Dynamic Impacts of School-based Internet Access on Student Learning *

Kevin Kho[1], Leah K. Lakdawala[2], and Eduardo Nakasone[3,4]

[1]U.S. Food and Drug Administration
[2]Wake Forest University
[3]Michigan State University
[4]International Food Policy Research Institute

This version: November 2021

Abstract

We investigate the impacts of school-based internet access on standardized test scores in Peru, using over 2 million student observations from a panel of 23,318 public primary schools from 2007 to 2016. We employ an event study approach to identify effects up to 6 or more years after installation, exploiting variation in the timing of internet access due to a large-scale, national program. We find that internet access has modest, positive short-run impacts, but importantly, these effects grow over time. These dynamics underscore the value of using extended evaluation windows to allow the benefits of school-based technology to materialize.

*Email addresses for correspondence: kevinkho256@gmail.com, lakdawl@wfu.edu, eduardo@msu.edu. We thank Jorge Mesinas at the Peruvian Ministry of Education for access to the data from the *Evaluacion Censal de Estudiantes*. We are very grateful to Chris Ahlin, Diether Beuermann, Prashant Bharadwaj, Julian Cristia, Pascaline Dupas, Steven Haider, Scott Imberman, Emilia Tjernström, and Maisy Wong for their insightful and helpful comments and suggestions. We are also thankful to seminar participants at the Michigan Development Day, Michigan State University, and University of Wisconsin-Madison. Previous drafts of this paper were circulated under the title "Impact of Internet Access on Student Learning in Peruvian Schools".

1 Introduction

In recent decades, developing countries have achieved large increases in school enrollment, particularly at the primary level. However, most remain far behind developed countries in terms of school quality as measured by student achievement (Glewwe and Kremer 2006). Traditional policies — such as hiring additional teachers or providing textbooks — do not appear to have improved student achievement in developing countries (Kremer et al., 2013).[1] In turn, new approaches to improving school performance, such as Information and Communication Technologies (ICTs), have garnered increasing interest. The promise of boosting modern-day digital competencies, promoting interactive student-centered teaching models, and providing up-to-date learning materials even in remote areas has encouraged developing countries to invest considerably in ICTs in schools (World Bank, 2018; Escueta et al., 2017; One Laptop per Child, 2016; UNESCO, 2012; Trucano, 2016; International Telecommunication Union, 2014).

Among ICTs, the internet in particular may have important pedagogical uses in developing countries. Internet access can provide underserved students with otherwise unavailable sources of information (Levin and Arafeh, 2002). Similarly, internet can expand teachers' access to references and teaching aids as well as their ability to share information among peers (Jackson and Makarin, 2018; Purcell et al., 2013). However, as with many new technologies, benefits materialize only after a period of learning and adaptation, suggesting the importance of understanding the dynamic effects of ICT interventions over time.

Despite the potential of the internet to improve learning, few studies have rigorously evaluated its impacts on student performance in developing countries. Though previous research in *developed* countries has led to ambivalent conclusions on the effectiveness of internet access as a learning input (Belo et al., 2014; Faber et al., 2015; Gibson and Oberg, 2004; Goolsbee and Guryan, 2006; Machin et al., 2007; OECD, 2015; Vigdor et al., 2014), school-based connectivity can be potentially more important in developing countries due to lower levels of teacher skills, larger class sizes, and limited access to other conventional inputs.[2] Additionally, since the broader literature on ICTs (Escueta et al. 2017 and Bulman and Fairlie 2016) typically examines bundles of interventions such as computer access, learning software, and internet expansion[3], it is not yet clearly understood how

[1]Some notable exceptions exist; for example, see Das et al. (2013) and Muralidharan and Sundararaman (2013).

[2]In a recent paper, Malamud et al. (2019) investigate the impact of home-based internet on Peruvian students' school performance, finding no statistically significant effect on standardized test scores 9 months after the implementation of the program. The authors posit that too little time might be spent on computers at home for any educational benefit to materialize. Relatedly, children might use internet as a tool for entertainment rather than learning. Both of these problems might be reduced when internet is provided at school rather than home.

[3]Some notable exceptions analyze the individual impact of computer access (Beuermann et al., 2015; Cristia et al., 2017; Barrera-Osorio and Linden, 2009; Mo et al., 2013; Toyama, 2015; de Melo et al., 2013; Sharma, 2014; Meza-Cordero, 2017; Bai et al., 2016) or learning software (Bando et al., 2016; Banerjee et al., 2007; Carrillo et al., 2010; He et al., 2008; Linden, 2008; Muralidharan et al., 2019) in developing countries. However, there is little evidence on the impact of internet access.

internet on its own influences learning.[4]

Moreover, most prior studies of internet access — and of ICTs more generally — have been based on short-term observation of small samples, and are thus only designed to detect somewhat larger and immediate treatment effects. Importantly, such studies may overlook potential longer term impacts that may follow from an initial learning period, during which teachers, students, and administrators adapt to new technology. Hence, detecting gains in learning that may arise over such a learning period requires a longer evaluation window.

We examine the impact of internet access on the performance of the universe of students that attended public primary schools in Peru that initially acquired internet between 2007 and 2020 or that remained unconnected by 2020, emphasizing its dynamic effects in schools over time. Over this period, more than 11,300 schools (which jointly enroll about 2 million students per year) gained access to internet. We link administrative data on school-based access to internet with their students' math and reading scores from a large-scale national test that covers nearly the universe of second graders in public schools in Peru between 2007 and 2016. We construct a panel dataset of around 23,300 schools where we observe the scores of about 2.3 million second grade students during our study period. To fully exploit the longitudinal structure of the school data and identify dynamic effects, we employ an event study framework in addition to a trend break analysis — approaches which also allow us to detect and control for pre-existing trends in student performance. Since we observe a large panel of schools over several years, we are also able to assess how other determinants of school performance change over time, tracing out the dynamics of student, teacher, and school-level inputs. This allows us to discuss potential channels through which internet affects student performance, as well as to thoroughly explore the possibility that other confounding factors drive our results.

We exploit plausibly exogenous variation in the timing of internet access and compare cohorts of students who attended these schools before and after they get connected to internet. Using within-school variation in the timing of internet installation, we find that internet access leads to initial modest test score improvements of 0.028 standard deviations in math and 0.018 standard deviations in reading in the first year after installation. Importantly, this advantage grows significantly over time, reaching 0.111 and 0.063 standard deviations five periods after installation for math and reading, respectively. It is important to underscore that our strategies only allow us identify school-level – as opposed to student-level – dynamics. That is, we find that a second grade student in a school that has had internet for 5 years performs 0.111 standard deviations better on math tests than a second grade student at a school that has not yet had internet installed. The trajectory of estimated effects implies that schools become more efficient in using the internet over time to produce improvements in students' test scores.

[4] Previous work (e.g., Cristia et al., 2014; Bet et al., 2014; Sprietsma, 2007) has assessed programs providing school-based internet as part of broader ICT expansion schemes. However, these papers do not aim to discern the effect of internet separately from that of other technologies.

We posit that this growth in our estimated impacts over time reflects an adaption period, during which schools must learn to integrate new technologies. Namely, we observe that schools respond to internet access by hiring teachers with formal training in digital skills, and that this process follows only gradually. In particular, schools are 30% more likely to have a computer-trained teacher by the fifth year after installation relative to before internet access. That the gradual growth over time in test scores shadows growth in the staffing of computer-trained teachers suggests that complementary investment in staff computer proficiency is needed to fully exploit internet-enabled classroom capabilities.[5] This finding aligns with a long line of previous literature examining the impacts of general purpose technologies and the complementary investments and organizational changes that ultimately drive long-run productivity gains (e.g. Brynjolfsson and Hitt 2000).

Furthermore, our data offer evidence for several additional channels through which internet access improves test scores. First, we use a completely separate data source (the Peruvian National Household Survey) to show that internet access at schools leads to meaningful increases in student use of internet. Five years after gaining connections at their schools, self-reported use among public primary school students is 50% higher relative to the period before their local schools become connected. This suggests that increased use of internet by students partially explains the improvements in student performance. Second, we present descriptive evidence of teachers' use of internet as a pedagogical tool. For this purpose, we use two nationally representative teacher surveys (the National Survey of Teachers and the National Survey of Educational Institutions). Teachers report that internet is one of the most important materials enhancing teaching. Moreover, teachers at schools with good quality internet connections also report less difficulty performing teacher activities than those without internet connections at school. Taken together, these findings indicate that access to online materials may boost student performance above and beyond the impacts of direct student use of internet-connected computers.We also find suggestive evidence that gains in test scores are larger in schools with high student-to-teacher ratios and that short-run gains in test scores appear largest for schools with relatively high teacher qualifications. However, the estimates in these subsamples are somewhat imprecisely estimated so differences in estimated effects across the groups are not statistically significant.

Our findings are robust to a number of alternative explanations. Concerning potential endogeneity in the timing of internet access, we find that, conditional on year and school fixed effects and a set of time-varying school characteristics (e.g., school enrollments, infrastructure, and resources), schools receiving access to internet do not exhibit systematic pre-trends in performance or different levels of scores prior to access. Second, we also find that our results are not explained by concurrent changes in other inputs (infrastructure, textbooks, teachers, or computers) or by pre-existing trends that differ by geographic areas, administrative units, or initial test performance. Third, while our

[5] Similarly, evaluations of laptop provision in the U.S. (Hull and Duch 2016) and computer assisted learning in China (Mo et al. 2015) estimate that the effects of ICT interventions grow over time. More generally, Jackson and Mackevicius (2021) find that the benefits of increased school spending rise with years of exposure.

main specifications are based on an unbalanced sample of schools, our results are very similar when using different sample restrictions (including a sample of non-attriting schools). Fourth, analyzing student composition within schools shows that our findings cannot be driven by endogenous sorting of students. Lastly, we show that our results are virtually identical using an alternative estimator that avoids negative weighting and is robust to heterogeneous treatment effects.

We contribute novel insights and perspective to a nascent body of research in developing countries on the educational benefits of school-based internet access, as well as to a wider literature concerning ICTs as schooling inputs.[6] The size and time span of our data present opportunities to complement and contextualize existing knowledge from randomized control trials (RCTs), which are the most common form of evaluating ICTs in relation to academic performance. Whereas RCTs often estimate short term effects (usually within one academic year)[7], we use our extended study period to analyze the effects of internet access more than 6 years after it is introduced to schools. Our results indicate that this longer evaluation window is highly relevant to understanding the impact of internet access, due to the dynamic effects of internet on learning over time.

Additionally, the large scale of our sample — containing over 23,000 public schools — provides power to detect the short-run impacts of internet, which appear to be modest in size. We discover average test score gains of 0.018-0.028 standard deviations in the first full year of internet access (depending on the subject), statistically significant at the 95% level of confidence. Many other evaluations of ICTs report slightly larger point estimates of short-run effects (ranging from 0.052-0.110 standard deviations 5-22 months after initial access; Bet et al. 2014, Barrera-Osorio and Linden 2009, Cristia et al. 2017, Beuermann et al. 2015, Mo et al. 2013), but none of these studies are able to statistically distinguish those effects from zero. One interpretation of our results is that the null findings in these other studies may be due in part to low statistical power stemming from much smaller samples of schools (ranging from 13 to 318 schools, depending on the study). Since we investigate a massive national policy (which affected public primary schools serving roughly 2 million children in a given year), we can assess the impacts of internet access in conditions relevant for policies implemented on a large scale.

Finally, we explicitly identify the gains that internet access produces over hardware resources alone. Anecdotally, the usefulness of school computers without internet access has been limited by lack of access to information (National Public Radio 2012) and the inability to obtain routine maintenance and software updates — particularly in remote, difficult-to-reach locations (One Laptop per Child 2011). Indeed, our data suggest that computers alone (in schools without internet

[6]Another study, Hopkins (2014), also examines the relationship between internet access and test performance in Peru using similar data. However, one important difference between our study and Hopkins (2014) is that we implement a school-fixed effects strategy whereas Hopkins (2014) compares internet-connected schools to non-connected schools. For this reason, we regard Hopkins (2014) as being important for establishing a statistical relationship between internet access and test scores, but one that is ultimately correlational rather than causal.

[7]However, evaluations beyond one year are becoming less rare; Mo et al. 2015 studies the impact of a 1.5 year computer assisted learning program.

access) have only modest impacts on student learning. To the best of our knowledge, the scale, longitudinal length, and setting of this study, along with the comprehensiveness of the available data uniquely address important gaps in the existing literature. More broadly, our work contributes to understanding the role of internet access in economic development. Prior research concludes that the spread of fast internet led to higher employment, incomes, and wealth in African countries (Hjort and Poulsen 2019). Our results imply that increased human capital production may factor importantly in this progress.

2 Setting and data

2.1 Education and ICT Access in Peru

Education in Peru is compulsory and free through the public school system beginning at age 3 and continuing until the end of secondary school. In the past few decades, Peru has greatly increased access to primary school (grades 1-6, approximately age 6-11), raising the net enrollment rate from 85.6% in 1980 to 97.9% in 2015 (The World Bank 2016). At the same time, however, the education budget has seen little growth, and thus greater enrollment over time has eroded per-student resources (Saavedra and Suarez 2002). The World Bank (2012) finds that, within Latin America, only the Dominican Republic has a lower education expenditure-to-GDP ratio than Peru.

This dearth of resources has been accompanied by Peru's poor performance in the OECD's Program for International Student Assessment (PISA) — an international standardized test among 15 year olds. In 2012, Peru ranked last out of 65 participating countries in all three evaluated subjects, with results revealing that most Peruvian students have serious deficiencies in math (75% deficient), science (69%), and reading (60%). In 2015, Peru jumped to the 64th place (out of 70 countries in the evaluation), nonetheless demonstrating that substantial progress remains to be made. Widespread under-preparedness is evident as early as primary school. In 2007, the Ministry of Education began administering yearly standardized tests, the National Student Assessment or *Evaluacion Censal de Estudiantes* (henceforth ECE, described below), to all second graders registered in classes with five or more students. The inaugural results of the ECE in 2007 showed that only 6% of students acquired skills mandated by the national curriculum in math and 12% in reading (Appendix Figure A.1). Despite improvement since then in test scores and in the proportion of students meeting expected skill levels, the quality of schooling has continued to prove inadequate for many children; even by 2016, less than 40% of second graders achieved proficiency in math (46% in reading).

In the early 2000s, the Peruvian government launched *Plan Huascarán*, which produced much of the variation in school internet access observed during our sample period. This project aimed to "incorporate information and communication technologies to increase the coverage, quality, decentralization, democratization, and equity of the Peruvian education system." Project planners ambitiously aimed to install hardware and internet in 32,000 schools by 2020.[8] *Plan Huascarán* targeted schools under public management, particularly in rural, peri-urban, and high poverty areas. Officially, selection into the program was rationed, with each Local Educational Management Unit (UGEL) allowed to request installation for a set quota of schools (see Appendix Figure A.2 for an excerpt of the official Ministry of Education flow chart that outlined the specific prioritization protocol under *Plan Huascarán*). As prerequisites for program selection, schools needed to have electricity and a computer lab with anti-theft measures (i.e., perimeter fencing). Within each UGEL and level, prioritization among qualified schools was officially based solely on the size of the student population, with larger schools receiving higher priority. Lists of eligible schools were aggregated to the regional level and then submitted to *Plan Huascarán* headquarters, accompanied by data sheets on the characteristics of each school listed, a sketch and description of each school's computer facilities, and the discussion minutes from each UGEL.[9] Officially, no school was integrated into the project without all required information.

As a consequence of initiatives such as *Plan Huascarán* and the One Laptop per Child program (OLPC, undertaken by the Peruvian government in 2008)[10], the ratio of students to computers in primary schools fell dramatically from 240 to 6 between 2000 and 2014 . In parallel, the government has steadily increased access to internet in schools (as described in Section 2.2.1). In 2013, the Ministry of Education announced plans to triple the number of schools having internet access.

2.2 Data

Our primary analysis uses school-level data from two sources administered by the Ministry of Education: the *Censo Escolar* (CE), an annual census of schools, and the *Evaluación Censal de Estudiantes* (ECE), an annual standardized test of second graders' skills.

2.2.1 Censo Escolar (CE) and School-based Internet Access

Each year, all school principals are required to submit two forms to the Ministry of Education. Between April and July, principals complete a form on enrollment (by grade and age), teachers (by qualification), available supplies and materials (e.g., books, computers, and laboratories), and infrastructure (e.g., access to utilities, building characteristics, and internet connectivity). Between December and February, another form is completed on year-end pupil outcomes (e.g., promotion

[8] Though teacher training was officially part of *Plan Huascarán*, in practice there was little emphasis on teacher training (Balarin (2013)).

[9] A translated version of the school data sheet is provided in Appendix Figure A.3.

[10] Peru has been the single largest buyer of OLPC laptops and to date has distributed close to one million laptops, mainly targeting school children in poor areas of the country. As our analysis to follow accounts for the total number of computers in a school, including OLPC laptops, we indirectly control for the influence of OLPC. For a discussion of the OLPC program in Peru, see Trucano (2012). In general, impact evaluations of OLPC in Peru suggest that the provision of laptops did not improve student performance (Beuermann et al. 2015; Cristia et al. 2017). We address the issue of concurrent increases in computing resources (including OLPC laptops) explicitly in Section 4.2.

and repetition rates, number of pupils transferring to other schools).[11] We refer to the CE for data on school characteristics such as internet access, enrollment, teachers, educational materials and resources, and physical infrastructure.[12] Between 2007 and 2020, around 31,100 public primary schools reported administrative information in the CE.

We use information from the CE to determine the timing of initial internet connection among schools in our sample. As detailed in Section 2.2.3 below, we focus on the period 2007-2016 (for which we have student test scores). However we use information from the CE until 2020 (the most recent round currently available) to identify the year in which schools initially install internet and assign them accordingly in our event time specification (described in Section 3.1).[13] This allows us to increase our sample size for periods before internet is installed in schools and better assess any potential pre-existing trends in schools prior to access to internet.

Administrators report in the first semester of every year whether their school currently has access to internet. Though some schools report gaps in internet access, the data do not allow us to distinguish between temporary outages and longer-term disruptions to connectivity.[14] Based on this information, we determine the first year in which a school reports gaining access to internet and interpret this as the time of connection. In our estimation framework, this implies a conservative estimate of the impact of internet access, because we treat schools that might have permanently lost their connections as still being connected. Another benefit of using initial internet connection rather than current access is that we avoid bias due to endogenous changes in access. We estimate that 11,310 schools — and the roughly 2 million primary school students in these schools each year — at some point gained internet connectivity between 2007 and 2020. This implies that the rate of internet connection in schools increased from 6% to more than 42% and that the share of students with internet connection in their schools jumped from 28% to 81% over this time period (Figure 1).

Most of the observed expansion in internet connectivity during this period was due to *Plan Huascarán*. In Appendix Table A.1 we verify that the official qualification and prioritization rules

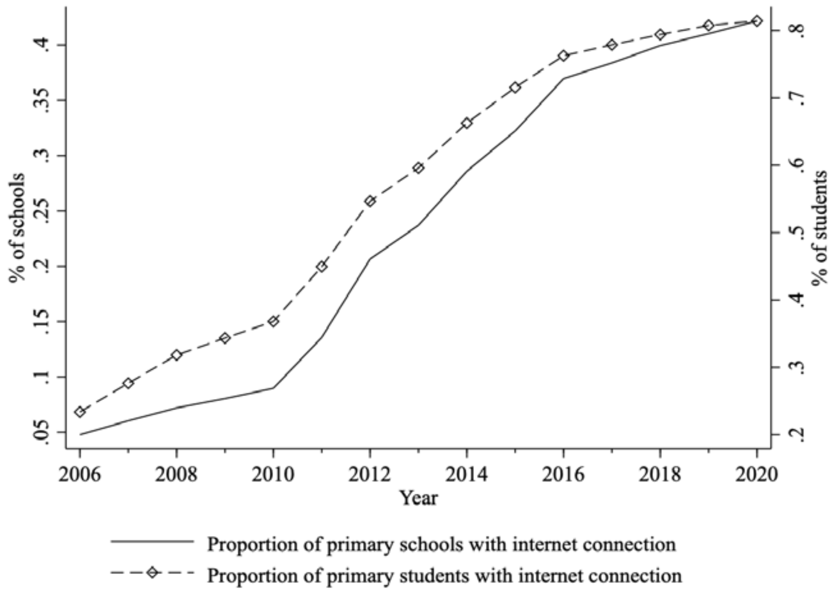
[11]The school year in Peru runs from March to December.

[12]While this information is self-reported by school principals, the Ministry of Education applies different filters and verifies the consistency of the data with secondary data sources. The CE forms are submitted to the Ministry of Education electronically and include consistency rules to avoid reporting errors. Once the electronic forms are submitted, the Ministry of Education validates the information with teacher payrolls, delivery records of materials, and historical information on enrollment. To provide further evidence that schools do not strategically misreport information (for example, inflating enrollment), we compare second grade enrollment as reported in the CE to the number of students scheduled to take the exam in the ECE data. The median discrepancy in these numbers is 0 and the average is 0.33 students (i.e. on average 0.33 more second grade students are reported in the ECE than in the CE); since the CE is reported earlier in the school year than the ECE, this discrepancy could arise naturally due to students entering after the CE forms have been submitted.

[13]CE data have been collected since 2000. However, the CE has only included variables for internet access since 2006.

[14]Out of the 31,111 public primary schools with at least one year of data in the CE between 2006 and 2020, 29.4% report not having access after having access in a previous year; about 44.1% of those schools regain internet access at a later point.

Figure 1: Internet Connectivity in Primary Public Schools, 2006-2020



The stock of schools that gained internet connections is based on the first year in which they report internet access in the Peruvian *Censo Escolar*.

set by the Ministry of Education do in fact predict actual installation.[15] Schools received priority primarily based on quotas by province (Local Education Management Units, UGELs), high poverty status, location in a rural or marginal urban area, public versus non-public management, the presence of required infrastructure (including electricity, a computer lab, and anti-theft measures such as perimeter fencing), and enrollment. Column 1 includes these characteristics and state-year fixed effects to control for aggregate and state-level trends in internet connectivity. To capture high poverty status, we include district-level fixed effects. We also include UGEL fixed effects to account for the UGEL-specific quotas. As expected, all required/prioritization characteristics positively and significantly predict internet access. In column 2, we add a control for perimeter fencing (available only for 2010 and later). Finally, in column 3, we include information from school data sheets (number of computers used for instruction, number of computers used for administrative purposes, and number of teachers), respectively. Since these factors predict internet access and are also likely to influence student performance directly, we control for all of these measures in our main specifications (except perimeter fencing due to data limitations).[16]

[15] Note that the official prioritization rank and program status of schools under Plan Huascarán is not available.

