,775	240,484
R-squared	0.892	0.903	0.909
Sample	Low initial fiber (mean coverage in 2008 = 0,2 %)	High initial fiber (mean coverage in 2008 = 33,0 %)	Excluding foreign born students

All regressions include a full set of year dummies and a full set of linear parish-specific trends. Standard errors clustered at the parish level. ***p<0.01, **p<0.05, *p<0.1

Table 7: Effect of fiber on dropout rate and tertiary eligibility

VARIABLES	Dropout	Eligibility
Fiber	-0.0149	-0.0285
	(0.0246)	(0.0284)
Sample mean	0.118	0.858
Observations	246,499	222,535
Students	246,499	222,535
R-squared	0.503	0.909

All regressions include 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, parish, and a parish-specific linear trend. Standard errors clustered at the parish level. ***p<0.01, **p<0.05, *p<0.1

Table 8: Student time use, PISA survey

VARIABLES	High use
Metro	0.00917
	(0.0143)
D_{2015}	0.157***
	(0.0119)
Metro*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

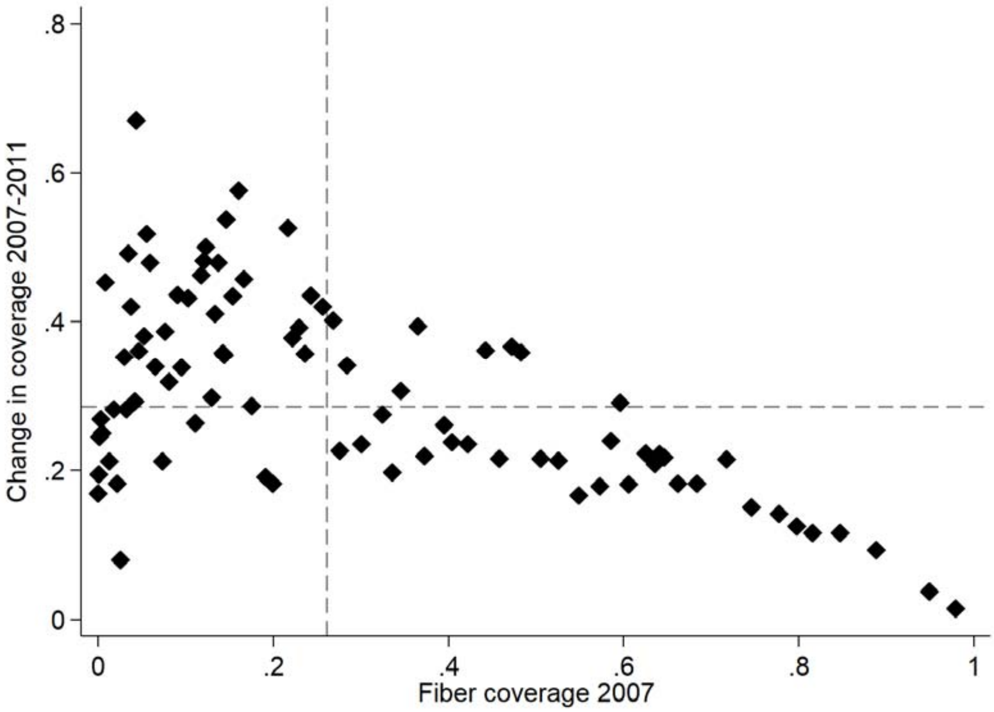

Figure 3: Change in fiber coverage plotted against the initial level of coverage, averages within 87 equal-size bins. Dashed lines represent sample means.

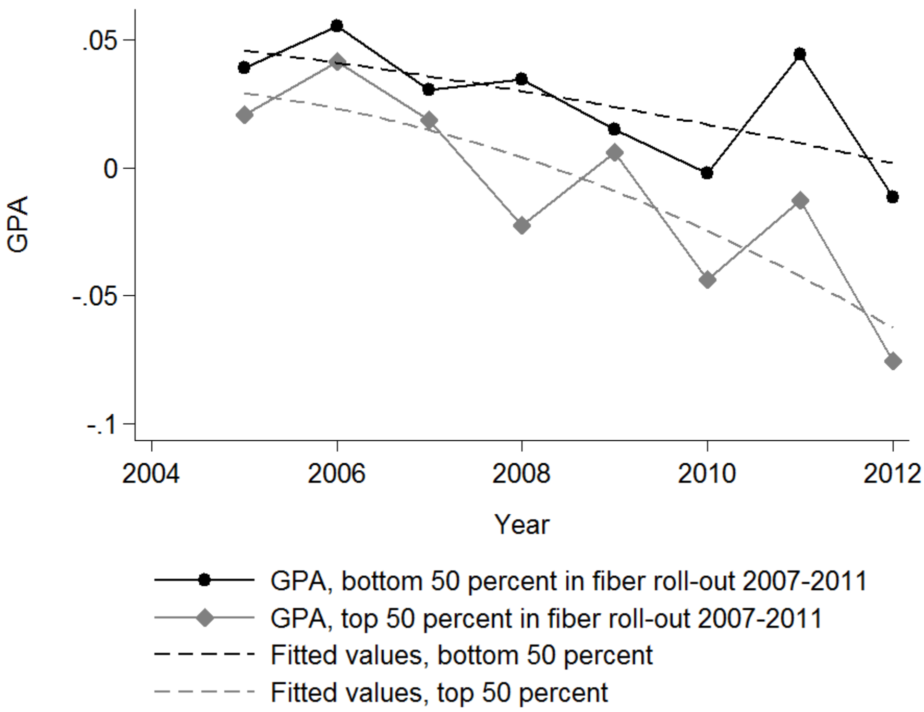

Figure 4: GPA trends, top and bottom quartiles of parish level fiber rollout 2007-2011. Only parishes with zero initial coverage plotted.

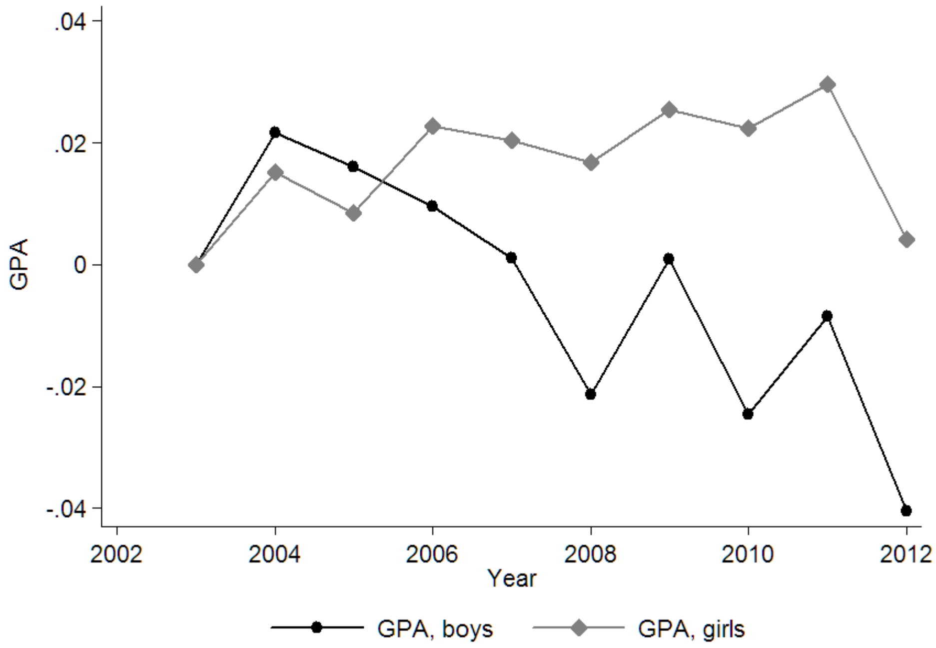

Figure 5: GPA, gender differential. To better illustrate the relative changes in, GPA is normalized to zero in 2003 for both boys and girls.