[16] It is not possible for us to calculate exactly how much of the overall increase in internet access over this period is directly attributable to *Plan Huascarán* versus other efforts. We discuss the robustness of our results to restricted samples of schools that likely gained access to internet only via *Plan Huascarán* in Section 4.3.

2.2.2 Evaluación Censal de Estudiantes (ECE)

The Ministry of Education also mandates the *Evaluación Censal de Estudiantes* (ECE), a yearly standardized assessment of second graders' skills, which is administered in late November or early December (before the end of the school year). In order to ensure uniform testing environments — and to prevent content leaks or influence from school personnel — the Ministry hires independent staff to administer the test in all schools simultaneously. As the same test is given to all schools, neither the content nor the testing environment varies by school characteristics. Furthermore, the ECE was designed for comparability of results over time: experts defined a set of basic competencies prior to the test's first administration. Hence, since its inauguration in 2007, the ECE has assessed the same skill sets with considerable consistency. We use student-level ECE scores from 2007 to 2016.[17] To account for differences in difficulty across cohorts and year-to-year changes in test score dispersion, we standardize ECE scores across the universe of test takers in public schools within each year. It is very important to highlight that the ECE is given to a *different* cohort of second grade students each year, so we have a repeated cross-section of students from a panel of schools. That is, we observe test scores for each student only once (in second grade).

The ECE gauges the academic performance of the vast majority of second graders in Peru, targeting all public and private schools that meet two criteria: 1) having at least five second graders enrolled during the test year, and 2) using Spanish as the primary language of instruction. The rationale for the first criterion is entirely budgetary, as smaller schools are often in remote areas and would take considerable resources to reach. As it stands, the ECE already requires about 40,000 field workers each year. Schools teaching in indigenous languages are covered under a separate testing schedule. In total, 14,000 – 19,000 primary public schools participated per year (48% to 65% of all primary schools; see Figure A.4a). About 20% – 43% of schools were exempt under the minimum enrollment or language criteria. The remaining schools (between 2% and 11%) were not tested due to logistical problems. The coverage of the test was nonetheless very broad: since the smallest schools were excluded by definition, and since schools in native language tend to have modest enrollments, between 80% and 89% of all second graders in the country were tested in the ECE in a given year (Appendix Figure A.4b).

2.2.3 Estimation Sample

In Appendix Table A.2 we account for the different steps we took to build our estimation sample. Columns 1-3 characterize all public schools that appear in the CE during our study period. We characterize these schools using information from 2007. When information for 2007 is not available, we use the earliest available data for each school. Overall, there were 31,111 schools that reported

[17] The ECE was not conducted in 2017 because of significant school absences due to El Niño and teacher strikes. Starting in 2018, the test has been given only to a very small set of schools (around 1.7% of the schools tested in 2016) and a new sample of schools is chosen each year. Given that our identification strategy (described in Section 3.1) includes school fixed effects, we can only use ECE scores through 2016.

information in at least one of the rounds of the CE during this period. Our empirical strategies exploit the timing of internet connection within schools. Therefore, in our sample we include all schools that initially installed internet in 2007 or later and all schools that remain unconnected by the end of our CE sample period (2020). We exclude all schools that were connected prior to 2007 (according to the 2006 CE), as we are unable to determine the initial year of access for these schools. This only excludes a small share of public schools, as only 1,366 (4.3%) were internet-equipped by 2007. This leaves us with 29,745 schools, over 95% of all public primary schools in Peru (Appendix Table A.2, columns 2 and 3). Though 18,435 schools (59.3%) remained unconnected by 2020, 11,310 schools (36.4%) gained access between 2007 and 2020. The sharp expansion in internet connectivity during this period allows us to form insights about the effect of internet access using variation from a large number of schools.

We then merge this information with annual test scores from the ECE, resulting in 24,748 matched schools (about 79.5% of all public schools; Appendix Table A.2, columns 5 and 6). The sample of schools in columns 2-3 (all schools who gained access to internet between 2007-2020 or did not have access by 2020) and columns 5-6 (for which we observe average scores) appear to be nearly identical on observable characteristics.

All in all, our estimation sample includes 23,318 schools that were tested in the ECE in our sample period (2007-2016), that gained internet between 2007 and 2020 or remain unconnected as of 2020, and that are contain all of the necessary covariate information for our main specification (described in Section 3.1). The final column of Appendix Table A.2 gives the summary statistics for the estimation sample.

While schools that remained without connection to the internet by 2020 aid us mostly in the identification of calendar year effects and covariate coefficients, the main source of variation in our event study comes from schools that adopted internet between 2007 and 2020. Importantly, the data suggest that schools that gained internet from 2007-2020 generally fall "between" the early adopters (who received internet before 2007) and non-adopters in various measures of school quality. Namely, early adopters appear to be schools with higher performance, larger enrollment, and better infrastructure and educational inputs (e.g., piped water, libraries, administrative offices, teachers, classrooms, computers, and textbooks). Conversely, non-adopters systematically appear worse in these areas. Thus, the adopters that provide the variation to identify the effect of internet access focuses neither on the best nor on the worst performing schools.

Within schools that gained access to internet in the study period, we note considerable variation in the timing of access for our analysis. This allows us to implement the event study approach described in Section 3.1. In Appendix Figures A.5 and A.6, we plot each "treatment cohort's" average math performance over time, which we normalize relative to year of internet access. For reference, Appendix Figures A.7a and A.7b represent the performance of schools that gained access prior to 2007 (the start of our sample period) and Appendix Figures A.7c and A.7d similarly display the scores over (calendar) time for schools that had not gained access to internet by 2020 (the end of our CE sample period).[18] Generally, schools that connected later or remained unconnected by 2020 exhibit lower average test scores, indicating that variation in internet access *across* schools is not random.

However, *within* a cohort of schools becoming connected in a given year (2007-2020), there do not appear to be systematic trends in scores prior to internet access. This suggests that within cohorts of treated schools, the timing of access is unrelated to test score trends on average. Furthermore, Appendix Figures A.7a and A.7b suggest that performance gains among treated schools are very small initially but grow over the medium term. In contrast, the relative performance of schools that had not been connected to the internet by 2020 declines over the period of analysis (Appendix Figures A.7c and A.7d). Thus, if we regard the trends in scores in this group as the counterfactual for performance in the absence of internet connections, the implied effects of internet on test scores are larger than the raw trends in Appendix Figures A.7a and A.7b suggest.

3 Empirical strategies and results

3.1 Event study specification

In order to analyze dynamic impacts of internet access over time, we estimate the following event study specification:

$$Y_{isr} = \sum_{t=-3}^{6+} \beta_t \mathbf{1}\left\{E_{sr} = t\right\} + \gamma X_{isr} + \alpha_s + \theta_{sr} + \varepsilon_{isr}, \quad t \neq -1 \tag{1}$$

Our primary outcomes of interest are standardized math and reading scores for second grade student i in school s observed in year r (Y_{isr}). Scores are normalized across the universe of Peruvian public schools within each year. α_s are school fixed effects, which capture all time-invariant observed and unobserved school-level determinants of performance.[19] X_{ir} is a set of individual and time-varying school characteristics that includes gender, class size, indicator variables for the number of second grade classes within a school, total school enrollment, number of second grade students scheduled for testing, facilities (piped water, library, administrative offices), and resources per student (classrooms, computers and teachers). It is important to control for school size relative to other schools within the UGEL because *Plan Huascarán* explicitly prioritized the schools with the highest enrollment within each UGEL, meaning that enrollment ranking within UGEL is highly predictive of access to internet. Larger schools are not only likely to differ systematically from small schools in level terms, they are also likely to experience different time-variant shocks to test scores.

[18] We cannot calculate event time for schools that gained access prior to 2007 or were unconnected in 2020 because we do not know the initial year of internet installation for these groups.

[19] Recall that the ECE is given only to second grade students each year, so we observe each student's scores only once (in second grade) though we have a panel of schools. Because of this data structure, we cannot include student fixed effects.

In light of this, our baseline specification includes year fixed effects that are specific to terciles of baseline enrollment within each UGEL, θ_{sr}.[20]

Let I_s denote the year in which school s gains internet connection (the first year in the dataset in which s reports internet access in the CE). E_{isr} represents time relative to internet access for each school; specifically, $E_{sr} = r - I_s$. Each of the event study dummy variables is set to zero for all schools that remain unconnected by 2020. We include these non-adopters in our estimation sample to help identify γ and the calendar year effects.[21] The coefficients on the set of event study dummy variables β_t capture the path of test scores relative to the year before a school receives internet access (i.e., relative to $t = -1$).[22] One important feature in the timing of the two datasets we use is that CE reports internet access in the beginning of the school year, while the ECE is a year-end test. Any school that installs internet after the CE (April-July) does not report internet access until the following calendar year. If internet installation occurs before the ECE exams (end of November - December), students are exposed to internet access during at least part of the year *prior* to reporting initial access in the CE. Therefore in merging internet information from the CE to test scores from the ECE, we match test scores from the ECE to the internet status in the CE of the following calendar year. This means that some schools acquire internet access in $t = 0$ (if installation occurred *before* submitting the CE information) while others acquire it in $t = 1$ (if installation occurred *after* submitting the CE information). Unfortunately, school-level information is not available for either the month of installation or completion of the CE, and so we are unable to tell how many schools receive internet in $t = 0$ versus $t = 1$. Thus, in interpreting estimates of β_t it is important to keep in mind that $t = 0$ is a partially treated year for some schools and a pre-treatment year for others, while $t = 1$ is a partially treated year for some schools and a (fully) treated year for others.

By exploiting variation in the timing of internet access *within* schools (as well as additionally controlling for aggregate year effects and a set of time-varying characteristics), we aim to identify the effects of internet access separately from potential confounders that are fixed at the school level. We consider this a refinement over Hopkins (2014) — who also examines the relationship between internet access and test performance in Peru — but compares internet-connected schools to non-connected schools. We use the event study framework to examine both pre-treatment trends and dynamic effects in a non-parametric fashion from three periods before to more than six periods after gaining internet access. We choose this event window prior to internet access because we have internet installation data from the CE until 2020, allowing us to correctly assign event time only as

[20]We use baseline enrollment to avoid any potentially endogenous changes in enrollment with respect to school internet access. Since the tercile is based on school enrollment in the first year a school is observed, it is time invariant. Results that include year effects that are specific to quintiles of enrollment are very similar (available upon request).

[21]Including non-adopters also avoids the multicollinearity problem identified in Borusyak et al. (2021).

[22]Recent work (for example, Goodman-Bacon (forthcoming)) has illustrated issues with estimating treatment effects using a two-way fixed effects model with staggered timing. In light of this, we show our results are robust to alternative estimation strategies in Section 4.

far back as $t = -3$ for all schools (given that our final year of test score data is 2016). We estimate coefficients for each year relative to internet access up to $t = 5$. Because the number of schools in our sample that had been connected to internet for more than five years declines substantially, we group all periods $t \geq 6$ for our estimation purposes. Standard errors are clustered at the school level to allow for arbitrary serial correlation in ε_{sr}.

Figures 2a and 2b display the results of estimating Equation 1 on our main sample when the outcomes considered are standardized math and reading scores, respectively. The full set of coefficient estimates for the event study dummy variables are reported in Appendix Table A.3. We find that, prior to school internet access ($t < 0$), math and reading performance was roughly constant from year to year. Importantly, there are no apparent trends in test scores prior to internet access, indicating that the timing of internet access within schools is unrelated to pre-trends in student performance. In particular, we rule out the case in which internet installation is budgeted endogenously as a reward for steadily improving test performance. While relative student performance rises in all years following initial connectivity in both math and reading, immediate gains are small in magnitude: 0.028 and 0.018 standard deviations in $t = 1$ (the first year of partial or full access to internet in our schools) for math and reading, respectively. However, test score growth following internet access is steady in both subjects. By year 5, scores are 0.111 standard deviations higher for math and 0.063 standard deviations higher for reading relative to the year prior to internet installation. Beyond five years of internet access, this effect increases to 0.143 and 0.078 for math and reading, respectively.

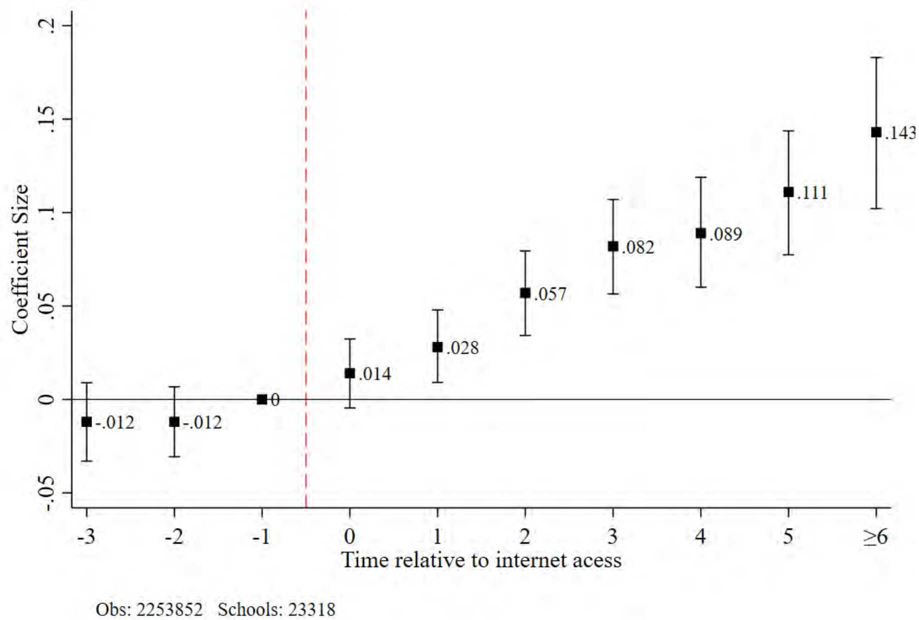
On the surface, this finding stands in contrast to other studies in developing countries that find limited or no impacts of ICTs on test scores. However, our short run estimates are in fact similar in magnitude to those from several studies of ICTs introducing hardware, such as laptops and tablets[23] —which range from 0.052 to 0.110 standard deviations, 5 to 22 months post-intervention (Bet et al. 2014, Barrera-Osorio and Linden 2009, Cristia et al. 2017, Beuermann et al. 2015 Mo et al. 2013) — though these other studies are unable to statistically distinguish estimates from zero.[24] Results from Figures 2a and 2b suggest that though classroom internet is beneficial to learning, improvement in the initial years post-intervention is small. The majority of the studies in this literature focus on impacts within the first 18 months post intervention, an early stage in which impacts may not be statistically detectable in smaller samples. The fact that our estimates

[23]To avoid confounding the impact of internet with that of access to hardware alone, all our regressions control for the number of computers per student. Hence, the estimated effects of internet access account for differential access to computers. We discuss the relationship between the effects of internet access and computer availability more explicitly in Section 4.

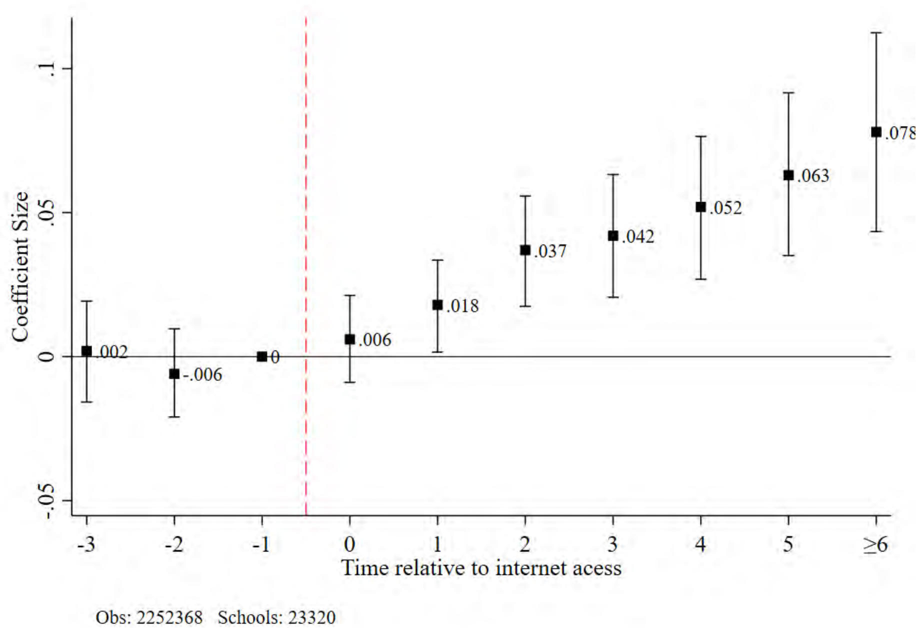
[24]For example, Bet et al. 2014 find that the point estimate for the effect of technology access on math scores is 0.069 standard deviations, which is larger than the effect we estimate one year after internet installation (0.028 standard deviations). However, the standard error in Bet et al. 2014 is 0.056, meaning that the estimate is not statistically significant at traditional levels. In our case, the standard error is 0.01, allowing us to statistically rule out a zero effect even though the point estimate is considerably smaller. Other studies (Angrist and Lavy 2002, Meza-Cordero 2017, Sharma 2014) find negative — though not always statistically significant — effects of hardware introduction on test scores.

Figure 2: Impact of Internet Access on Test Scores

(a) Standardized Math Scores



(b) Standardized Reading Scores



The above figures plot the coefficients and 95% confidence intervals from estimating equation 1. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Coefficients capture the increase in test scores relative to the year before internet installation ($t = -1$). Note that due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ and some receive it in $t = 1$. For more details, see Section 3.1. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school.

grow over time, at least through the medium-term, is also consistent with two other longer-term studies of ICTs in education — which have also supported the need for an adaptation period to fully utilize new technologies (Hull and Duch 2016; Mo et al. 2015).

In the medium-run (3-5 years), the increase in test scores is substantial, though somewhat smaller than those typically found in evaluations of computer assisted learning and related interventions (0.18 to 0.59 standard deviations); e.g., see Bando et al. 2016, Banerjee et al. 2007, Beg et al. (2019), Carrillo et al. 2010, He et al. 2008, Linden 2008, Muralidharan et al. 2019. Our smaller albeit statistically significant estimates may owe partly to the fact that introduction of internet into Peruvian schools was unaccompanied by any particular pedagogical software service, or complementary interventions.

Interestingly, the dynamic path of tests scores appears to be linear with respect to school exposure to internet for both subjects. This aligns with recent evidence from Jackson and Mackevicius (2021), who find that the effects of increased school spending on educational attainment grow linearly with years of exposure.

In comparing our estimates to the existing literature, it is important to reiterate that we identify school-level - as opposed to student-level - dynamics. Thus the increase in test scores that we observe over time after internet installation are reflection of how schools adapt to a new technology and become more efficient at improving test scores over time. In this way, impacts beyond the short run in our context are not directly comparable to estimates one might expect from studies that follow individuals over time.

3.2 Trend break specification

Though the shape of Figures 2a and 2b suggest a steadily increasing effect of internet access on test scores over time, it does not explicitly test for a break in the trajectory of scores at the time of internet installation. To do so, we estimate a linear trend break specification as follows:

$$\begin{aligned} Y_{isr} &= \phi_1 \text{Post-internet Access}_{ir} + \phi_2 \text{Event Time}_{ir} \\ &\quad + \phi_3 \text{Post-internet Access}_{ir} \times \text{Event Time}_{ir} + \gamma X_{isr} + \alpha_s + \theta_{sr} + \varepsilon_{isr} \end{aligned} \tag{2}$$

Here, Post-internet Access is a dummy variable that is equal to one in all periods after internet installation ($t \geq 0$). Event Time is a linear term for time relative to the year prior to access, $t = -1$. The control set (X_{isr}) is otherwise identical to that described in Section 3.1. In this specification, ϕ_1 captures the level shift in test scores in response to internet access; ϕ_3 represents the change in the linear time trend in math scores after schools gain internet access; and ϕ_2 accounts for any pre-existing linear trend. Based on the results in Section 3.1, it is unlikely that there are any existing pre-trends. However, one benefit of this specification is that even in the presence of any linear pre-trends in test scores, ϕ_3 measures the impact of internet access on the growth in test scores *apart* from any such trends.

Table 1: Internet Access & Test Scores: Trend Break Results

	Dependent Variable: Standardized Test Score	
	Math	Reading
	(1)	(2)
Post-internet Access	0.014	0.015
	(0.011)	(0.009)
Post-internet Access X Event Time	0.013**	0.011**
	(0.005)	(0.004)
Event Time	0.008	-0.000
	(0.005)	(0.005)
p-value of Test of Joint Significance for Post and PostXEventTime:	0.0520	0.0620
Observations	2,253,852	2,252,368
Number of schools	23,318	23,320

Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Coefficients capture the increase in test scores relative to the year before internet installation ($t = -1$). Note that due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ and some receive it in $t = 1$. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school. Significance levels denoted by: *** p< 0.01, ** p< 0.05, * p< 0.1.

Results from estimating equation 2 are displayed in Table 1. Though we observe a level shift in both math and reading scores upon installation (0.014 and 0.015 standard deviations, respectively), the immediate effect is small and we are unable to statistically distinguish these effects from zero. On the other hand, the year on year gain in test scores in both math and reading are significant and the magnitude is meaningful (0.013 and 0.011 standard deviations, respectively). This stands in contrast to the estimated pre-trends, which are close to zero and not statistically significant. We also find that the level shift (captured by ϕ_1) and the trend break (captured by ϕ_3) are jointly significant at the 90% level for both subjects.

4 Robustness checks

In this setting, identification of the impact of school-based internet access on student performance may be confounded if the timing of internet access is non-random within schools. However, conditional on school fixed effects, year effects, and a set of school-specific time varying controls, our analysis in Section 3 offers no signs that the timing of internet access relates to prior test performance. In this section, we address several other potential challenges to identification, namely endogenous changes in sample composition in terms of both students and schools (including non-random attrition), concurrent changes in school resources, pre-existing trends that might differ by geographic area, administrative unit, or prior student performance, and other potentially nonexogenous sources of variation in the timing of internet access.[25] Finally, we show that our results are robust to an alternative estimation method that addresses many of the concerns raised by the recent literature on two-way fixed effects estimation with staggered timing of treatment.

4.1 Changes in sample composition

As we use students from an unbalanced panel of schools (schools are included when they participate in the ECE) observed over a limited window of time (2007-2016, years for which ECE data is available), it is possible that our estimated treatment effects to some extent reflect changes in sample composition. Namely, identification of pre-trends and treatment effects might rely on students from different samples of schools. In Section 4.1.1, we find that our estimates are not contaminated by this issue or by school-level attrition. Additionally, we consider also that the composition of students *within* schools may change in response to internet access. For instance, internet access may attract a different pool of students to a school (either from other schools or from non-school activities). Section 4.1.2 presents evidence that this manner of endogenous student sorting does not occur in our sample.

4.1.1 Unbalanced panel and attrition

Since some schools in our main sample are observed only prior to internet installation while others are observed only after, it is possible that the pre-internet coefficients and trends are identified from a different set of schools than those identifying the post-internet treatment effects. Unfortunately, the likelihood of observing a school's pre-internet years versus post-internet years is furthermore naturally influenced by the date of internet installation (in the extreme case, we of course do not observe the pre-period of any school that installed internet in 2007). If these schools in actuality experienced non-zero pre-trends, our estimates will not take these into account. On the other hand, schools that installed internet in 2016 are not observed post-internet and can only be used to identify pre-internet coefficients/trends. Even though these schools may show no pre-trends, it is possible that they could potentially go on to experience zero (unobserved) effects of internet access.[26]

[25] We also perform an exact randomization exercise as follows. We randomly reassign schools' initial year of internet access (including non-adoption) in our sample. We maintain the actual distribution of installation years, thus ensuring that the sample sizes of each of our "treatment cohorts" matches our data. We then use the randomly assigned installation years to generate false event study dummies, which we use to re-estimate equation 1. We repeat this process 500 times and plot the median, 5th percentile, and 95th percentile of the resulting coefficients in Appendix Figure A.8. For the pre-internet periods, the coefficients from the baseline specification are close to the median of the coefficients from the placebo exercise (essentially zero), whereas the post-internet coefficients fall well outside the 5th and 95th percentile of the placebo coefficients. We take this as additional evidence that our coefficients yield statistically significant estimates of the impact of internet access, and that the inference in our baseline specification is appropriate.

[26] Relatedly, we find no evidence that the effects of internet access vary across schools that adopt internet earlier versus later. To explore this possibility, we run our trend break specification allowing for heterogeneous effects across early and late adopters (split by the median year of internet adoption). We find that the trend break estimates across

Figures 3a and 3b and columns 2 and 5 of Appendix Table A.5 suggest that our main findings are not driven by this school-level sample composition issues. Specifically, we restrict the sample to students in schools that can appear at least twice prior to and twice following internet installation (i.e., schools that installed internet between 2009 and 2015), for which we can observe *both* pre-trends and treatment effects within the same school. Imposing this restriction drops about 2,000 schools from the sample, but the remaining sample is comparable along many observable dimensions (see Appendix Table A.4). We continue to include all non-adopters (schools without access by 2020) to identify year effects. In this sample, we find no statistically significant trends in performance prior to internet access, and the estimated effects are similar in magnitude and statistically significant and that show similar dynamics as those using the full sample.

School-level attrition from the panel may pose another compositional issue. Overall attrition in our sample is 40.7% (missing school-year observations). School-level attrition can happen for several reasons. First, as mentioned in Section 2.2.1, only schools with at least 5 second grade students and in which the language of instruction is Spanish are required to administer the ECE exam. Therefore, observations will be missing when schools fall below the threshold of 5 students (or which switch to an instructional language other than Spanish). Some schools might have experienced permanent reductions in their second grade enrollment (and drop from the sample at some point), while others might alternately meet and fall below the ECE enrollment threshold from year to year (e.g., a school might have five second graders during a year and only four during the next year). Appendix Table A.6 shows that over two thirds of overall attrition is likely due to a school dropping below the enrollment threshold.[27] The remaining attrition is either due to missing ECE scores for another reason or missing CE (covariate) information. Only a very small portion of attrition is due to school closures.

All in all, a considerable proportion of attrited observations come from missing scores in the ECE (82.8% of attrited observations). Because the ECE is administered and graded by the Ministry of Education, schools cannot selectively decide whether to report their ECE scores or not. However, it is possible that — if changes in enrollment are negatively correlated with changes in school quality[28] — certain schools might fall below the ECE enrollment threshold in years in which they might have potentially performed worse. As suggested by our results so far, if internet does improve school performance, then we would be more likely to observe ECE scores for schools after they gain internet connection. To formally rule out this possibility, we estimate equations 1 and 2 on the restricted sample of schools with ECE scores and the full set of covariates (i.e., the sample of non-attritors) in *every* calendar year 2007-2016. The results are displayed in Figures 4a and 4b and columns 3 and 6 of Appendix Table A.5. Even among schools that are observed in every year, we

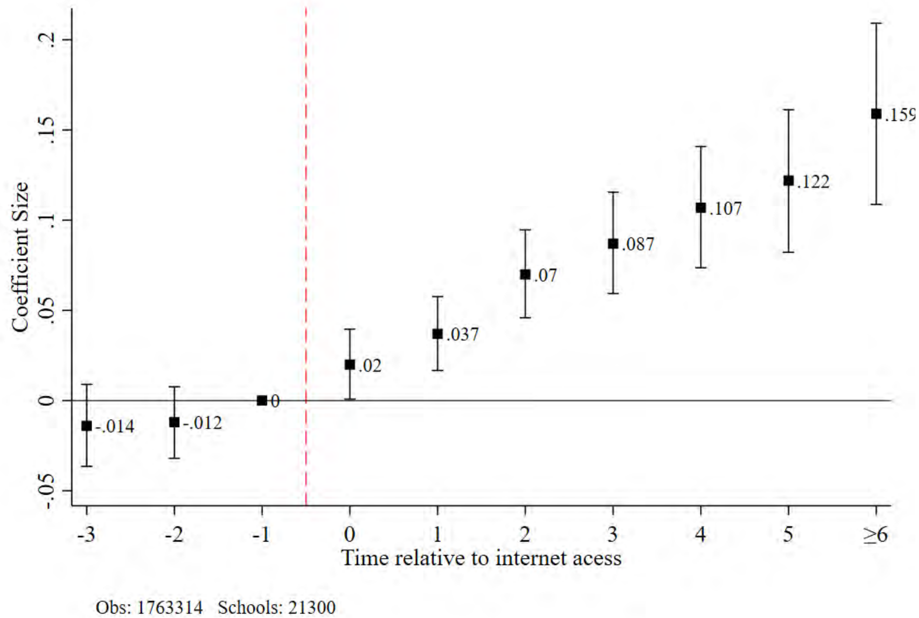
the two samples are not statistically significantly different from each other for either subject.

[27] About half of attrition is due to schools having fewer than 5 second grade students, and an additional 15% is explained by having enrollment "near" the threshold (defined as having 5-8 second grade students).

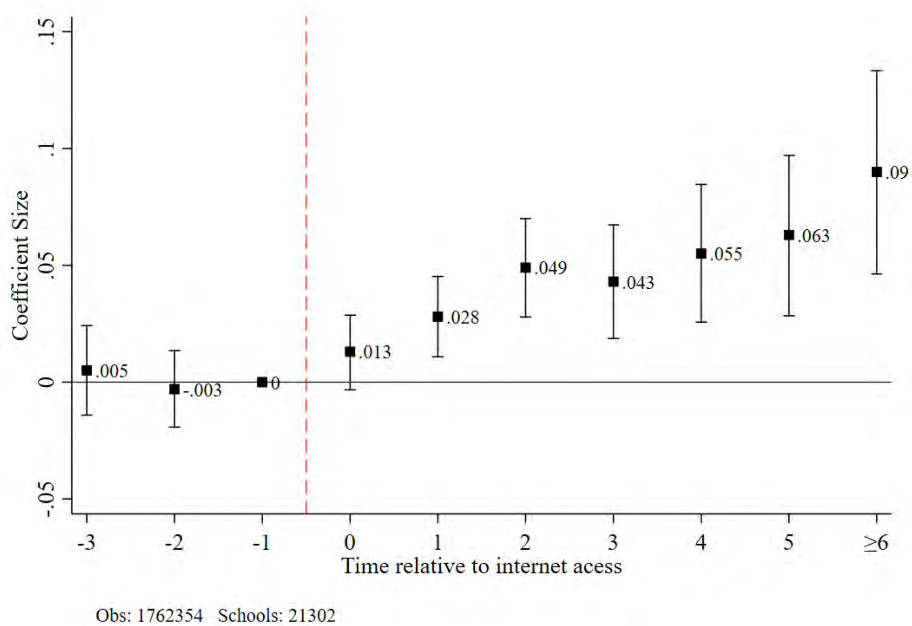
[28] In Section 4.1.2, we analyze the possibility that internet access increases enrollment.

Figure 3: Effect of Internet Access in Schools Connected between 2009 and 2015: Event Study Results

(a) Standardized Math Scores



(b) Standardized Reading Scores



The sample includes all grade 2 students from all public schools that gain internet between 2009 and 2015 (i.e. observed at least twice prior to and twice after installing internet) or remain unconnected by 2020. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ and some receive it in $t = 1$. Control variables include sex, class size, indicator variabes for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Standard errors are clustered by school.

find that the estimated effects of internet access are sizable and grow over time. The trend break results in columns 3 and 6 of Appendix Table A.5 show that the estimated yearly gain in test scores due to internet access is similar to the baseline results (columns 1 and 3), though the trend break coefficients are not statistically significant. It is worth noting that the sample observed in all years is somewhat similar to the baseline sample along many observable dimensions, albeit higher achieving and (somewhat naturally) larger in terms of enrollment (Appendix Table A.4).[29][30]

4.1.2 Student composition and endogenous sorting

Another related issue is that the composition of students *within* schools may change in response to internet access. A priori, it is hard to tell the direction of the bias that this would entail. For instance, internet access as a signal of increased school quality may imply countervailing possibilities with regard to student composition. Parents who would otherwise not have sent their kids to school might decide to enroll their children in a school connected to internet. If internet connectivity attracts students that would have otherwise performed poorly, then our estimates of treatment effects are likely conservative. Alternatively, motivated parents seeking learning opportunities for their children may decide to transfer students from schools without internet to schools that gained connectivity. If these new students are better achievers on average, then our findings of positive treatment effects may owe to upward bias from changes in student composition.

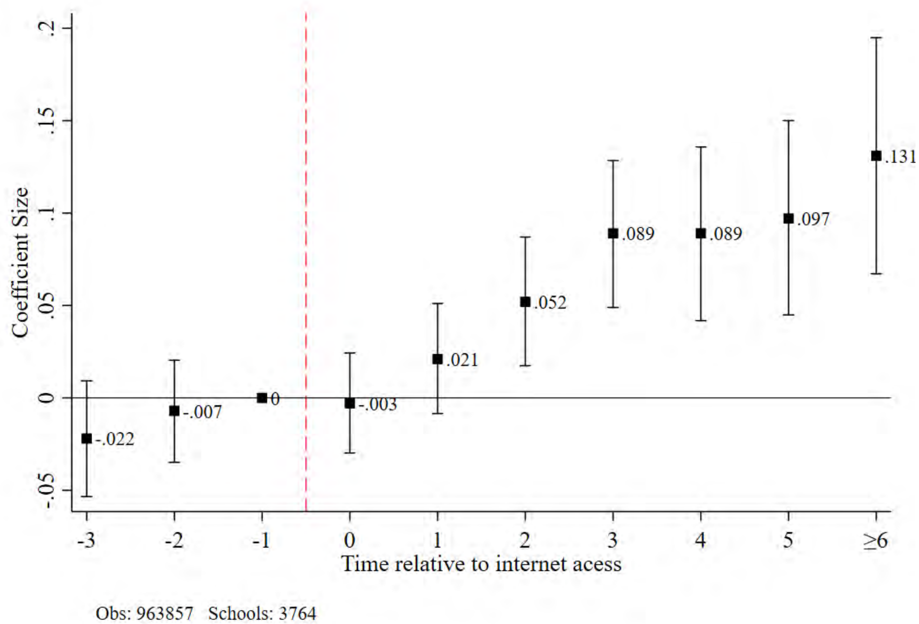
Overall, it does not appear plausible that an influx of high-achieving transfers or re-entrants explains the performance gains in our main results. To rule out this possibility, we first analyze the response of grade 2 transfers, re-entry, and total enrollment to internet access in columns 1-3 of Table 2. Transfers are students enrolled in the current year who were enrolled in a different school in the previous year. Re-entrants are students that are currently enrolled but who were not enrolled in any school during the previous year (i.e., dropouts who come back to school). If anything, it appears that there are fewer transfers over time after a school becomes connected to the internet,

[29] We also conduct the analysis using Lee Bounds to show that attrition does not likely explain our results. Define q_t as the proportion of observed (non-attrited) observations in time t. We identify the event time with the highest attrition rate and calculate the fraction of observed outcomes for that event time (q_{min}). We then calculate $Q_t = \frac{q_t - q_{min}}{q_t}$ for each event time t and restrict the sample to observations that have test scores within the percentiles (Q_t, $1 - Q_t$) for each t. We then re-run our trend break specification (equation 2) on this restricted sample. Even after trimming the sample in this way, we find that the estimated level shifts and trend breaks are of similar magnitude to the baseline results and are significant at the 5% level, with the exception of the trend break for math. Results available upon request.

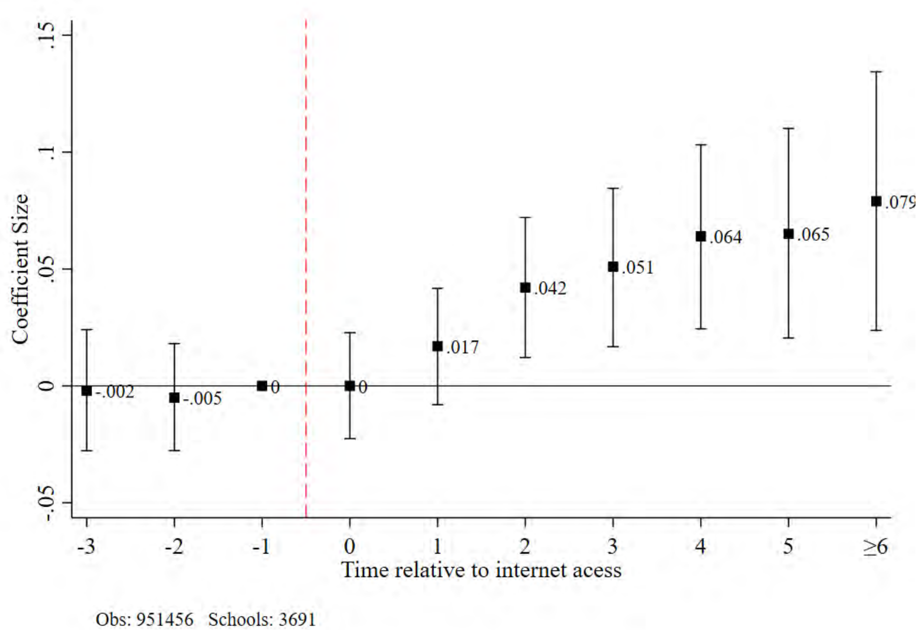
[30] We also show that our results are not being driven by the use of an unbalanced panel of schools by showing robustness to a very restricted sample that includes only students from schools that appear in all event years in Appendix Figure A.9. Note that with 10 years of data and a 10-period event window, we have no variation in internet timing if we require all schools to appear in all event years. Thus we shorten the event window to allow for this sample restriction. The (post-internet) point estimates are all as large or larger in the fully balanced sample as in the baseline sample, with only a single exception ($t = 0$ for math). Note that the fully balanced sample is a very small and select sample; it includes only around 2100 schools, which is less than 9% of our baseline sample of schools. This is because, in order to observe each school in each event year, we can only include 3 "treatment cohorts" of schools (whereas in our full sample we use 14 cohorts). Therefore the estimates from the fully balanced sample are considerably noisier and that our our estimates are only relevant for a limited set of schools.

Figure 4: Effect of Internet Access in the Sample of Schools Observed in All Calendar Years 2007-2014: Event Study Results

(a) Standardized Math Scores



(b) Standardized Reading Scores



The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020*and* that are observed for the entire sample period, i.e. for each year 2007-2016 (non-attritors). Coefficients capture the increase in test scores relative to the year prior to a school receiving internet access ($t = -1$). Note that due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ while some receive it in $t = 1$. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Standard errors are clustered by school.

though the point estimate is small (column 1). There are no apparent effects of internet access on grade 2 re-entry or enrollment (columns 2 and 3).

Nevertheless, even if total enrollment remains relatively unchanged in response to internet connectivity, the makeup of the students that take the test could still change. While students might be enrolled in school, they only take the ECE if they attend school on the day in which the test is administered. If internet availability induces attendance, then a different set of students will be present to take the test after a school gains internet access. We investigate this possibility in columns 4 and 5 of Table 2. Column 4 examines the effect of internet on the proportion of enrolled students that actually take the ECE. It does not appear that students become more likely to take the test following the installation of internet. These results are consistent with Cristia et al. (2017) and He et al. (2008), who find that neither hardware nor CAI/CAL interventions has any significant effects on attendance. In column 5, we further explore whether student background changes in response to internet access. The only information on the background of students in the *Censo Escolar* is the proportion of native Spanish speakers enrolled.[31] To the extent that native language captures student background, it does not appear that internet access attracts more advantaged students. In Appendix Figure A.11, we show the results of the event study analysis when considering the outcomes in Table 2. Consistent with the trend break results in Table 2, we find no statistically significant patterns in the effects on these outcomes. Overall, the evidence in Table 2 and Appendix Figure A.11 does not seem to indicate that endogenous student sorting drives our estimated impacts of internet access.

4.2 Concurrent changes in school resources

Timing of internet access may also possibly correlate with changes in other school resources.[32] For example, it might be that internet provision is bundled with other inputs in a multifaceted approach to improve quality of schooling.[33] If this is the case, the improvement of students' performance that we observe might be due to increases in these other resources. For the most part, we do not find that the timing of internet access is correlated with increases in other observable inputs (Table 3). Classrooms, textbooks, overall teachers (excluding computer teachers, which are separately discussed in Section 5.3.1), and qualified teachers (those with a pedagogical or university degree) per student actually *fall* slightly after internet access (though the point estimate is very small; columns 1-4). However, in column 5, we note a positive and non-negligible level increase

[31] The proportion of Spanish-speaking students is positively related to higher test scores, even after conditioning on state-year fixed effects and the controls listed in Section 3.1.

[32] All of our specifications include school and UGEL-specific enrollment tercile-year effects as well as time-varying school characteristics. We show that our results do not depend on the inclusion of covariates in Appendix Table A.7.

[33] It could also be the case that internet access at schools is correlated with alternative sources of internet. For example, students who gain internet access at school may already have internet connections at home or via cyber-cafes. However, we find that only 15% (29%) of students who used internet at school also use it at cyber-cafes (or at home) according to the 2016 Peruvian National Household Survey (ENAHO). Additionally, our results are unchanged if we include a control for whether the town nearest the school has a cyber cafe.

Table 2: Effect of Internet Access on Transfers, Re-entry, Test Taking, and Student Composition

	Grade 2 Transfers (1)	Grade 2 Re-entry (2)	Grade 2 Enrollment (3)	Proportion of Enrolled Students that Took Test (4)	Proportion of Native Spanish Speakers in Grade 2 (5)
Post-internet Access	0.140	0.052	0.394	0.005	-0.003
	(0.251)	(0.064)	(0.350)	(0.004)	(0.003)
Post-internet Access X Event Time	-0.230**	-0.008	-0.109	-0.003	-0.001
	(0.106)	(0.032)	(0.173)	(0.002)	(0.001)
Event Time	0.139	-0.020	-0.011	0.000	0.000
	(0.100)	(0.033)	(0.153)	(0.002)	(0.002)
p-value of Test of Joint Significance for Post and PostXEventTime:	0.004	0.305	0.166	0.189	0.620
Pre-internet mean of dep. variable	2.783	0.417	33.18	0.894	0.845
Observations	2,253,852	2,253,852	2,253,852	2,253,218	2,253,218

Transfers are students enrolled in the current year who were enrolled in a different school in the previous year. Re-entrants are students that are currently enrolled but who were not enrolled in the previous year. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school. Post-internet access is a dummy variable for whether a school has gained internet access (i.e. $t >= 0$). Event time is years relative to internet access. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Columns 1-3 controls for both second grade enrollment and enrollment in other grades separately, not including transfers or re-entrants when specified as an outcome variable. Column 4 also controls for the number of second grade students scheduled to take the test.

in computing resources (which includes OLPC laptops) at the time of internet installation; on the other hand, the estimate of the trend break is negative and significant.

Table 3: Effect of Internet Access on School Resources

	Classrooms per Student (1)	Textbooks per Student (2)	(Non-computer) Teachers per Student (3)	Qual. Teachers per Student (4)	Computers per Student (5)
Post-internet Access	-0.001	0.021	-0.0004	-0.0003	0.008***
	(0.001)	(0.074)	(0.000)	(0.000)	(0.002)
Post-internet Access	-0.000	-0.029	-0.0003	-0.0004**	-0.002**
X Event Time	(0.000)	(0.033)	(0.000)	(0.000)	(0.001)
p-value of Test of Joint Significance for Post and PostXEventTime:	0.326	0.242	0.274	0.108	0.000
Pre-internet mean of dep. variable	0.0667	3.878	0.0538	0.0413	0.118
Observations	2,253,852	1,703,334	2,253,852	2,253,852	2,253,852
Number of schools	23,318	21,866	23,318	23,318	23,318

In column 2, the dependent variable is the number of 2nd grade textbooks per 2nd grade student. Textbook information is not available for 2012. In column 3, the number of teachers per students excludes computer teachers (discussed separately in Section 5.3.1). In column 4, the dependent variable is the number of teachers with a pedagogical or university degree per student. In column 5, the dependent variable is the number of instructional computers per student. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school. Post-internet access is a dummy variable for whether a school has gained internet access (i.e. $t >= 0$). Event time is years relative to internet access. Control variables include (excluding the dependent variable) sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects.
Significance levels denoted by: *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$.