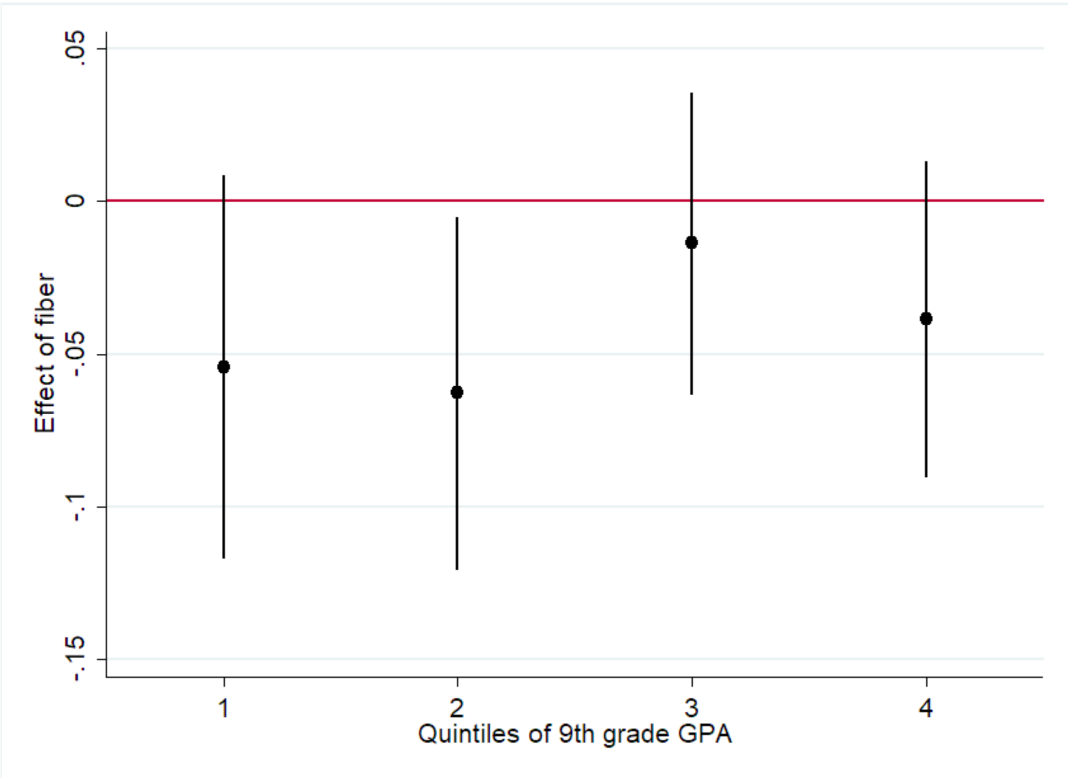

Figure 6: Heterogeneity in the effect of fiber by student ability, as proxied by 9[th] grade GPA.

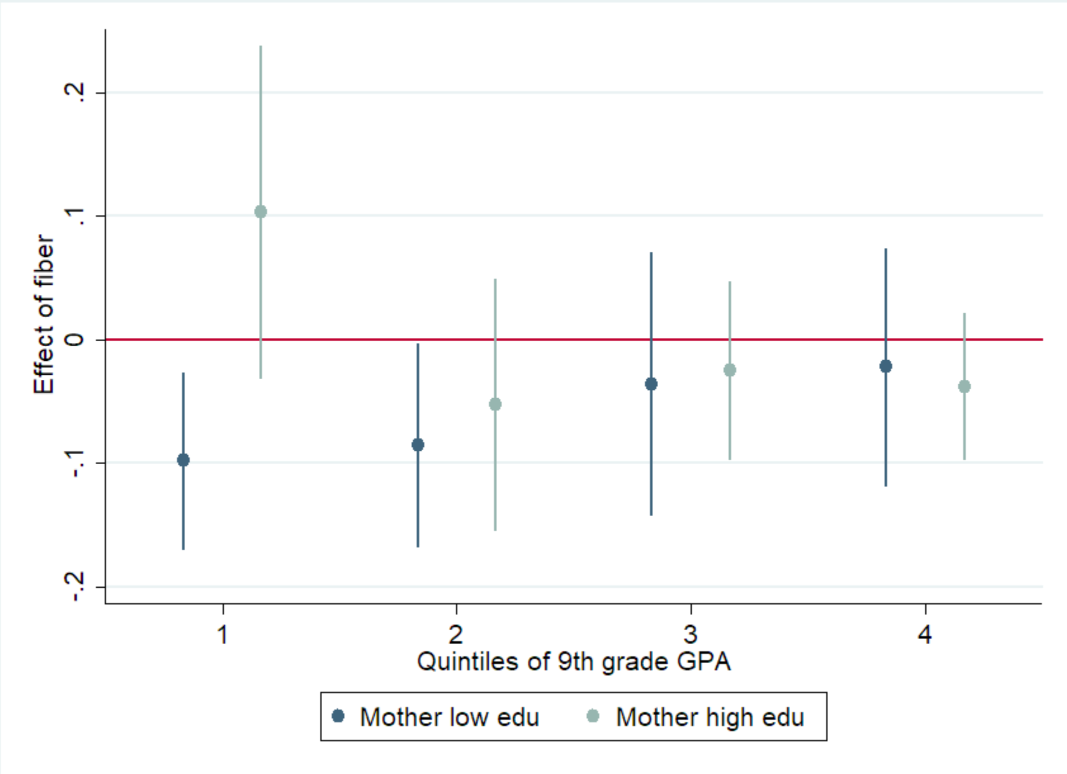

Figure 7: Heterogeneity in the effect of fiber by student ability by mothers' education.