By definition, internet connectivity is complementary to computer access (i.e., generally students cannot use the internet *without* computers). However, students might benefit from computers without access to internet (e.g., using preloaded software and resources installed from flash drives / DVDs, etc.). Joint increases in computer and internet availability might thus possibly imply that our estimates capture (at least partially) the effect of computers themselves apart from internet access. To better understand the potential for increases in these resources to confound our estimates of the impact of internet access, we begin by showing that non-internet enabled computers alone do not seem to improve test scores. In Appendix Table A.8, we regress test scores on computers per student for the sample of schools that do not gain access to the internet during our sample period.[34] In each successive column, we add in lags of computers per student to allow for dynamic effects of these resources on student performance. For schools than remain unconnected by 2020, we find few statistically significant positive correlations between computers and test scores and those

[34]These regressions include school fixed effects, UGEL-specific enrollment tercile by year fixed effects, and all of the controls from our baseline specification (equation 1).

correlations are only for the 2-year lag and are small in magnitude. Though we control for many confounding factors (by including covariates, school effects, and UGEL-specific enrollment tercile by year fixed effects), we ultimately lack a source of purely exogenous variation in school computers. We admit that these results are not definitive proof that non-internet connected computers fail to raise test scores. However, we believe these results are nonetheless informative. Moreover, these results are also well-supported by the existing literature (most often based on RCTs, where internal validity is not a concern), which generally finds no significant contemporaneous impact of computers alone on student performance (Bet et al. 2014, Barrera-Osorio and Linden 2009, Cristia et al. 2017, Beuermann et al. 2015 Mo et al. 2013).[35]

Next, we perform some back of the envelope calculations. These calculations are based on the following reasoning: if internet connections increase availability of computers; how much of the performance improvements of schools that gain access to internet can be explained by their subsequent access to computers alone? By our calculations (displayed in Appendix Table A.9), we find that the increases in computers alone explain very little of the observed rise in test scores (at most, 0.91% for math and 0.64% for reading). For these calculations we use two pieces of information: (i) the impact of internet access on computer resources and (ii) the impact of computer resources alone (i.e., without connections to the internet) on student test scores. Our estimate of (i) comes from column 4 of Table 3 above. To approximate (ii), we use the most "generous" estimates from Appendix Table A.8.[36] For example, let us examine the calculations for computers in $t = 0$ (first row of bottom panel of Appendix Table A.9). Computers per student rise by 0.008 when internet is introduced (column 4, Table 3). The largest estimate of contemporaneous (non-internet connected) computers' impact on test scores is 0.016 (from column 2 of Appendix Table A.8). Based on both estimates, we would expect a test score gain of $0.008 \times 0.016 = 0.00013$ standard deviations at $t = 0$. From our event study specification, our overall estimated effect of internet access at $t = 0$ is 0.014 standard deviations. Therefore, we estimate that the increase in computers per student alone at $t = 0$ explains only about $0.00013/0.014 = 0.91\%$ of the rise in test scores that we observe in response to internet access. The calculations for other post-internet periods are more complex when the effect of resources is allowed to be dynamic. However, our most "generous" estimates indicate that, at most, acquisition of more computers explains only a small portion of the estimated gains in test scores following internet access. These calculations aside, we include

[35] We cannot use the estimates of the effect of computers directly from these papers in the back of the envelope calculations presented in Appendix Table A.9 for two reasons. First, none of these papers estimates the dynamic effect of computers (which are necessary to compare to the dynamic effects we estimate in our main specification). Second, all except Cristia et al. 2017 report overall program treatment effects (where each specific program varies across papers) rather than effects scaled by computers per student. If we use the imputed contemporaneous estimated effect of computers on test scores from Cristia et al. 2017 (which is *not* statistically significant), we find that at best, the increase in computer per student that occurs at $t = 0$ can explain only 2.8% of the estimated effect of internet access on math scores at $t = 0$. Cristia et al. 2017 estimated a negative (non-significant) impact of computers on reading scores.

[36] To get the "most generous" estimates, we take the largest positive individual estimated effect of computers for each lag across all specifications in Appendix Table A.8, regardless of significance level.

time-varying measures of these resources in our control set to ensure that the internet effects we identify are conditional on the computers available to the students. Therefore, we do not believe that our results are driven by concurrent access to computers in schools.

4.3 Differential pre-trends and other sources of endogenous internet connections

Another possibility is that access to internet is correlated with pre-existing trends in test scores. For example, districts with faster growing economies might be better able to finance internet expansions, increase public spending on education, or otherwise improve student learning. In Table 4 we show that our results are robust to allowing for an array of group-specific pre-trends. These include groups that are defined administratively (i.e., by Local Educational Management Unit, UGEL), geographically (district, the finest geographical unit we observe), and by initial academic performance (captured by pre-internet test score deciles). Including UGEL-specific pre-trends (column 2) is important for ruling out endogenous selection for internet installation at higher levels of government (e.g. if the central government allocated more internet funds to UGELs with higher test score growth, even though the official UGEL-level quotas were determined solely by student enrollment). Allowing for pre-trends that are specific to initial test performance (column 4) is also useful in ruling out possible reversion to the mean, as we allow schools with initial poor performance to be on a separate score trajectory than high performing schools. The point estimates of the level shift and trend break in test scores are all stable in terms of sign, magnitude, and statistical significance across specifications; if anything, the magnitudes and statistical significance are slightly higher when we include district- and initial scoring decile-specific pre-trends. Overall, we take the evidence in Table 4 to indicate that pre-existing trends in test scores and reversion to the mean do not confound our estimates of the effect of internet access.

Finally, we show that our results are not likely driven by schools that are potentially gaining internet access endogenously. While *Plan Huascarán* is likely to be the most important source of variation in internet access during our study period, schools whose principals become more resourceful or whose parents improve their motivation might have gained access through other sources. To account for this possibility, we perform two additional checks. First, we show that the results to excluding schools in non-complying UGELs, defined as UGELs in which more public primary schools become connected to the internet than planned under the official *Plan Huascarán*; see columns 2 and 5 of Appendix Table A.10.[37] In this restricted sample, the estimated effects of

[37] We calculate the quotas as follows: first, we obtain annual quotas per UGEL from the Ministry of Education (published in 2004). Second, we multiply the annual quota by 13 (to reflect 13 years between when the quotas were published in 2004 and the end of our sample period, 2016) and then by 0.7 (to reflect that 50% of the quota was for primary (only) schools and 20% was for integrated (primary and secondary) schools). We then define complying UGELs as those with no more connected schools than the imputed quota for 2016. This is likely to overstate the number of non-complying UGELs, as the program expanded in 2007 when it became part of Directorate General of Educational Technologies (*DIGETE*).

Table 4: Allowing for Differential Linear Pre-trends by Various Groups

	Dependent Variable: Standardized Math Scores			
	Baseline	UGEL-specific Trends	District-specific Trends	Trends by Initial Scoring Decile
	(1)	(2)	(3)	(4)
Post-internet Access	0.014	0.014	0.014	0.013
	(0.011)	(0.011)	(0.011)	(0.011)
Post-internet Access X Event Time	0.013**	0.013**	0.016***	0.020**
	(0.005)	(0.005)	(0.005)	(0.006)
p-value of Test of Joint Significance for Post and PostXEventTime:	0.0520	0.0460	0.007	0.001
Observations	2,253,852	2,253,852	2,253,852	1,852,351
Number of schools	23,318	23,318	23,318	22,024
Number of groups		219	1826	10
	Dependent Variable: Standardized Reading Scores			
	Baseline	UGEL-specific Trends	District-specific Trends	Trends by Initial Scoring Decile
	(5)	(6)	(7)	(8)
Post-internet Access	0.015	0.014	0.013	0.013
	(0.009)	(0.009)	(0.009)	(0.009)
Post-internet Access X Event Time	0.011**	0.011**	0.013***	0.020***
	(0.004)	(0.004)	(0.004)	(0.005)
p-value of Test of Joint Significance for Post and PostXEventTime:	0.0620	0.0500	0.274	0.000
Observations	2,252,368	2,252,368	2,252,368	1,851,590
Number of schools	23,318	23,318	23,318	22,028
Number of groups		222	1826	10

"Number of groups" refers to the number of groups used to create additional controls listed in each column; e.g., in col. 2 including UGEL-specific linear trends adds an additional regressor per UGEL (222). The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. In columns 4 and 8, the sample is further restricted to schools that are observed prior to internet access. Initial scoring decile is defined as the school's scoring decile based on pre-internet scoring average; the decile is calculated based on all schools that have pre-internet scores (including those that are not connected by 2020). Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Post-internet access is a dummy variable for whether a school has gained internet access. Event time is years relative to internet access. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, school fixed effects, and the additional trends indicated in each column heading. Standard errors are clustered by school. *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$.

internet are larger than in the baseline specification and statistically significant. Second, the results are also robust to excluding schools that potentially gained access to internet through means other than *Plan Huascarán*, i.e. those schools that are located in areas with some alternate form of internet access (proxied by the existence of local cyber cafes) *prior* to school-based internet access (columns 3 and 6 of Appendix Table A.10). The idea behind this restriction is that it would be very difficult for schools to install internet other than through *Plan Huascarán* in places where there is no existing internet infrastructure (and thus no cyber cafes). If anything, these results indicate that the effect of school-based internet access is slightly stronger in this sample of schools (though not statistically significant for math), whose students and teachers are very unlikely to have other means of accessing the internet.

4.4 Two-way fixed effects with staggered treatment timing

Recent work (for example, Goodman-Bacon (forthcoming)) raises some important issues with using two-way fixed effects estimation when there is staggered timing of treatment. We show that these issues do not explain our main estimates by illustrating robustness to the estimator proposed in Sun and Abraham (forthcoming). Similar to Callaway and Sant'Anna (forthcoming), the Sun and Abraham (forthcoming) method identifies cohort- (defined by installation year) and time-specific average treatment effects and then aggregates those individual treatment effects (the weighting method differs across the two papers)[38]. Doing so avoids potential biases that may be present under the standard two-way fixed effects event study methodology. The results using the Sun and Abraham (forthcoming) estimator are displayed in Figure A.10. The patterns and magnitudes of these estimates are very similar to our main estimates.

5 Explaining dynamics and identifying potential mechanisms

To better understand the mechanisms behind our main results, we now turn to alternate sources of data. We begin by examining student-level mechanisms and show that school-based internet access increases students' internet usage. Next, we use descriptive data to investigate teacher-level mechanisms, i.e., teachers' use of internet over and above students' internet usage in the classroom. We find that public primary school teachers in internet-connected schools find a variety of teaching activities to be less difficult than those without access. We also show that the effects of internet access on school performance appear somewhat larger for schools with high student to teacher ratios and - in the short run - where teachers have more qualifications, though the differences are not statistically significant. Finally, we turn to school-level mechanisms. We analyze the effects of internet installation on the hiring of specially-trained teachers and find evidence that schools make

[38] Sun and Abraham (forthcoming) weigh these cohort-specific effects by the distribution of cohorts and the relative time indicators.

complementary investments in these types of teachers following internet access.

5.1 Student-level Mechanisms

5.1.1 Does access to internet in schools increase internet use among students?

We show that internet access in school effectively translates into increased use by students. To do this, we turn to the Peruvian National Household Survey (ENAHO) because the CE data do not contain any information on individual students or internet usage. We construct a repeated cross section of all primary school students enrolled in public schools from the 2007-2016 rounds of the ENAHO. The ENAHO collects two crucial pieces of data. First, it collects individual information about internet usage during the 30-day period previous to the survey. Second, it gathers household GPS locations (village location in rural areas or the centroid of the neighborhood block in urban areas). While the ENAHO does not elicit the particular primary school that each student attends, we can infer the school that a student is most likely to attend by assigning each student enrolled in a public primary school to the nearest public primary school.[39] In the ENAHO data, we observe over 55,000 students that are likely to attend 4,693 of the 23,318 schools (20.1%) in our main estimation sample.[40]

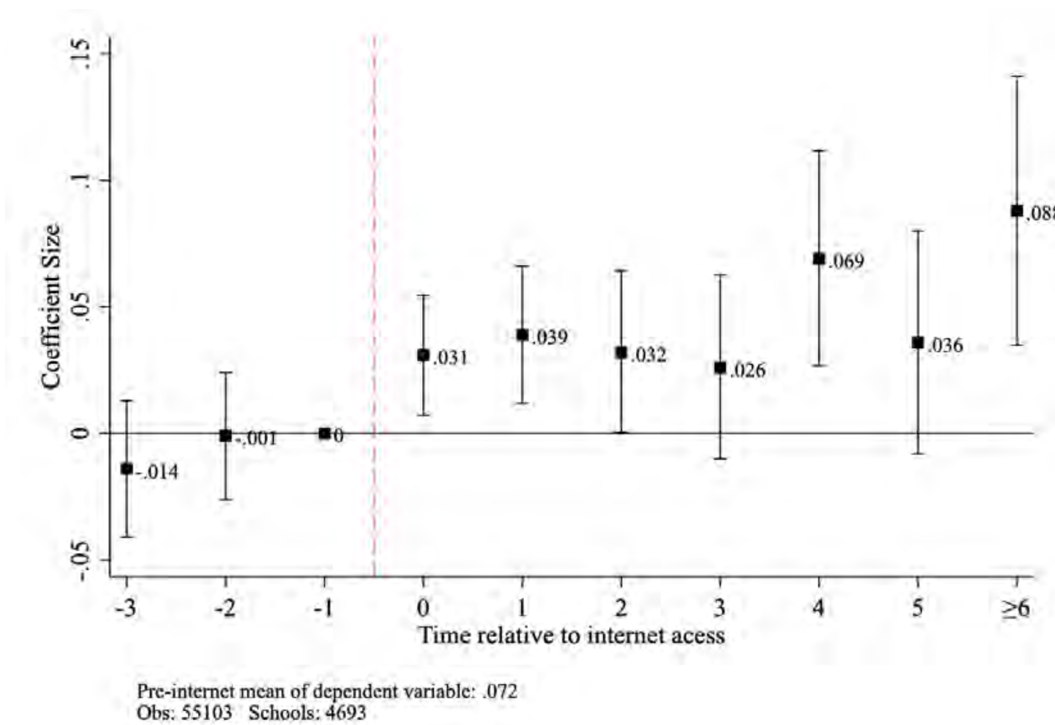
We then estimate our event study specification (Equation 1) with an indicator variable for students' internet use (from the ENAHO) as the dependent variable. We regress this variable on event time dummies based on access to internet in the student's closest school (from the CE data). We include the same set of school control variables as in Section 3.1 and add household and individual controls (student's age, grade, native language, household size, and the age, sex, and education of the household head). Importantly, our specification also includes school fixed effects. Thus, our event study coefficients capture whether the changes in internet access *within* schools increase internet usage among students likely to attend them.

We present these estimates in Figure 5. We find that students become much more likely report using the internet after their nearest public school has gained access to the internet. We see no evidence of pre-trends with respect to student internet access. After 6 or more years of in-school access, students' probability of having used the internet in the last thirty days increases by 8.8 percentage points, or 122% over the mean prior to school-based internet connections. This suggests that increased use of internet might be able to explain a portion of the improvements in student performance that we document in Section 3.

[39] In matching students to schools, we match only if there is a public primary school within 10km of the household and that is operating in the current survey year. Using this method, the median distance between a student and his/her matched school is 0.33 km. The estimates below are robust to increasing the maximum distance to 20 km or reducing it to 5 km.

[40] The reason we are only able to match 20.1% of schools is that our school data cover the universe of schools, while the household survey covers only a random sample of students each year; by its nature, it will not cover students from all schools in the ECE and CE data. In fact, the ENAHO samples about 0.4% of households in Peru in a given year. Even with this low coverage of households, we are able to match over 20% of schools.

Figure 5: Impact of Internet Access at Schools on Students' Internet Use



The above figures uses data from household surveys (ENAHO), 2007-2016. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The estimation sample includes all sampled students attending public primary schools that gain internet after 2007 or remain unconnected by 2020. Standard errors are clustered by school.

5.2 Teacher-level Mechanisms

To complement these results, we present suggestive evidence based on descriptive statistics from two nationally representative surveys of schools in Peru: the 2014 National Survey of Educational Institutions (ENIE) and the 2014 National Survey of Teachers (ENDO). These surveys directly interview teachers and include information about their use of internet use in classrooms and their perceptions of the advantages of ICTs in education. Unfortunately, neither survey provides school identifiers linking these surveys to the CE or ECE. However, they allow us to characterize teachers' approaches to internet use in Peruvian schools by the end of our period of analysis.

The ENIE suggests a considerable degree of student exposure to internet in the classroom by 2014: among those who use it, 76% report using the internet at least once per week (and, on average, for 1.85 hours per week). Information in the ENDO also supports the notion of regular internet use in classrooms. As of 2014, 31% of Peruvian public primary school teachers at internet-connected schools listed internet-connected computers among the top 3 most-used classroom tools. Among schools with internet access, only very basic materials (such as photocopies and flip charts) score higher in this ranking, relative to internet-connected computers. Moreover, teachers view the internet as critical for the success of their students. 83.2% of second grade teachers in public schools with internet access believe internet increases students' access to information that is otherwise unavailable to them and 81.7% state that it improves collaborative learning among students. Among primary school teachers, the most often-cited school factor considered to negatively affect student learning is lack of access to new technologies, including the internet.

5.2.1 How do teachers use the internet?

On the other hand, student use of internet-equipped computers is only one way in which internet access could benefit students; it also can serve as a tool for teachers themselves. For example, teachers can access "off the shelf" lesson plans, repositories of practice questions, instructional aids, etc. According to Sandro Marcone (the former director of *Plan Huascarán*), what teachers demanded most from these programs was not only classroom resources themselves but lesson plans, saying essentially 'tell me what to teach' (Balarin (2013)). In the 2014 ENDO, 63% of teachers at internet-equipped primary schools considered access to internet and technology a top 3 factor in enhancing teaching performance— a larger proportion than those who listed either reference materials (38%) or networking with colleagues (38%). Over 64% of teachers at internet-connected public primary schools use the internet to access virtual pedagogical courses; the only other category with higher reported use is for email correspondence. Furthermore, school-based internet may be very important for teachers who otherwise might not have access to the internet. Among public primary school teachers, nearly half (44%) do not have internet access at home.

We use the data in the ENDO to estimate some suggestive correlations of teachers' ease to perform certain tasks and their access to internet. Public primary school teachers that have access to school-based internet in "good condition" report that teaching activities are less difficult than those without access, even conditional on teacher and school observable characteristics (Appendix

Table A.11). These activities include communicating with and motivating students, selecting and making good use of methodology and materials, using time effectively in the classroom, teaching according to different levels of student learning, and addressing the academic problems of students. Thus, teacher access to online materials may boost student performance above and beyond the impacts of direct student use of internet-connected computers.

5.2.2 Heterogeneity by class size

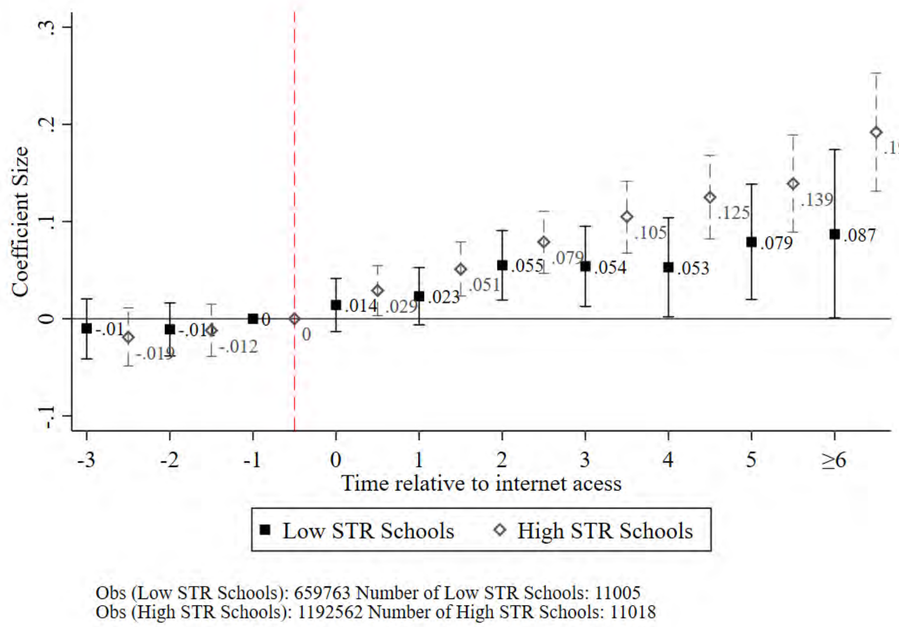
We next examine heterogeneity along the lines of class size, following Barrow et al. (2009). One might expect that the effects of internet (and other ICTs) may be stronger in larger classes for a variety of reasons. For example, previous work suggests that ICTs may provide students with more individualized attention than they would otherwise receive from teachers. If ICTs reduce the time teachers spend in group activities (for which internet might aid through interactive tools), they might be able to increase the time they allocate to individualized instruction. In particular, teachers assigned to larger classes might be more constrained in providing individualized instruction, and thus may be expected to see larger gains from ICTs.

Alternatively, internet access might be especially useful in strengthening the effectiveness of group work. For example, children may be more likely to focus on a group learning activity that involves watching a video or playing an educational game online than more traditional paper- or text-based activities. As mentioned, the overwhelming majority (81.7%) of teachers in internet-connected schools believe that internet access enhances collaborative learning among students (ENDO, 2014). This implies that internet connectivity would be very useful in classrooms with many students, where teachers rely heavily on group work. For both of these reasons, we expect internet access to matter more for schools with high versus low student to teacher ratios (STR).

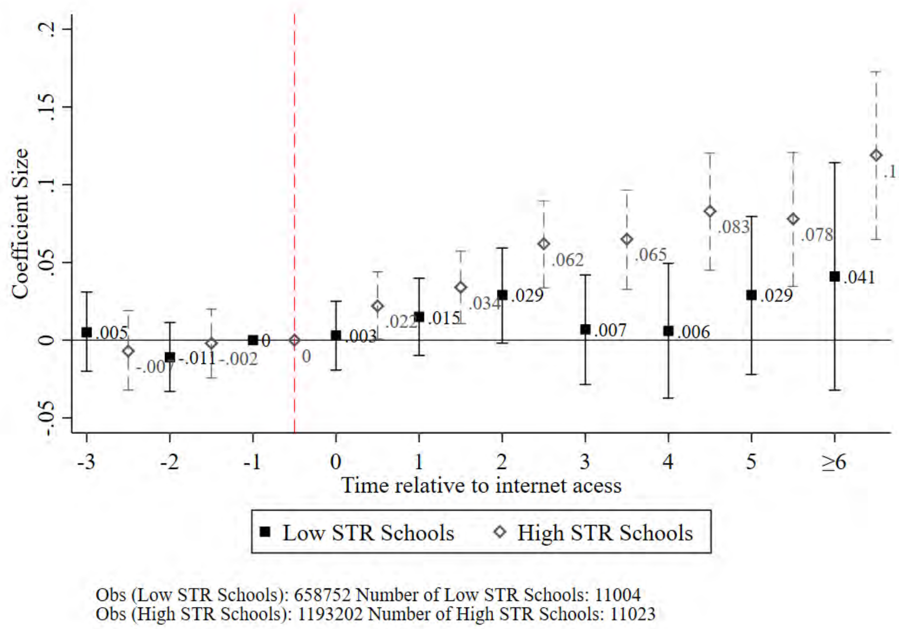
Splitting schools by the pre-internet median STR, we find that the positive effects of internet access are concentrated among schools with high STRs. We define "high STR" and "low STR" groups as follows. First, we calculate the total number of teachers per second grade student (we do not use the number of teachers exclusively dedicated to second grade, because many smaller schools assign teachers to multiple grades). Then, we calculate each school's pre-internet average STR (time-invariant); this includes all observations for schools that remain unconnected by the end of the CE sample period (2020). Finally, we divide the schools into high and low STR groups based on having a pre-internet average STR above or below the median. In Figures 6a and 6b, the high and low STR trends in test scores prior to internet access are similar, but diverge once internet is introduced. In low STR schools, the effects are much smaller (and close to zero for reading). From Figures 6a and 6b, we can see that the 95% confidence intervals of the event study indicators for low and high STR schools overlap. However, the overall pattern seems to suggest larger average gains for high STR schools. We test this in our trend break specification. Columns 1 and 2 Appendix Table A.12 confirm that the level shift and trend break in test scores are larger in high STR schools (though not statistically significant so). Thus we take these results as suggestive that the effects of internet are overall stronger in larger classes.

Figure 6: Heterogeneity by Student to Teacher Ratios: Event Study Results

(a) Math



(b) Reading



For each of the above figures, the sample is split based on each school's pre-internet average ratio of second graders to total teachers (STR) — high (low) STR schools fall above (below) the median pre-internet STR. The pre-internet median is calculated using all schools, including ones that do not have internet by 2020. Coefficients capture the increase in test scores relative to the year prior to a school receiving internet access ($t = -1$). Note that due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ while some receive it in $t = 1$. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school.

5.2.3 Heterogeneity teacher qualifications

Another possibility is that ICTs generate gains in student learning because they compensate for the lack or low quality of other inputs. For example, Jackson and Makarin (2018), determine that the benefits — in terms of math achievement — of providing teachers with online access to "off the shelf" lesson plans were larger among students with weaker teachers. Relatedly, some have found that the success of ICT interventions may depend on whether they displace traditional instruction or constitute additional learning activities outside of traditional classroom hours (as part of an after school tutoring program, for example, as in Linden 2008). In cases where ICTs substitute for traditional instruction, impacts may depend on the quality of instruction that the new technology is displacing. Such is hypothesized in Bulman and Fairlie (2016, p. 20), "[...] Interestingly, evidence of positive effects appears to be the strongest in developing countries. This could be due to the fact that the instruction that is being substituted for is not as of high quality in these countries."

Conversely, ICTs could be more effective when teachers are more highly qualified if, for example, teachers with post-secondary degrees may be better able to adapt to and use new technologies. Using the 2014 ENDO, we observe that more qualified teachers (e.g. those holding a college degree or higher) are more likely to report internet-connected PCs as one of the three most often used classroom tools than those without a college degree (34% versus 28%).

To shed some light on which effect is stronger, we examine heterogeneity in results by the level of qualifications that a school's teachers have obtained. In Figures 7a and 7b, we see that there is very little difference in the estimated effects across schools with low and high teacher qualifications. Here, we measure teacher qualification as the per student number of teachers with a pedagogical or university degree. We estimate the average ratio of qualified teachers-to-students by school using the pre-internet period (including all observations from schools that remain unconnected by 2017), and split the sample in two groups based on the sample median across schools. Those with ratios above (below) the sample median are classified as schools with "high" ("low") teacher qualifications. The only noticeable distinction is that the short-run gains in test scores (within the first 3 years of access) appear to be higher when schools have teachers with more qualifications; this can also be seen in columns 3 and 4 of Appendix Table A.12), where the level-shift appears larger for high qualification schools (though not statistically significantly so). This is consistent with the possibility that teachers with more qualifications are better able to immediately use the internet efficiently, while it takes longer for teachers without these qualifications to do so. However, after year 4, gains are similar across schools with high and low teacher qualifications.

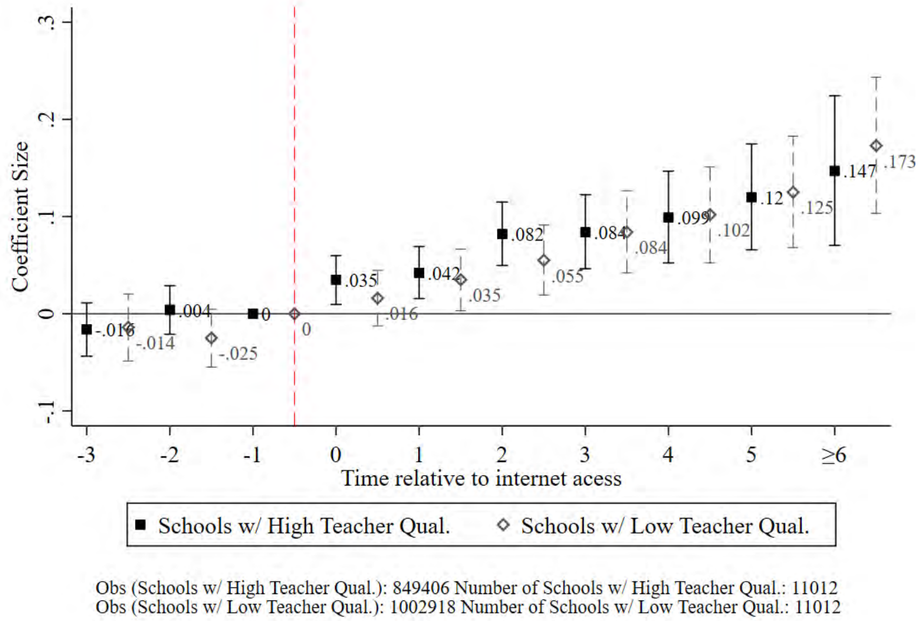
5.3 School-level Mechanisms

5.3.1 Complementary investments in trained teachers

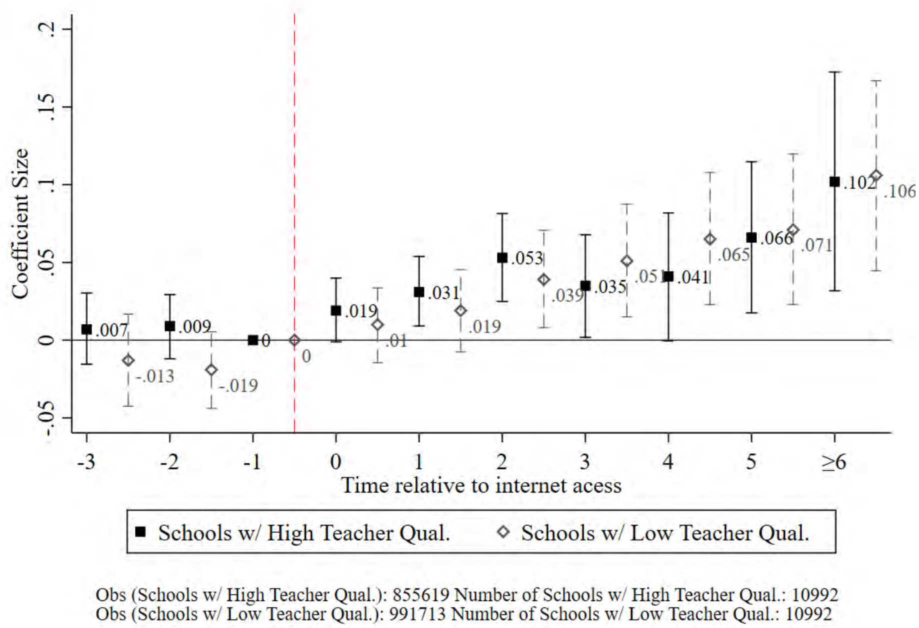
One explanation for why we observe delayed impacts of internet access may be that schools require teachers with digital and internet skills in order to incorporate the new technology into the classroom. To investigate this possibility, we study whether schools respond to internet access by hiring

Figure 7: Heterogeneity by Teacher Qualifications: Event Study Results

(a) Math



(b) Reading

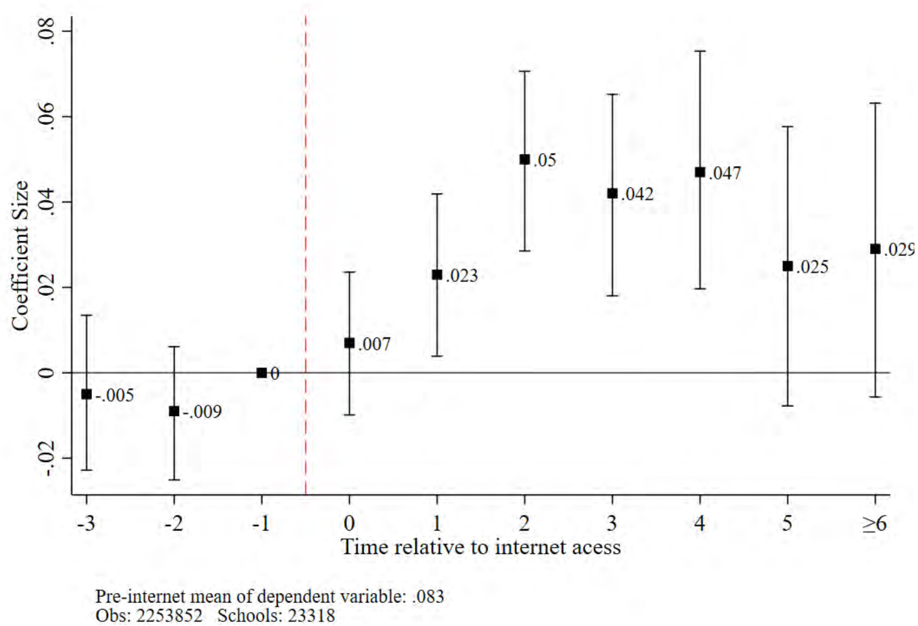


For each of the above figures, the sample is split based on each school's pre-internet average number of teachers with a pedagogical or higher education degree per student over the sample period relative to the median of all schools' sample averages. The pre-internet median is calculated using all schools, including ones that do not have internet by 2020. Coefficients capture the increase in test scores relative to the year prior to a school receiving internet access ($t = -1$). Note that due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ while some receive it in $t = 1$. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school.

teachers with expertise in "computer and information technology." This includes both teachers trained to teach computer skills, as well as teachers who themselves underwent advanced education relating to computers; hereafter, these are referred to as "computer teachers." We estimate equation 1 using an indicator for the presence of a computer teacher as the outcome.

Figure 8 shows that internet access is accompanied by a steady increase in the likelihood a school has a computer teacher that levels off 2 years post-internet installation; by year 4, this results in a 57% increase in the probability that a school has a computer teacher over the pre-internet likelihood. When taken together, the findings for computer teachers and test scores are consistent with the idea that schools may need time to make complementary investments to fully exploit new classroom technologies, such as teachers with computer training. However, because the presence of a computer teacher is a function of internet access (and is not exogenously given), we are not fully able to show the impact of the complementarity of both inputs on school performance.[41]

Figure 8: Internet Access and Presence of a Computer Teacher



Coefficients capture the increase in the likelihood of having a computer teacher on staff relative to the year prior to a school receiving internet access ($t = -1$). Note that due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ while some receive it in $t = 1$. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school.

[41]For example, we do not estimate equation 2 including a triple interaction between post-internet access, event time, and presence of a computer teacher because as shown in Figure 8, schools hire computer teachers *as a lagged response* to gaining internet access.

6 Conclusions

We find evidence that the introduction of internet to Peruvian primary schools produces economically meaningful improvements in student performance as measured by standardized test scores for grade 2. Gains increase over time, growing from 0.018-0.028 standard deviations in the year of installation to 0.063-0.111 standard deviations 5 years after installation (depending on the subject). Importantly, there are no apparent pre-existing trends in test scores prior to internet access, suggesting little role for reverse causality. Using a trend break specification, we confirm that there is a trend break in test scores that occurs at the time of internet access. In the medium term, the yearly gain in test scores is about 0.011-0.013 standard deviations. These results, based on over 2.2 million students from a large panel dataset of more than 23 thousand schools, are robust to a number of potential confounding factors, including changes in sample composition with respect to either schools or students, changes in school resources, and endogenous timing of installation with respect to prior trends in test performance. In our setting, the nationwide scale of roll-out, large sample of students and schools, and extended time frame uniquely enable the analysis of this technology's application at the farthest-reaching level of policy.

On the one hand, previous research on ICTs has found that providing ICT hardware with few or no complementary learning tools has little immediate impact on student performance (Bet et al. 2014; Barrera-Osorio and Linden 2009; Cristia et al. 2017; etc.). Our short run results (based on up to 1 year after internet installation) confirm that any effects of school-based internet access are small in magnitude — and thus perhaps impossible to detect in smaller samples of schools. On the other hand, medium run gains are sizable, pointing towards the necessity of a longer evaluation window for understanding the effectiveness of ICT interventions. Ultimately, our estimated effects of internet access between the second year and more than five years after internet installation still fall below prior estimates of the impact of computer assisted learning and instruction. While school-based internet does not fully confer the benefits of individualized pedagogical tools, it may provide access to learning resources that are otherwise unavailable to many students in developing countries.

Based on two additional nationally representative surveys of households and teachers, we present evidence about students' and teachers' use of internet. On one hand, internet connection in schools does make students more likely to directly access this tool. On the other, it appears that internet helps teachers increase their effectiveness through access to additional online teaching resources (e.g., off-the-shelf materials, teaching materials, etc.). Gains in test scores are somewhat concentrated among schools that have high student-teacher ratios. Hence, school-based internet may generate important gains in learning particularly when the number of teachers per student is constrained below the optimum. We also find that the short-run improvements in test scores are higher in schools where more teachers hold pedagogical or university degrees (though not statistically so), suggesting that teacher qualifications may be an important complementary input to new technology. However, we find that, over time, schools with lower teacher qualifications are able to benefit from access to internet as well.

We provide supporting evidence that achievement gains are slow to emerge because schools need time to adapt to new technologies. Specifically, after installing internet public schools require time to augment their staff with teachers experienced in computers and information technology. We thus concur with several prior studies finding that student achievement begins to increase only as teachers learn to integrate new technology into their curricula (Hull and Duch 2016, Mo et al. 2013, Sprietsma 2007).

However, the interpretation of the results presented is subject to a number of limitations. While we characterize teachers' and students' use of internet through secondary data sources, we are unable to directly incorporate measures of their use of technology as a mediating factor in our analysis. We are largely unable to explore heterogeneity in the effectiveness of school-based internet based on student characteristics. Indeed, previous work suggests that individual heterogeneity - especially with regard to initial achievement - significantly determines how technology affects the learning process (Bai et al. 2016; Barrow et al. 2009; Linden 2008; He et al. 2008; Muralidharan et al. 2019). Future research on heterogeneous impacts of internet in education could bear broad implications for inequality within and across learning environments.

Perhaps most notably, our results speak only to school-level dynamics and may mask important individual-level dynamics. Thus it is still an open question whether access to school-based technology might produce longer-term effects – for example, at higher schooling levels or later in the labor market. We consider longer-term individual-level studies (such as follow-ups of previous RCTs which evaluated relatively short-term impacts of technology) as an avenue for relevant future research.

References

Angrist, J. and Lavy, V. (2002), 'New Evidence on Classroom Computers and Pupil Learning', *The Economic Journal* **112**(482), 735–765. 13

Bai, Y., Mo, D., Zhang, L., Boswell, M. and Rozelle, S. (2016), 'The impact of integrating ICT with teaching: Evidence from a randomized controlled trial in rural schools in China', *Computers & Education* **96**, 1–14. 1, 38

Balarin, M. (2013), Las pol'íticas TIC en los Sistemas Educativos de América Latina: Caso Perú, Technical report, Programa TIC y Educación Básica, UNICEF, Buenos Aires, Argentina. 6, 31

Bando, R., Gallego, F., Gertler, P. and Romero, D. (2016), 'Books or Laptops? The Cost-Effectiveness of Shifting from Printed to Digital Delivery of Educational Content', *NBER Working Paper* (22928). 1, 15

Banerjee, A. V., Cole, S., Duflo, E. and Linden, L. (2007), 'Remedying Education: Evidence from Two Randomized Experiments in India', *Quarterly Journal of Economics* **122**(3), 1235–1264. 1, 15

Barrera-Osorio, F. and Linden, L. L. (2009), The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia, Policy Research Working Paper 4836, The World Bank, Washington, DC. 1, 4, 13, 25, 37

Barrow, L., Markman, L. and Rouse, C. (2009), 'Technology's Edge: The Educational Benefits of Computer-Aided Instruction', *American Economic Journal: Economic Policy* **1**(1), 52–74. 32, 38

Beg, S. A., Lucas, A. M., Halim, W. and Saif, U. (2019), 'Beyond the basics: Improving post-primary content delivery through classroom technology', *NBER Working Paper* (w25704). 15

Belo, R., Ferreira, P. and Telang, R. (2014), 'Broadband in School: Impact on Student Performance', *Management Science* **60**(2), 265–282. 1

Bet, G., Ibarraran, P. and Cristia, J. (2014), The Effects of Shared SchoolTechnology Access on Students' Digital Skills in Peru, IDB Working Paper Series 476, Inter-American Development Bank, Washington, DC. 2, 4, 13, 25, 37

Beuermann, D. W., Cristia, J., Cueto, S., Malamud, O. and Cruz-Aguayo, Y. (2015), 'One Laptop per Child at Home: Short-Term Impacts from a Randomized Experiment in Peru', *American Economic Journal: Applied Economics* **7**(2), 53–80. 1, 4, 6, 13, 25

Borusyak, K., Jaravel, X. and Spiess, J. (2021), 'Revisiting event study designs', *Unpublished Manuscript* . 12

Brynjolfsson, E. and Hitt, L. M. (2000), 'Beyond computation: Information technology, organizational transformation and business performance', *Journal of Economic perspectives* **14**(4), 23–48. 3

Bulman, G. and Fairlie, R. W. (2016), 'Technology and Education: Computers, Software, and the Internet', *NBER Working Paper* (22237). 1, 34

Callaway, B. and Sant'Anna, P. H. (forthcoming), 'Difference-in-differences with multiple time periods', *Journal of Econometrics* . 28

Carrillo, P., Onofa, M. and Ponce, J. (2010), Information Technology and Student Achievement: Evidence from a Randomized Experiment in Ecuador, IDB Working Paper Series 223, Inter-American Development Bank, Washington, DC. 1, 15

Cristia, J., Cueto, S., Ibarraran, P., Santiago, A. and Severin, E. (2017), 'Technology and Child Development: Evidence from the One Laptop per Child Program', *American Economic Journal: Applied Economics* **9**(3), 295–320. 1, 4, 6, 13, 22, 25, 37

Cristia, J., Czerwonko, A. and Garofalo, P. (2014), 'Does Technology in Schools Affect Repetition, Dropout and Enrollment? Evidence from Peru', *Journal of Applied Economics* **17**(11), 89–111. 2

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K. and Sundararaman, V. (2013), 'School Inputs, Household Substitution, and Test Scores', *American Economic Journal: Applied Economics* **5**(2), 29–57. 1

de Melo, G., Machado, A., Miranda, A. and Viera, M. (2013), Impacto del Plan Ceibal en el Aprendizaje. Evidencia de la Mayor Experiencia OLPC, Serie Documentos de Trabajo DT 13, Instituto de Economía, Facultad de Ciencias Económicas y de Administración, Unversidad de la República, Montevideo, Uruguay. 1

Escueta, M., Quan, V., Nickow, A. J. and Oreopoulos, P. (2017), 'Education Technology: An Evidence-based Review', *NBER Working Paper* (23744). 1

Faber, B., Sanchis-Guarner, R. and Weinhardt, F. (2015), 'ICT and Education: Evidence from Student Home Addresses', *NBER Working Paper* (21306). 1

Gibson, S. and Oberg, D. (2004), 'Visions and Realities of Internet Use in Schools: Canadian Perspectives', *British Journal of Educational Technology* **35**(5), 569–585. 1

Glewwe, P. and Kremer, M. (2006), Schools, Teachers, and Education Outcomes in Developing Countries, *in* E. A. Hanushek and F. Welch, eds, 'Handbook of the Economics of Education', Vol. 2, North Holland, chapter 16, pp. 945–1017. 1

Goodman-Bacon, A. (forthcoming), 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics* . 12, 28