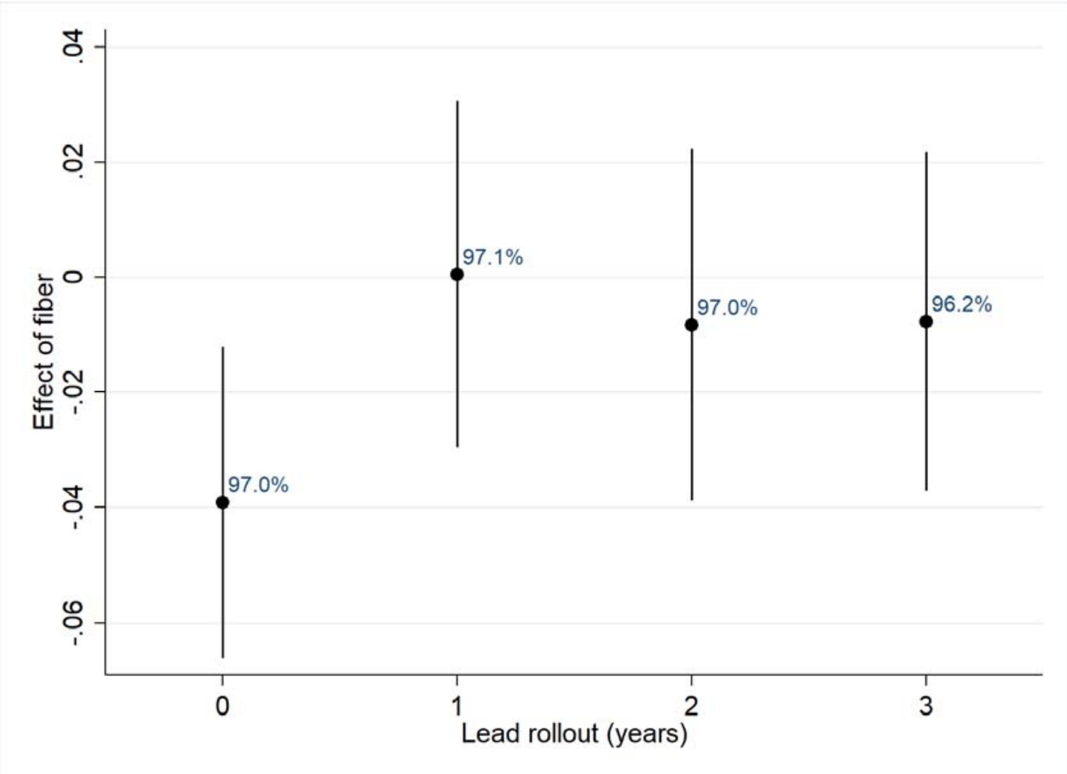

Figure 8: Effect of fiber when the rollout is pushed forward by up to 5 years (e.g., 2002-2006 instead of 2007-2011). The labels denote the share of students that we can match to a parish for all three years of upper secondary school. The "0" estimate corresponds to column 2 of Table 4.

Appendix

Parish matching sequence using register data

Match	Cumulative share of students with location at time t.
Students at time t	37.3%
Parents at time t, conditional on mother and father residing in the same parish	52.8%
Younger siblings at time t*	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%

*As students graduating in 2012 is the last cohort on record in our data, we are missing a lot of younger siblings.

About AgriFood Economics Centre

AgriFood Economics Centre provides economic expertise in the fields of food, agriculture, fishing and rural development. The Centre is a cooperation for applied research between the Swedish University of Agricultural Sciences (SLU) and Lund University. The aim is to supply government bodies with a solid scientific foundation supporting strategic and long-term policy choices.

Publications can be ordered free of charge from **www.agrifood.se**

AgriFood Economics Centre
PO Box 730
SE-220 07 Lund
SWEDEN

www.agrifood.se
mail: info@agrifood.se

AgriFood economics centre