Goolsbee, A. and Guryan, J. (2006), 'The Impact of Internet Subsidies in Public Schools', *The Review of Economics and Statistics* **88**(2), 336–347. 1

He, F., Linden, L. L. and MacLeod, M. (2008), How to Teach English in India: Testing the Relative Productivity of Instruction Methods within the Pratham English Language Education Program. 1, 15, 22, 38

Hjort, J. and Poulsen, J. (2019), 'The arrival of fast internet and employment in africa', *American Economic Review* **109**(3), 1032–79. 5

Hopkins, A. (2014), Internet in Schools. The effect on educational performance in Peru: 2007-2011. 4, 12

Hull, M. and Duch, K. (2016), One-to-one Technology and Student Outcomes, Unpublished manuscript, Department of Economics, University of North Carolina at Greensboro, Greensboro, NC. 3, 15, 38

International Telecommunication Union (2014), Final WSIS Target Review: Achievements, Challenges and the Way Forward, Technical report, Partnership on Measuring ICT for Development, Geneva, Switzerland. 1

Jackson, C. K. and Mackevicius, C. (2021), 'The distribution of school spending impacts', *NBER Working Paper* (w28517). 3, 15

Jackson, K. and Makarin, A. (2018), 'Can online off-the-shelf lessons improve student outcomes? evidence from a field experiment', *American Economic Journal: Economic Policy* **10**(3), 226–54. 1, 34

Kremer, M., Brannen, C. and Glennerster, R. (2013), 'The Challenge of Education and Learning in the Developing World', *Science* **340**(6130), 297–300. 1

Levin, D. and Arafeh, S. (2002), The Digital Disconnect: The Widening Gap between Internet-Savvy Students and their Schools, Technical report, American Institutes for Research, Washington, DC. Report prepared for the Pew Internet and American Life Project. 1

Linden, L. L. (2008), Complement or substitute?: The effect of technology on student achievement in India, InfoDev Working Paper 22928, World Bank, Washington, DC. 1, 15, 34, 38

Machin, S., McNally, S. and Silva, O. (2007), 'New Technology in Schools: Is There a Payoff?', *The Economic Journal* **117**(522), 1145–1167. 1

Malamud, O., Cueto, S., Cristia, J. and Beuermann, D. W. (2019), 'Do Children Benefit from Internet Access? Experimental Evidence from Peru', *Journal of Development Economics* **138**, 41–56. Mimeo. Inter-American Development Bank. 1

Meza-Cordero, J. A. (2017), 'Learn to Play and Play to Learn: Evaluation of the One Laptop per Child Program in Costa Rica', *Journal of International Development* **29**(1), 3–31. 1, 13

Mo, D., Swinnen, J., Zhang, L., Yi, H., Qu, Q., Boswell, M. and Rozelle, S. (2013), 'Can One-to-One Computing Narrow the Digital Divide and the Educational Gap in China? The Case of Beijing Migrant Schools', *World Development* **46**, 14–29. 1, 4, 13, 25, 38

Mo, D., Zhang, L., Wang, J., Huang, W., Shi, Y., Boswell, M. and Rozelle, S. (2015), 'Persistence of learning gains from computer assisted learning: Experimental evidence from China', *Journal of Computer Assisted Learning* **31**(6), 562–581. 3, 4, 15

Muralidharan, K., Singh, A. and Ganimian, A. J. (2019), 'Disrupting education? experimental evidence on technology-aided instruction in india', *American Economic Review* **109**(4), 1426–60. 1, 15, 38

Muralidharan, K. and Sundararaman, V. (2013), 'Contract Teachers: Experimental Evidence from India', *NBER Working Paper* (19440). 1

National Public Radio (2012), 'One Child, One Laptop ... And Mixed Results In Peru', Article for Weekend Edition Saturday on October 11, 2012. Last accessed March 2016.
URL: *http://www.npr.org/2012/10/13/162719126/one-child-one-laptop-and-mixed-results-in-peru* 4

OECD (2015), *Students, Computers and Learning: Making the Connection*, OECD Publishing. 1

One Laptop per Child (2011), 'Internet Connectivity: The Achilles Heel of the Peru Deployment'. Last accessed March 2016.
URL: *http://www.olpcnews.com/countries/peru* 4

One Laptop per Child (2016), '2016 Annual Report: A Year of Growth'.
URL: *http://one.laptop.org/magazine/annualreport/index.html* 1

Purcell, K., Heaps, A., Buchanan, J. and Friedrich, L. (2013), How Teachers are using Technology at Home and in their Classrooms, Technical report, PEW Research Center, Washington, DC. 1

Saavedra, J. and Suarez, P. (2002), El Financiamiento de la Educación en el Perú: El Rol de las Familias, Documento de Trabajo 38, Grupo de Análisis para el Desarrollo (GRADE), Lima, Peru. 5

Sharma, U. (2014), Can Computers Increase Human Capital in Developing Countries? An Evaluation of Nepal's One Laptop per Child Program. Paper presented at the 2014 Annual Meeting of the Agricultural and Applied Economics Association. 1, 13

Sprietsma, M. (2007), Computers as Pedagogical Tools in Brazil: A Pseudo-Panel Analysis, ZEW Discussion Paper 07-040, Centre for European Economic Research, Mannheim, Germany. 2, 38

Sun, L. and Abraham, S. (forthcoming), 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics* . 28, 52

The World Bank (2016), 'World Development Indicators'. Accessed August 2017.
URL: *https://openknowledge.worldbank.org/handle/10986/23969* 5

Toyama, K. (2015), *Geek Heresy: Rescuing Social Change from the Cult of Technology*, first edn, PublicAffairs, New York, NY. 1

Trucano, M. (2012), Evaluating One Laptop Per Child (OLPC) in Peru. Post in the EduTech blog.
URL: *http://blogs.worldbank.org/edutech/olpc-peru2* 6

Trucano, M. (2016), 'Technologies in Education Across the Americas: The promise and the peril and some potential ways forward', *World Bank Education, Technology & Innovation: SABER-ICT Technical Paper* (12). 1

UNESCO (2012), ICT IN EDUCATION IN LATIN AMERICA AND THE CARIBBEAN: A regional analysis of ICT integration and e-readiness, Technical Report 22923, UNESCO Institute for Statistics, Montreal, Canada. 1

Vigdor, J. L., Ladd, H. F. and Martinez, E. (2014), 'Scaling the Digital Divide: Home Computer Technology and Student Achievement', *Economic Inquiry* **52**(3), 1103–1119. 1

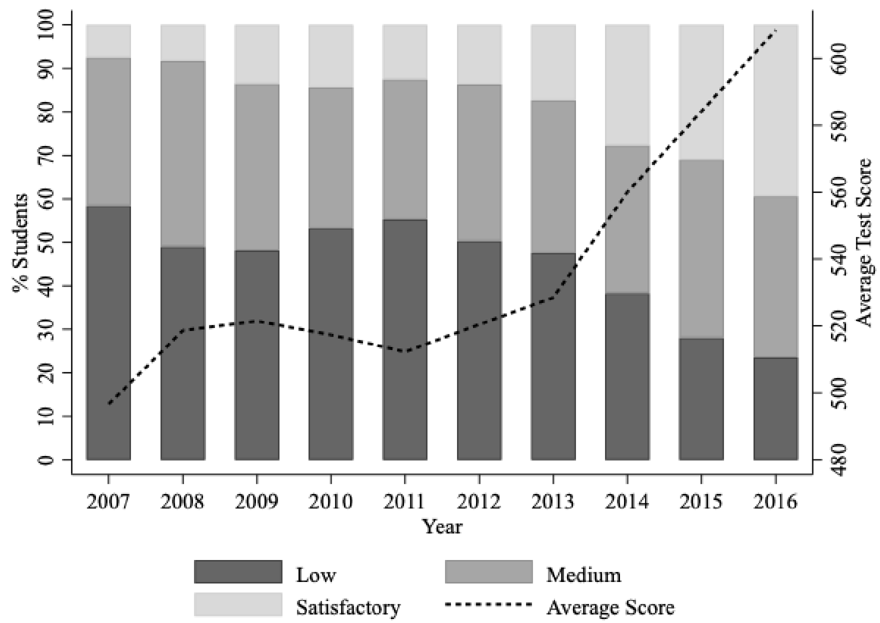
World Bank (2012), Public Expenditure Review for Peru:Spending for Results. Report No. 62586 - PE. 5

World Bank (2018), World Development Report 2018: Learning to Realize Education's Promise, Technical report, The World Bank, Washington, DC. 1

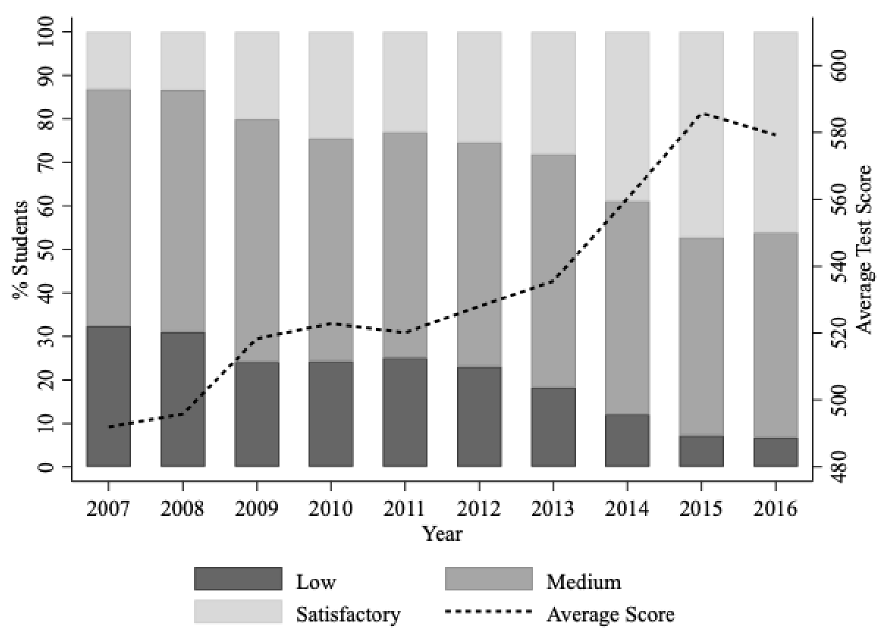
7 Appendix Figures and Tables

Figure A.1: Performance of Grade 2 Students in Public Schools on the ECE (2007-2014)

(a) Average Test Scores and Performance Categories in Math

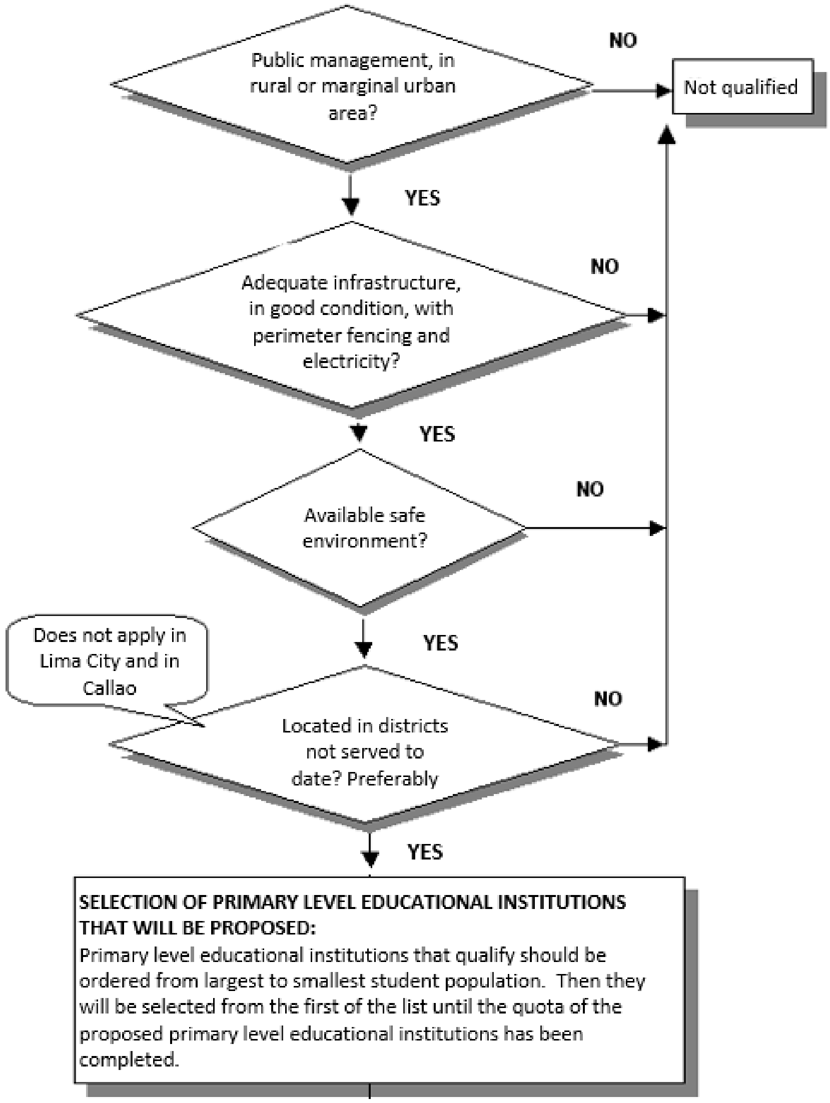


(b) Average Test Scores and Performance Categories in Reading



Source: Peru Ministry of Education (MINEDU)

Figure A.2: Translated Excerpt of Flow Chart for Prioritization under *Plan Huascarán*



Authors' translation. Original document in Spanish can be found here: http://www.minedu.gob.pe/normatividad/directivas/Dir083VMGP2003.php. <Accessed October 4, 2017>

Figure A.3: Translated School Data Sheet for *Plan Huascarán*

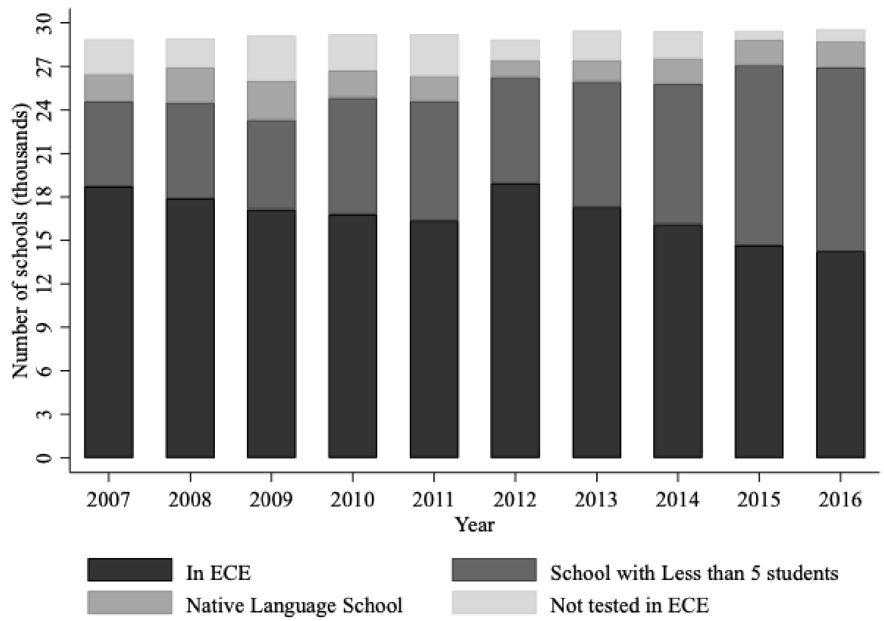
ANNEX N ° 2

DATA SHEET OF THE EDUCATIONAL INSTITUTION		
Name of Educational Institution		
School Site Code		
Address		
Department		
Province		
District		
Town Center		
Phone		
Principal's name		
Direct intermediary body		
Geographical area (urban, rural)		
Type of Management (State, Parish, Cooperative, Supervised, etc.)		
Number of computers for school use (only Pentium I or more)		
Number of computers for administrative use (only Pentium I or more)		
Do you have electricity?		
Number of hours of electricity		
Number of students and teachers per level	Students	Teachers
Initial		
Primary		
High school		
Number of students and teachers per shift	Sections	
Morning		
Late		
Night		
Number of sections per shift	Sections	
Morning		
Late		
Night		
Number of sections per level	Students	Teachers
Initial		
Primary		
High school		
Is there home-based telephone in the locality?		
Native language of students		
Distance to the nearest Huascarán Program educational institution		

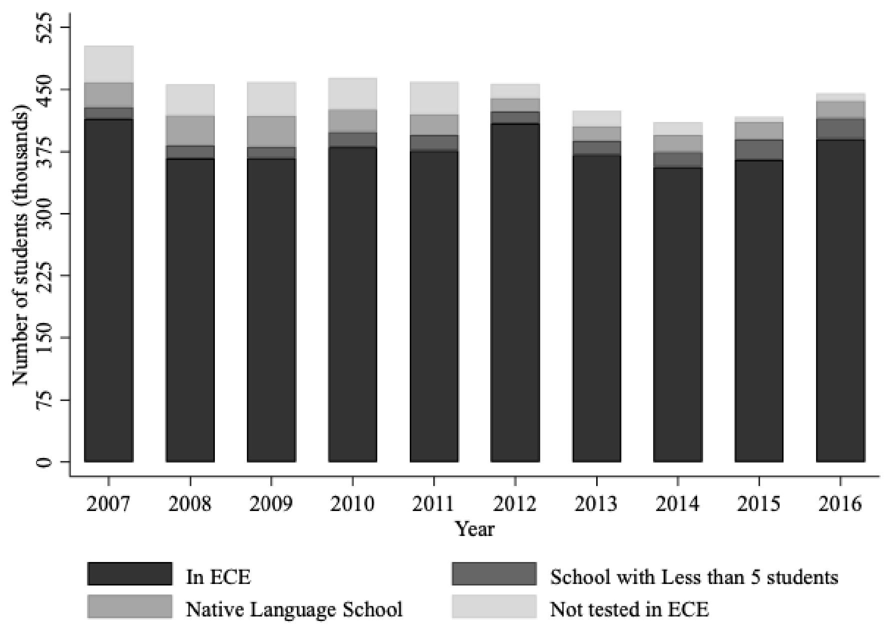
Authors' translation. Original document in Spanish can be found here: http://www.minedu.gob.pe/normatividad/directivas/Dir083VMGP2003.php. <Accessed October 4, 2017>

Figure A.4: Schools and Students in ECE, 2007-2014

(a) Primary Schools tested in ECE

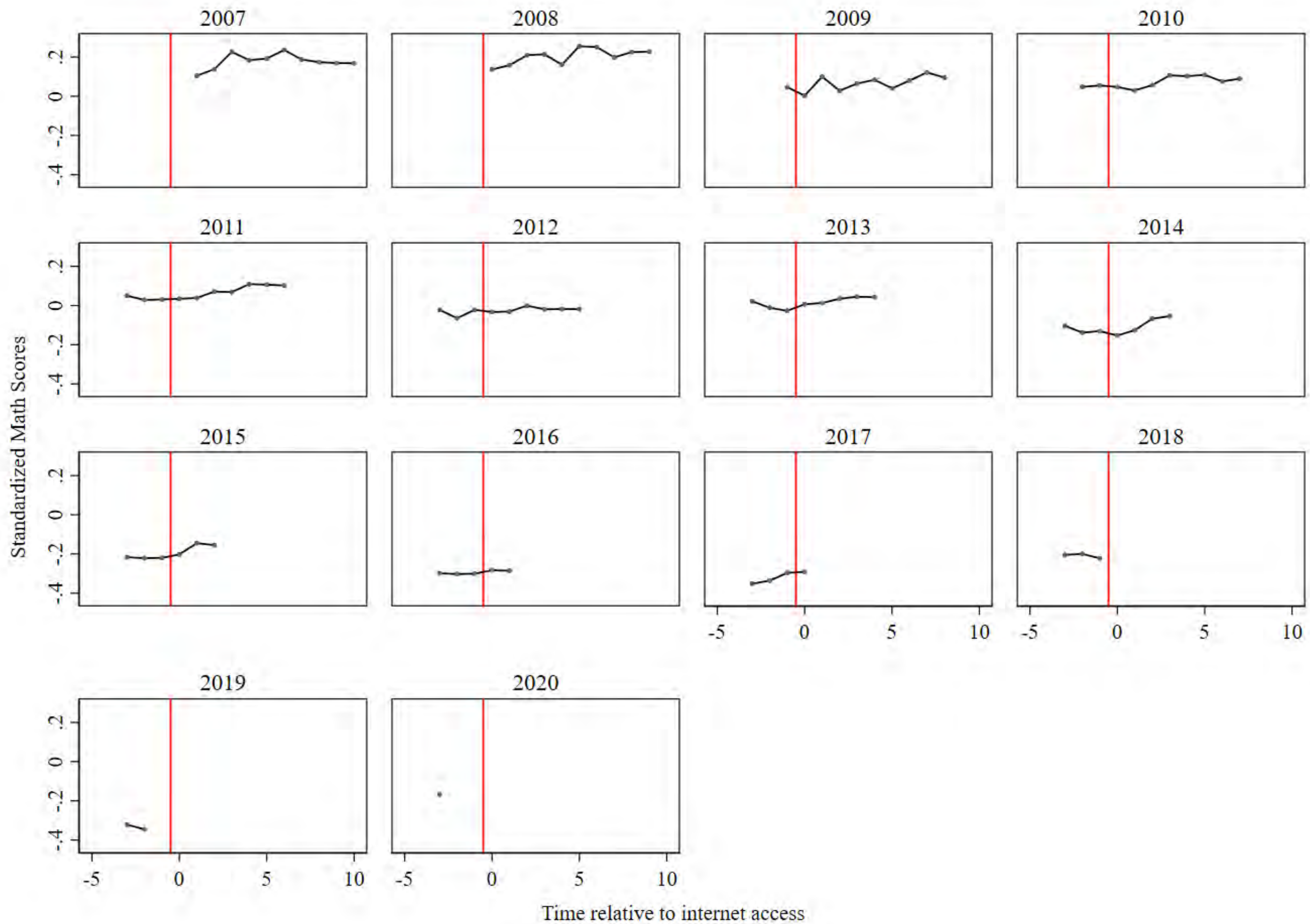


(b) Primary Students tested in ECE



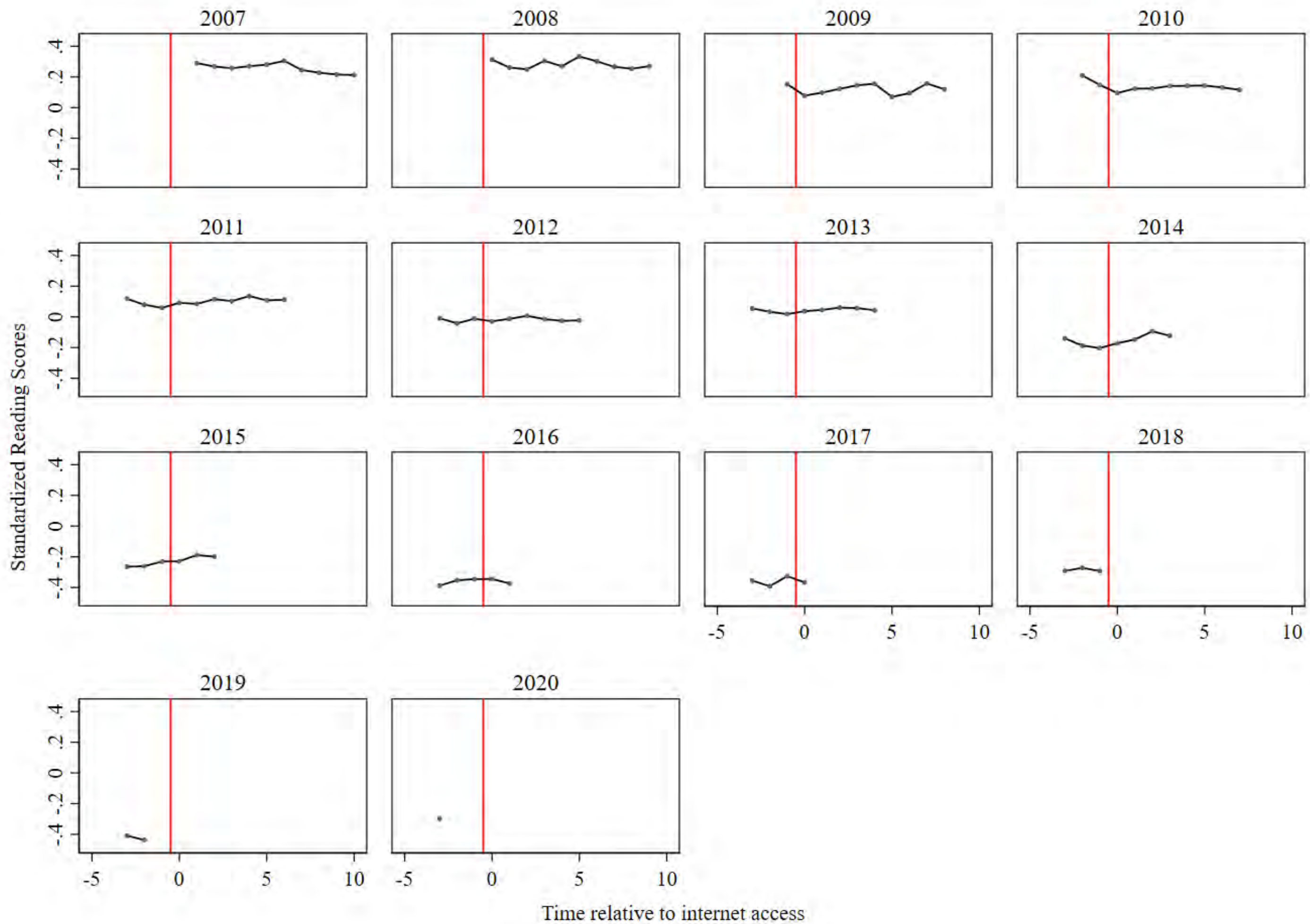
Source: Authors' calculations based on the Peruvian *Censo Escolar* (CE) and *Evaluacion Censsal de Estudiantes* (ECE), 2007-2016.
* Note: Some schools both have fewer than five second graders and teach primarily in native languages. For simplicity, the graph includes these under "Fewer than five students."

Figure A.5: Standardized Math Scores over Time, by Year of Initial School Access



This figure plots the standardized test scores against time relative to internet access, separately for by the year of initial internet connection.

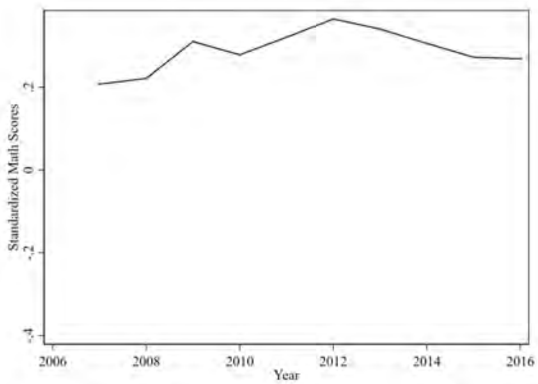
Figure A.6: Standardized Reading Scores over Time, by Year of Initial School Access



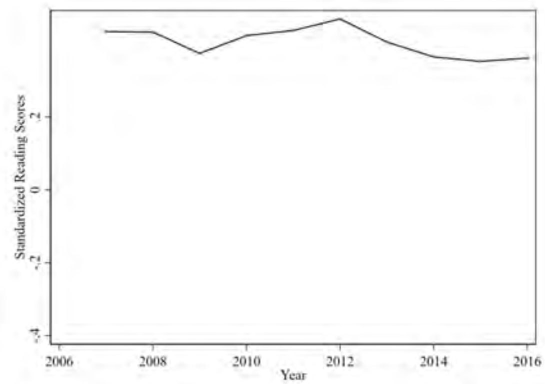
This figure plots the standardized test scores against time relative to internet access, separately for by the year of initial internet connection.

Figure A.7: Standardized Test Scores over Time in Schools Installing Internet before 2007 and in Schools without Access by 2020

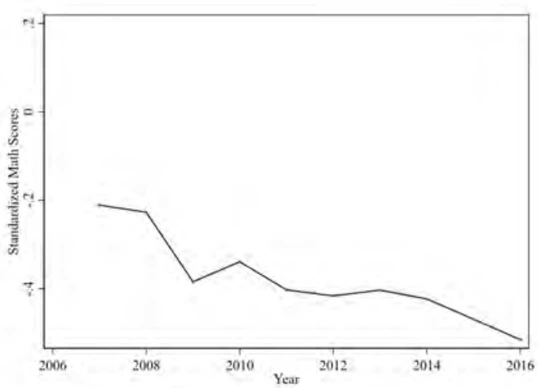
(a) Math: Schools with Access before 2007



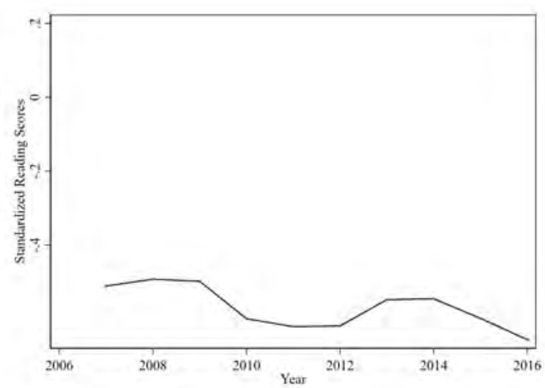
(b) Reading: Schools with Access before 2007



(c) Math: Schools without Access by 2020



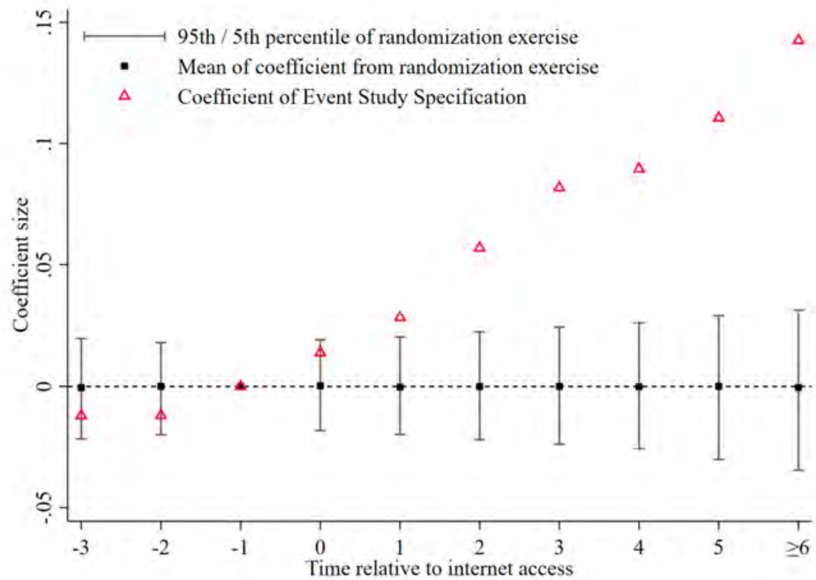
(d) Reading: Schools without Access by 2020



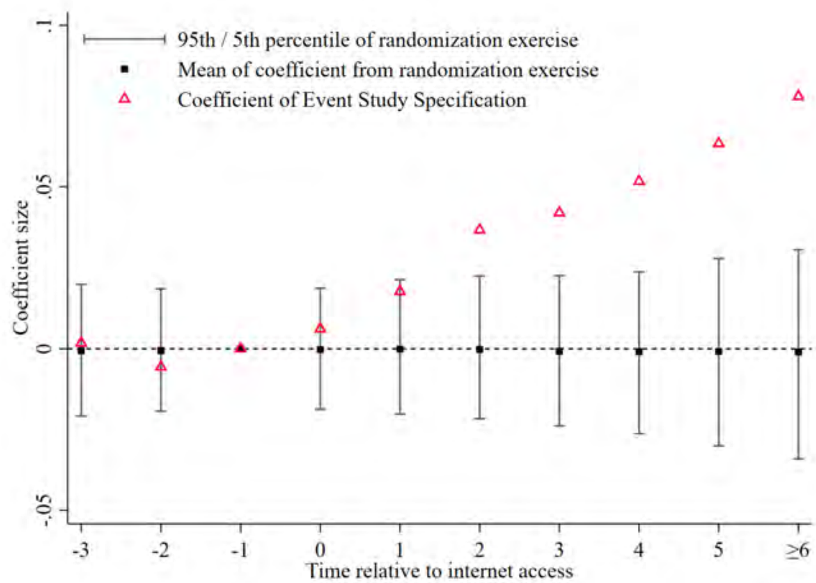
Figures A.7a and A.7b plot the standardized test scores for students from all public schools that had an internet connection prior to 2007. Figures A.7c and A.7d plot the standardized test scores for students from all public schools that had not been connected to the internet as of 2020.

Figure A.8: Randomization Exercise

(a) Standardized Math Scores



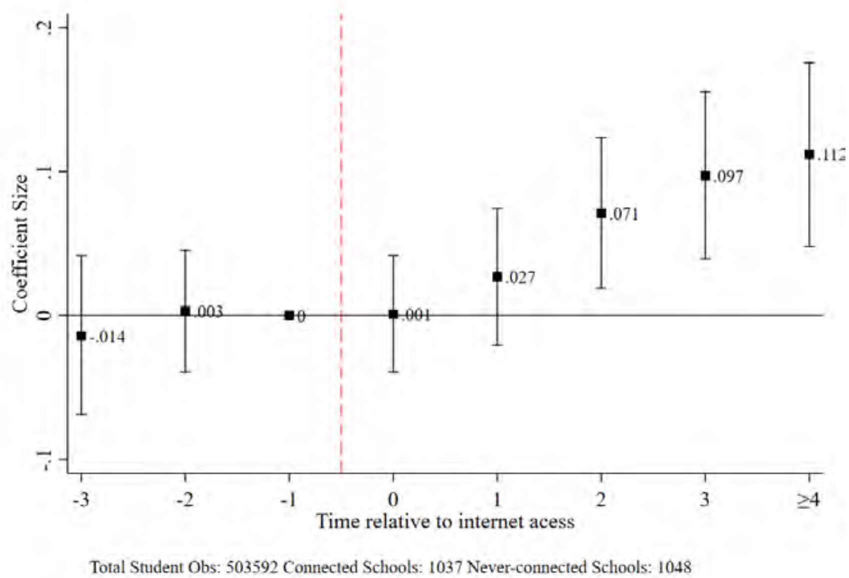
(b) Standardized Reading Scores



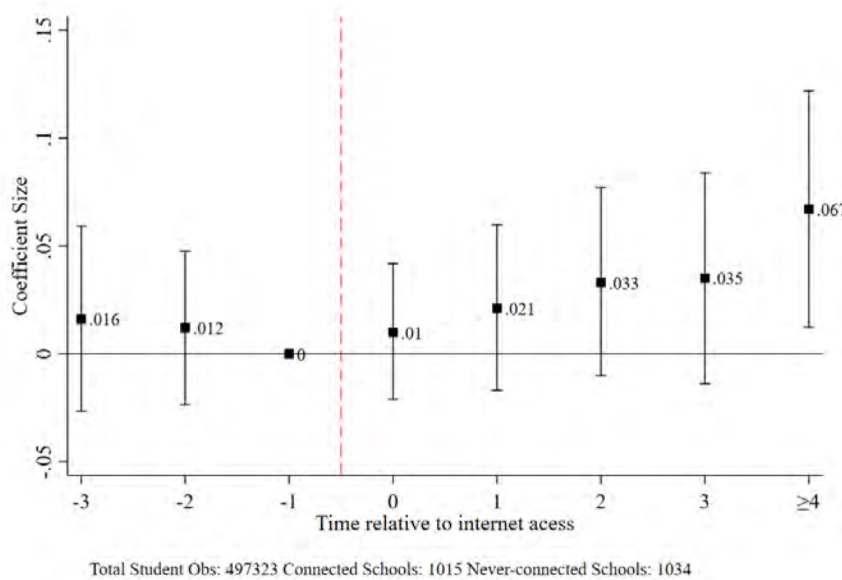
These figures plots the median as well as the 5th and 95th percentile of the coefficients from the following placebo test: We randomly reassign internet installation years to schools in our sample (including non-adoption), maintaining the actual distribution of installation years and thus ensuring that the sample sizes of each of our "treatment cohorts" matches our baseline specification. We then use the randomly assigned installation years to generate the false event study variables and estimate equation 1. We repeat this process 500 times. For reference, we also plot the estimates from our baseline specification (equation 1). Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school.

Figure A.9: Balanced Panel Estimates

(a) Standardized Math Scores



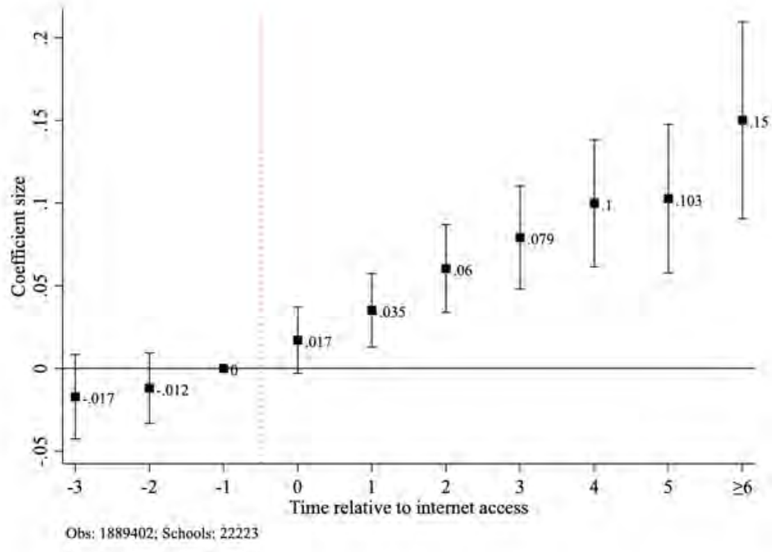
(b) Standardized Reading Scores



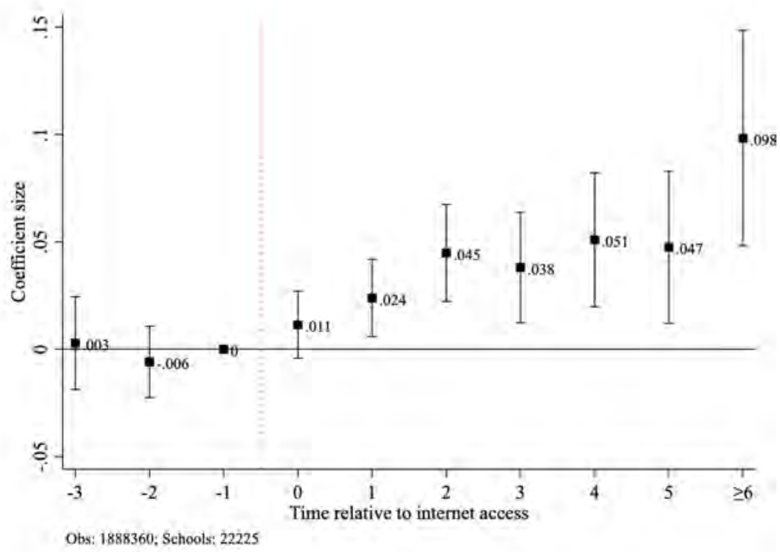
The sample includes all grade 2 students from all public schools using only internet-connected schools that appear in all event periods and non-connected schools that appear in all calendar years of the sample period. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ and some receive it in $t = 1$. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Standard errors are clustered by school

Figure A.10: Sun and Abraham Estimates

(a) Math



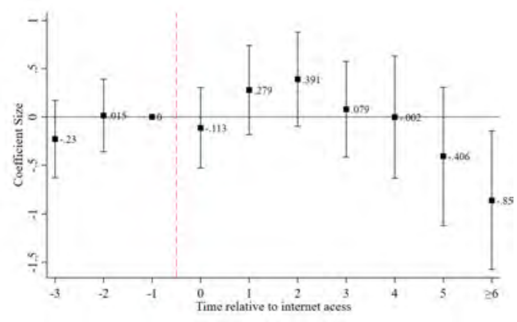
(b) Reading



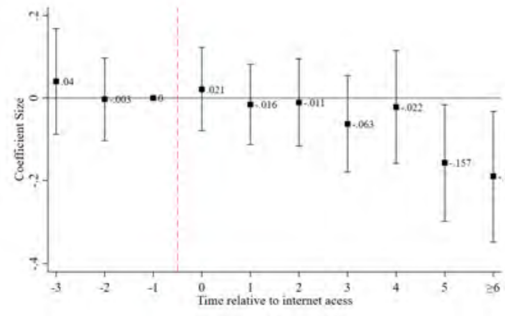
These estimates are calculated using the method described in Sun and Abraham (forthcoming). The sample includes all grade 2 students from a restricted set of cohorts that first installed internet between 2010 and 2018, as required by the estimation procedure. The standard errors are bootstrapped and are estimated with 500 replications. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools.

Figure A.11: Internet Access and Grade 2 Transfers, Re-entry, Enrollment, Test Taking, and Student Composition

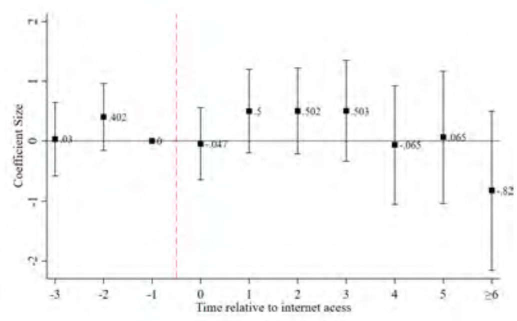
(a) Transfers



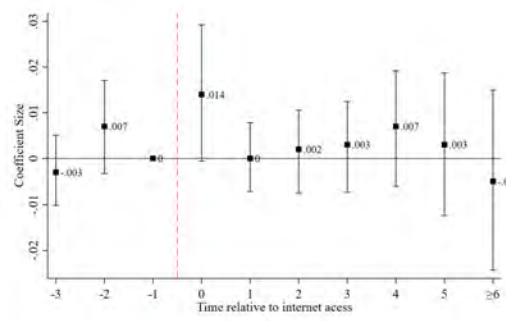
(b) Re-entrants



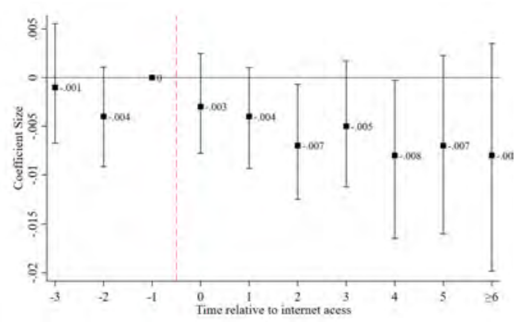
(c) Enrollment



(d) Proportion of Students Taking the ECE



(e) Proportion of Native Spanish Speakers



Transfers are total (net) grade 2 transfers from other schools. Re-entrants are total grade 2 students entering school after not being enrolled in any school the previous year. Enrollment includes all grade 2 students at the school except the individual. Proportion of students taking the ECE is the proportion of enrolled grade 2 students that take the ECE exam. Proportion of Native Spanish Speakers is the proportion of grade 2 students that speak Spanish at home. Each of the above figures plots the coefficients and 95% confidence intervals from estimating equation 1. Coefficients capture the change in outcomes relative to the year before internet installation ($t = -1$). Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The sample includes all grade 2 students in public schools that gain internet after 2007 or remain unconnected by 2020. Standard errors are clustered by school.

Table A.1: Predictors of Internet Access

	Dependent Variable: School Has Gained Internet Access		
	Including Characteristics Prioritized by *Plan Huascarán*	Adding a Control for for Fencing (2010 and Later)	Adding Characteristics in School Data Sheets
	(1)	(2)	(3)
School has a computer room	0.102***	0.071***	0.098***
	(0.005)	(0.005)	(0.006)
School has electricity	0.073***	0.119***	0.061***
	(0.006)	(0.008)	(0.006)
Total Enrollment, in 100s of students	0.030***	0.027***	0.024***
	(0.001)	(0.002)	(0.003)
School is in an urban area	0.152***	0.173***	0.141***
	(0.009)	(0.010)	(0.009)
School has a full perimeter fence		0.036***	
		(0.006)	
Number of computers for instruction			0.001***
			(0.000)
Number of computers for administration			0.001*
			(0.001)
Total number of staff with teaching responsibilities			0.002***
			(0.001)
Observations	2,251,549	1,720,337	1,978,996
Number of schools	23,318	20,714	22,934

The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered at the school level. Prioritized characteristics include enrollment and facilities (computer room, electricity), district fixed effects (to capture poverty status) and UGEL-specific enrollment tercile by year fixed effects. Information on the existence of a perimeter fence is only available for 2010 and later. Additional characteristics on school data sheets include number of computers used for instruction, number of computers used for administrative purposes, and number of teachers. Significance levels denoted by: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table A.2: Summary Statistics (2007 or Earliest Available Year) of Public Primary Schools

	A. All Public Schools			B. Public Schools in ECE 2007-2014			C. Estimation Sample
	Internet before 2007	Internet 2007-2020	No Internet by 2020	Internet before 2007	Internet 2007-2020	No Internet by 2020	Internet 2007-2020 or No Internet by 2020 Matched to ECE, All Covariate Info.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Standardized math score				0.19	-0.04	-0.29	-0.11
				(0.43)	(0.63)	(0.93)	(0.74)
Standardized reading score				0.42	-0.04	-0.59	-0.20
				(0.41)	(0.64)	(0.85)	(0.75)
Rural school	0.07	0.61	0.94	0.07	0.58	0.94	0.78
	(0.25)	(0.49)	(0.23)	(0.25)	(0.49)	(0.24)	(0.41)
One teacher per grade	0.94	0.44	0.06	0.94	0.47	0.06	0.24
	(0.24)	(0.50)	(0.23)	(0.24)	(0.50)	(0.24)	(0.43)
Total enrollment	551.30	155.98	44.54	553.44	166.31	50.83	102.80
(Grades 1-6)	(348.57)	(196.79)	(47.93)	(347.83)	(201.01)	(47.48)	(150.10)
Enrollment in 2nd Grade	89.32	27.30	8.95	89.66	29.04	10.00	18.59
	(58.75)	(33.73)	(9.10)	(58.65)	(34.41)	(8.94)	(25.70)
Schools connected to public	0.87	0.51	0.27	0.87	0.52	0.29	0.39
water drinking network	(0.34)	(0.50)	(0.45)	(0.34)	(0.50)	(0.45)	(0.49)
School has library	0.70	0.34	0.24	0.70	0.35	0.24	0.29
	(0.46)	(0.47)	(0.42)	(0.46)	(0.48)	(0.43)	(0.45)
School has administrative	0.80	0.41	0.22	0.80	0.43	0.24	0.32
office(s)	(0.40)	(0.49)	(0.42)	(0.40)	(0.49)	(0.43)	(0.47)
Total number of textbooks	359.85	112.32	37.96	360.61	118.78	41.87	78.56
in school	(269.55)	(156.47)	(44.08)	(269.45)	(160.27)	(45.06)	(121.40)
Number of teachers	6.12	5.51	4.97	6.13	5.54	5.15	5.32
	(7.87)	(7.91)	(7.71)	(7.88)	(7.94)	(7.67)	(7.78)
Number of classrooms	20.47	6.82	2.57	20.54	7.17	2.77	4.75
	(11.62)	(6.08)	(2.10)	(11.60)	(6.11)	(2.21)	(4.92)
Computers in school	17.89	2.33	0.31	17.95	2.44	0.34	1.28
	(19.58)	(6.89)	(2.23)	(19.60)	(6.56)	(2.39)	(4.88)
Number of schools	1366	11310	18435	1360	10438	14310	23318

Panel A includes all public schools in the CE from 2007 to 2020. Panel B restricts the sample to schools that participated in the ECE at least once between 2007 and 2016. Panel C includes all schools that either received internet in 2007-2020 or had not received internet by 2020 and that contain all necessary covariate information for the baseline specification (see Section 3.1). We present data for 2007 (or the earliest available year with information in the sample, when data for 2007 is not available). Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. In the Peruvian school system, schools might be *unidocente* (only one teacher in the school teaches all grades), *multigrado* (more than one teacher, but each might teach more than one grade in the same classroom), or *polidocente completo* (there is one teacher per grade in the school). Standard deviations in parentheses.

Table A.3: Effect of Internet Access on Standardized Test Scores

	Dependent Variable: Standardized Test Scores	
	Math	Reading
	(1)	(2)
t=-3	-0.012	0.002
	(0.011)	(0.009)
t=-2	-0.012	-0.006
	(0.010)	(0.008)
t=0	0.014	0.006
	(0.009)	(0.008)
t=1	0.028***	0.018**
	(0.010)	(0.008)
t=2	0.057***	0.037***
	(0.012)	(0.010)
t=3	0.082***	0.042***
	(0.013)	(0.011)
t=4	0.089***	0.052***
	(0.015)	(0.013)
t=5	0.111***	0.063***
	(0.017)	(0.014)
t=6	0.143***	0.078***
	(0.021)	(0.018)
p-value for Test of Joint Significance for All $t < 0$:	0.383	0.621
p-value for Test of Joint Significance for All $t \geq 0$:	0.000	0.001
Observations	2,253,852	2,252,368
Number of schools	23,318	23,320

Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Note that due to the timing of the *Censo Escolar* relative to the ECE exam, some schools receive internet access in $t = 0$ while some receive it in $t = 1$. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. Standard errors are clustered by school. Significance levels denoted by: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table A.4: Summary Statistics for Alternate Estimation Samples

	Baseline Sample (1)	Observed at least Twice Before and After Installation (2)	ECE Scores in All Years (3)
Standardized Math Score	-0.085	-0.147	-0.002
	(0.994)	(0.991)	(0.977)
Standardized Reading Score	-0.122	-0.212	0.016
	(0.987)	(0.983)	(0.962)
Years of internet access	1.793	1.034	2.194
	(2.540)	(1.713)	(2.682)
Number of Students Taking the Test	55.101	43.773	67.070
	(53.861)	(43.989)	(54.070)
School Has Library	0.463	0.410	0.515
	(0.499)	(0.492)	(0.500)
School Has Administrative Office(s)	0.451	0.420	0.509
	(0.498)	(0.493)	(0.500)
Ratio of Classrooms to Students	0.060	0.063	0.052
	(0.053)	(0.056)	(0.036)
Ratio of Computers to Students	0.135	0.143	0.117
	(0.231)	(0.252)	(0.166)
Ratio of Teachers to Students	0.050	0.052	0.045
	(0.045)	(0.049)	(0.022)
Total School Enrollment	342.902	269.058	419.338
	(336.259)	(272.205)	(337.270)
Observations	2,245,942	1,774,410	960,859
Number of schools	23,318	20,300	3,764

Column 1: The baseline sample includes students from all public schools that gained internet access between 2007 and 2020 or that remained unconnected by 2020. Column 2: The sample includes grade 2 students from schools observed at least twice prior to and twice after internet access, i.e. that installed internet between 2009 and 2015. Column 3: The non-attritor sample are schools that are observed for each year 2007-2016. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools.

Table A.5: Understanding the Role of School-level Compositional Changes: Trend Break Results

	Dependent Variable: Standardized Test Score					
	Math			Reading		
	Baseline Sample	Observed at least Twice Before and After Installation	Observed in All Years	Baseline Sample	Observed at least Twice Before and After Installation	Observed in All Years
	(1)	(2)	(3)	(4)	(5)	(6)
Post-internet Access	0.014	0.014	0.000	0.015	0.021**	0.013
	(0.011)	(0.012)	(0.016)	(0.009)	(0.010)	(0.014)
Post-internet Access X Event Time	0.013**	0.015**	0.009	0.011**	0.013**	0.010
	(0.005)	(0.006)	(0.008)	(0.004)	(0.005)	(0.006)
Event Time	0.008	0.007	0.014*	-0.000	-0.002	0.001
	(0.005)	(0.006)	(0.008)	(0.005)	(0.005)	(0.007)
p-value of Test of Joint Significance for Post and PostXEventTime:	0.0520	0.0560	0.322	0.0620	0.0380	0.323
Observations	2,253,852	1,763,314	963,857	2,252,368	1,762,354	951,456
Number of schools	23,318	21,300	3,764	23,320	21,302	3,691

Columns 1 and 4 reproduce the baseline results using all public schools that gained internet access between 2007 and 2020 or that remained unconnected by 2020. Columns 2 and 5 present results using the restricted sample of schools that gained internet access between 2009 and 2015 (i.e. observed at least twice prior to and twice after installing internet) or that remained unconnected by 2020. Columns 3 and 6 present results using all schools that gained internet access between 2007 and 2017 or that remained unconnected by 2017 *and* that are observed for the entire sample period, i.e. for each year 2007-2016 (non-attritors). Standard errors are clustered by school. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Post-internet access is a dummy variable for whether a school has gained internet access (i.e. $t >= 0$). Event time is years relative to internet access. ontrol variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Significance levels denoted by: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table A.6: Attrition (Number of School-Year Observations)

	N	%
All Information Observed (Not Attrited)	123,477	59.26
Attrited	84,899	40.74
School Permanently Closed	5,375	2.58
Missing ECE Scores		
Second Grade Enrollment less than 5	46,228	22.18
Second Grade Enrollment between 5 and 8	12,578	6.04
Other	11,471	5.50
Missing Census (CE) Information	9,320	4.43
Other	17	0.01

Based on each school's initial year of internet connection and our event study window, we determine all the periods that each should be included in our panel dataset. Enrollment is measured from the CE (reported at the beginning of each year). Only schools with five or more second graders by the end of each year are tested in the ECE. Schools that have 5-7 students at the beginning of the year might not have been included in the ECE if they fell below the 5-student threshold by the end of the year. CE information is used to calculate infrastructure (and, importantly, internet access) and school resources (control variables in our regressions).

Table A.7: Assessing the Role of Covariates: Trend Break Results

	Dependent Variable: School Average Standardized Test Score					
	Math			Reading		
	No Controls (1)	Only Enrollment (2)	Adding Resources (Baseline) (3)	No Controls (4)	Only Enrollment (5)	Adding Resources (Baseline) (6)
Post-internet Access	0.014	0.014	0.014	0.014	0.014	0.015
	(0.011)	(0.011)	(0.011)	(0.009)	(0.009)	(0.009)
Post-internet Access X Event Time	0.013**	0.013**	0.013**	0.011**	0.011**	0.011**
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
p-value of Test of Joint Significance for Post and PostXEventTime:	0.0400	0.0370	0.0520	0.0490	0.0470	0.0620
Observations	2,253,852	2,253,852	2,253,852	2,252,368	2,252,368	2,252,36
Number of schools	23,318	23,318	23,318	23,320	23,320	23,320

The sample includes all grade 2 students in all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. All columns include UGEL-specific enrollment tercile by year fixed effects and school fixed effects. Columns 1 and 4 do not include any time-varying controls. Columns 2 and 5 add in controls for total student enrollment and number of second grade students that took the ECE. Columns 3 and 6 add in facilities (computer room, library, administrative offices) and resources per student (classrooms, computers, and teachers) and sex, class size, indicator variables for the number of 2nd grade classes at the school. This is our baseline specification. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Post-internet access is a dummy variable for whether a school has gained internet access (i.e. $t >= 0$). Event time is years relative to internet access. Significance levels denoted by: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table A.8: Effects of Computers in Non-internet Schools

	Standardized Math Score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Computers per Student	0.011	0.016	-0.004	-0.007	0.000	-0.001	-0.005
	(0.012)	(0.014)	(0.014)	(0.015)	(0.016)	(0.019)	(0.023)
Computers per Student (1 Year Lag)		-0.003	-0.012	-0.021	-0.028*	-0.022	-0.017
		(0.013)	(0.014)	(0.015)	(0.016)	(0.018)	(0.022)
Computers per Student (2 Year Lag)			0.027**	0.024*	0.027**	0.027*	0.023
			(0.013)	(0.013)	(0.013)	(0.015)	(0.023)
Computers per Student (3 Year Lag)				-0.019	-0.013	-0.012	0.006
				(0.014)	(0.015)	(0.017)	(0.023)
Computers per Student (4 Year Lag)					-0.031	-0.020	-0.029
					(0.021)	(0.021)	(0.022)
Computers per Student (5 Year Lag)						0.027	0.036
						(0.028)	(0.031)
Computers per Student (6 Year Lag)							0.014
							(0.100)
Observations	607,114	582,735	481,945	393,088	318,989	255,778	199,694
Number of schools	13,791	13,556	12,297	11,134	9,753	8,531	7,253
	Standardized Reading Score						
	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Computers per Student	-0.001	0.003	-0.010	-0.012	-0.012	-0.014	-0.017
	(0.011)	(0.012)	(0.013)	(0.014)	(0.015)	(0.017)	(0.021)
Computers per Student (1 Year Lag)		0.000	-0.006	-0.018	-0.021	-0.021	-0.020
		(0.012)	(0.013)	(0.013)	(0.014)	(0.016)	(0.020)
Computers per Student (2 Year Lag)			0.021**	0.019*	0.021*	0.025**	0.022
			(0.011)	(0.011)	(0.011)	(0.012)	(0.021)
Computers per Student (3 Year Lag)				-0.012	-0.007	-0.002	0.014
				(0.014)	(0.014)	(0.015)	(0.024)
Computers per Student (4 Year Lag)					-0.029	-0.022	-0.029
					(0.021)	(0.021)	(0.021)
Computers per Student (5 Year Lag)						0.024	0.039
						(0.025)	(0.028)
Computers per Student (6 Year Lag)							-0.003
							(0.095)
Observations	607,114	582,735	481,945	393,088	318,989	255,778	199,694
Number of schools	13,793	13,556	12,297	11,136	9,754	8,530	7,253

The sample includes all schools that report ECE scores during 2007-2016 and that had not gained internet access by 2020. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Standard errors are clustered by school. Significance levels denoted by: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table A.9: Effects of Concurrent Computer Investments: Back of the Envelope Calculations

Math

Time Relative to Internet Access (1)	Total Predicted Rise in Computers per Student (2)	Total Dynamic Effect of Computers on Scores (3)	Estimated Effect of Internet Access on Scores (4)	Share of Total Internet Effect Explained by Teachers (%) (5)
t=0	0.008	0.00013	0.014	0.91%
t=1	0.006	0.00007	0.028	0.26%
t=2	0.004	0.00026	0.057	0.46%
t=3	0.002	0.00023	0.082	0.28%
t=4	0.000	-0.00002	0.089	-0.02%
t=5	-0.002	0.00021	0.111	0.19%
t=6	-0.004	0.00016	0.143	0.11%

Reading

Time Relative to Internet Access (6)	Total Predicted Rise in Computers per Student (7)	Total Dynamic Effect of Computers on Scores (8)	Estimated Effect of Internet Access on Scores (9)	Share of Total Internet Effect Explained by Computers (%) (10)
t=0	0.008	0.00002	0.006	0.40%
t=1	0.006	0.00002	0.018	0.10%
t=2	0.004	0.00021	0.037	0.57%
t=3	0.002	0.00027	0.042	0.64%
t=4	0.000	0.00001	0.052	0.02%
t=5	-0.002	0.00028	0.063	0.44%
t=6	-0.004	0.00029	0.078	0.37%

Columns 2 and 7 give the total predicted rise in computers in each period using the parameter estimates from the trend break regression of resources per student on Post-internet access, event time, and the interaction between the two (from Table 3). Columns 3 and 8 calculate the total dynamic effect of computers as of time t using (i) the largest positive parameter values from regressing ECE scores on computers per students and lags using the sample of non-internet connected schools (from Appendix Table A.8), without regard to significance level and (ii) the total predicted rise in computers from columns 2 and 7. Columns 4 and 9 display estimated effects of internet access on test scores from the baseline event study specification (Appendix Table A.3). Columns 5 and 10 express the total effect of computers as a percent of the total effect of internet access (column 3 (8) divided by column 4 (9)).

Table A.10: Restricting the Sample to Likely *Plan Huascarán* Compliers: Trend Break Results

	Dependent Variable: Standardized Test Score					
	Math			Reading		
	Baseline Sample	Only Schools in Complier UGELs	Only Schools in Localities w/o Other Sources of Internet	Baseline Sample	Only Schools in Complier UGELs	Only Schools in Localities w/o Other Sources of Internet
	(1)	(2)	(3)	(4)	(5)	(6)
Post-internet Access	0.014	0.022*	0.027	0.015	0.022**	0.027*
	(0.011)	(0.013)	(0.017)	(0.009)	(0.011)	(0.015)
Post-internet Access X Event Time	0.013**	0.015**	0.013	0.011**	0.013**	0.014*
	(0.005)	(0.006)	(0.009)	(0.004)	(0.005)	(0.008)
p-value of Test of Joint Significance for Post and PostXEventTime:	0.0520	0.0530	0.231	0.0620	0.0390	0.141
Observations	2,253,852	1,600,349	935,414	2,252,368	1,599,131	934,491
Number of schools	23,318	16,055	18,444	23,320	16,056	18,446

Columns 1 and 4 reproduce the baseline results using all public schools that gained internet access between 2007 and 2020 or that remained unconnected by 2020. Columns 2 and 5 present results excludes all schools in UGELs where more public primary schools become connected than allowed under the official quota under *Plan Huascarán*. We calculate the quotas as follows: first, we obtain annual quotas per UGEL from the Ministry of Education (published in 2004); note that these are only available for 206 out of the 219 UGELs. Second, we multiply the annual quota by 11 (to reflect 11 years between when the quotas were published in 2004 and the end of our sample period, 2014) and then by 0.7 (to reflect that 50% of the quota was for primary (only) schools and 20% was for integrated (primary and secondary) schools). Columns 3 and 6 present results excluding schools in areas that had an alternate source of internet access (i.e. a cyber cafe) prior to schools gaining access. Standard errors are clustered by school. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Post-internet access is a dummy variable for whether a school has gained internet access (i.e. $t >= 0$). Event time is years relative to internet access. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Significance levels denoted by: *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$.

Table A.11: Ease of Conducting Teaching Activities and School-based Internet Access

	Reported Ease of Teaching Activities (1=Very Difficult, 6 = Very Easy)					
	Communicating w/ & Motivating Students (1)	Selecting & Making Good Use of Methods & Materials (2)	Using Class Time Effectively (3)	Teaching at Different Learning Levels (4)	Addressing Students' Academic Problems (5)	Overall Activity Index (Average of 1-5) (6)
School has internet	0.001	0.113*	0.110*	0.063	0.149**	0.087*
in good condition	(0.054)	(0.059)	(0.061)	(0.066)	(0.064)	(0.045)
School does not have internet	0.000	0.037	0.079	0.034	0.022	0.035
in good condition	(0.066)	(0.073)	(0.075)	(0.081)	(0.078)	(0.055)
Observations	3,474	3,474	3,474	3,474	3,474	3,474
Mean of dependent variable	4.975	4.644	4.640	4.311	4.442	4.602
Std. dev. of dependent variable	1.061	1.104	1.155	1.263	1.215	0.883
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Teacher-level controls	Yes	Yes	Yes	Yes	Yes	Yes

The sample includes teachers in public primary schools that appear in the 2014 ENDO. "School has internet in good condition" takes the value of 1 if a school has internet access in "good condition" and is zero for all schools that do not have access or that have access in poor condition. "School does not have internet in good condition" takes the value of 1 if a school has internet access that is not in "good condition". Teacher- and school-level control variables include sex, age, experience, fixed effects for education level, and a dummy variable for urban status. Significance levels denoted by: *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$.

Table A.12: Heterogeneity in the Impact of Internet Access across Schools by Student to Teacher Ratios and Teacher Qualifications: Trend Break Results

	Dependent Variable: Standardized Test Scores			
	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)
Post-internet Access	0.000	0.003	0.017	0.023*
	(0.016)	(0.013)	(0.015)	(0.012)
Post-internet Access	0.006	0.007	0.008	0.013**
× Event Time	(0.008)	(0.007)	(0.007)	(0.006)
Post-internet Access	0.021	0.019		
× High STR	(0.022)	(0.018)		
Post-internet Access	0.005	0.002		
× Event Time × High STR	(0.011)	(0.009)		
Post-internet Access			-0.008	-0.017
× Low Teach. Qual.			(0.022)	(0.018)
Post-internet Access			0.005	-0.007
× Event Time × Low Teach. Qual.			(0.011)	(0.009)
	Effects in High STR Schools		Effects in Schools with Low Teacher Qualifications	
...for Post-internet Access	0.021	0.022	0.009	0.006
	[p-val.=0.159]	[p-val.=0.0814]	[p-val.=0.568]	[p-val.=0.709]
...for Post-internet Access × Event Time	0.011	0.009	0.013	0.006
	[p-val.=0.148]	[p-val.=0.170]	[p-val.=0.138]	[p-val.=0.409]
Observations	1,852,351	1,851,980	1,852,351	1,847,358
Number of Schools	22,024	22,027	22,024	21,984

High and Low STR schools are defined based on each school's pre-internet average ratio of grade 2 students to total teachers in the school (STR) over the sample period relative to the median of all schools' sample averages (including schools that remained unconnected by 2020). High and Low Teacher Quality schools are defined based on each school's pre-internet average number of teachers with a pedagogical or higher education degree per student over the sample period relative to the median of all schools' sample averages (including schools that remain unconnected by 2020). Note that degree information is not available for all schools. The samples include students from all public schools that gain internet between 2007 and 2020 or remain unconnected by 2020. or that remain unconnected by 2017. Standard errors are clustered by school. Scores are standardized within each calendar year to have mean zero and standard deviation of one across the universe of test takers in public schools. Post-internet access is a dummy variable for whether a school has gained internet access (i.e. $t >= 0$). Event time is years relative to internet access. Control variables include sex, class size, indicator variables for the number of 2nd grade classes at the school, total school enrollment, number of second grade students that took the ECE, facilities (computer room, library, administrative offices), resources per student (classrooms, computers, and teachers), UGEL-specific enrollment tercile by year fixed effects, and school fixed effects. Significance levels denoted by: *** p< 0.01, ** p< 0.05, * p< 0.1.